UC San Diego
UC San Diego Electronic Theses and Dissertations

Title

Fluorescent Puromycin Derivatives Containing Functionalized Adenosine Analogues—Synthesis, Photophysics, and Biological Applications

Permalink

https://escholarship.org/uc/item/7j36j5s2

Author

Hadidi, Kaivin

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Fluorescent Puromycin Derivatives Containing Functionalized Adenosine Analogues—Synthesis, Photophysics, and Biological Applications

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Chemistry

by

Kaivin Hadidi

Committee in charge:

Professor Yitzhak Tor, Chair
Professor Thomas Hermann
Professor Simpson Joseph
Professor Ulrich Muller
Professor Victor Nizet

2022

The dissertation of Kaivin Hadidi is approved and it is acceptable in quality and form

for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

I dedicate this thesis to my family and friends.

EPIGRAPH

"Nobody ever figures out what life is all about, and it doesn't matter. Explore the world. Nearly everything is really interesting if you look into it deeply enough."

-Richard P. Feynman

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to give my sincerest thanks to Dr. Yitzhak Tor, for truly turning me into the scientist and scholar I am today. My entire academic career was built through his guidance, and I sincerely hope I can provide justice for that beyond my time in UCSD. Words cannot describe how indebted I feel for bringing a young undergrad with limited knowledge into the lab and turning him into a true intellectual—it has forever changed my life.

To my committee members, who may or may not read this, thank you for being so open to discussion and mentorship. Especially in times like these, where pandemics are surging and wars are raging, it was truly nice to have the support of the faculty aiding me despite everything going on.

I would also like to thank Dr. Ezekiel Wexselblatt, Dr. Andrea Fin, and Dr. Alex Rovira for their outstanding mentorship and guidance, even beyond their time at UCSD. They have been amazing people to turn to in times of need, or when you just want company for a drink.

During my time at UCSD, I have seen many iterations of Tor lab members come and go. To the old lab members, I cherish the memories I have working with you in the lab-all the late nights, the jokes, the outings, and the companionship. I hope your endeavors have been fruitful, and I hope you all know deeply you have impacted me-both professionally and personally. To those that are still there, I will miss you guys. I know things seem difficult and stressful, but one step after another, and before you know it time has flown.

To the Esko and Godula lab members, I sincerely thank you for welcoming me into your disciplines. You have all been amazing people to work with, and I truly grateful for your expertise, assistance, and friendship. I wish you all the best in the future, wherever that may take you.

I've said this before, and I'll say it again. I would like to thank my family and my brothers, for their loyalty and support in me, even when I didn't have faith in myself. You are the reason I've come so far, and why I keep on going.

VITA

2015 Bachelor of Science in Chemistry, University of California San Diego

2017 Master of Science in Chemistry, University of California San Diego

2022 Doctor of Philosophy in Chemistry, University of California San Diego

PUBLICATIONS

Hadidi, K.; Tor, Y.; Azetidines-Containing Fluorescent Purine Analogs: Synthesis and Photophysical Properties. *Chem. Eur. J.,* **2022**, accepted.

Ortiz-Rodríguez, L.; Fang, Y.; Niogret, G.; Hadidi, K.; Hoehn, Sean; Folkwein, H.; Jockusch, S.; Tor, Y.; Cui, G.; Levi, L.; Crespo-Hernández, C.; Development of Thieno[3,4-*d*]pyrimidin-4(1H)-thione: An Effective Heavy-Atom-Free Photosensitizer for Cancer Cells Independent of the Oxygenation Status. *J. Am. Chem. Soc.,* **2022**, submitted.

Hadidi, K.; Bellucci, M. C.; Dall'Angelo, S.; Leeson-Payne, A.; Rochford, J. J.; Esko, J. D.; Tor, Y.; Volonterio, A. Guanidinoneomycin-Maleimide Molecular Transporter: Synthesis, Chemistry and Cellular Uptake. *Org. Biomol. Chem.* **2021**, *19* (29), 6513–6520.

Hadidi, K.; Wexselblatt, E.; Esko, J. D.; Tor, Y. Cellular Uptake of Modified Aminoglycosides. *The Journal of Antibiotics* **2018**, *71* (1), 142–145.

ABSTRACT OF THE DISSERTATION

Fluorescent Puromycin Derivatives Containing Functionalized Adenosine Analogues—
Synthesis, Photophysics, and Biological Applications

by

Kaivin Hadidi

Doctor of Philosophy in Chemistry

University of California San Diego, 2022

Professor Yitzhak Tor, Chair

The synthesis, photophysical analysis, and biological investigation of puromycin analogues containing fluorescent adenosine derivatives are discussed. We analyzed the photophysical properties of two novel fluorophores, *N,N*-dimethylthieno[3,4-*d*]pyrimidin-4-amine and isothiazolo[4,3-*d*]pyrimidine-7(6H)-thione. While the photophysical properties of these two compounds were found to be lackluster, substitution of the dimethylamine substituents for modified and unmodified azetidines were found to significantly augment the brightness of the

fluorophores by nearly three orders of magnitude in the thienopurines and two orders of magnitude in the isothiazolo derivatives.

We then synthesized C-nucleoside analogues of puromycin employing a thieno[3,4-*d*]pyrimidine heterocycle in place of the native purine, along with other derivatives employing azetidine and 2,2-difluoroazetidine at the 6 position (following native purine numbering). A convergent synthesis utilizing lithium halogen exchange between a halogenated thiophene-based purine analogue and a benzyl protected 3'-deoxy-3'-azido-ribolactone produced a desired C-glycoside with reactive handles at the 3'position of the sugar (for peptide coupling) and at the 6-position of the modified purine (for S_nAr reactions) providing synthetic accessibility to numerous fluorescent antibiotic scaffolds.

Biological evaluation of the synthesized antibiotics indicates similar bioactivity of the novel fluorophores compared to native puromycin, both in *in vitro* protein expression systems and bactericidal assays. Treatment of the modified derivatives in HEK293T cells indicated localization of the compounds with ribosomes following fixation immunofluorescence treatments. One derivative, the 2,2-difluoroazetidine modified analogue, exhibited fluorescence in live cells and neurons at sub-cytotoxic concentrations, one of the first and few times a fluorescent nucleoside derivative has been used in this context given the effect of size restriction on the chromophores in question.

CHAPTER 1: Introduction

Cellular machinery is perhaps one of, if not the most complex system ever to be studied and deconvoluted by human scientific prowess. Ingenuity powering scientific discoveries and innovations focusing on nature's own design has propelled fields such as medicine, agriculture, and cosmetics benefiting society tremendously. Within the cellular machinery, three classes of biomacromolecules comprise highly significant and diverse functions necessary for life: deoxyribonucleic acid (DNA), ribonucleicacid (RNA), and proteins. In this work, we will primarily focus on RNA and proteins in the context of neuronal signaling and extracellular response.

1.1 Introduction to proteins and protein synthesis

Proteins are a class of biomacromolecules commonly referred to as the workhorses responsible for the myriad of intra- and -extracellular functions. Their biochemical versatility is the main reason behind their involvement in nearly every cellular process. Such range of adaptability is a unique result of amino acids, the residues that comprise proteins and the structural diversity within them.[1] More specifically, amino acids are organic compounds that contain carboxylate and α-amino functional groups, along with side chains consisting of various other chemical functionalities including carboxylic acids, indoles, imidazoles, phenols, and more. Twenty canonical amino acids, both essential and non-essential (figure 1.1) along with other non-cannonical structures,[2] encompass this group of building blocks that are stitched together through amide linkages to produce fully functional proteins. The unique sequence of amino acids, along with the functional group variety of amino acid side chains, is responsible for the various secondary and tertiary structures observed and further leads to biochemical function. Metabolism, cell signaling, cell division, DNA replication and transcription, as well as the

production and processing of new proteins are among a few significant processes modulated through their unique chemistries.[1]
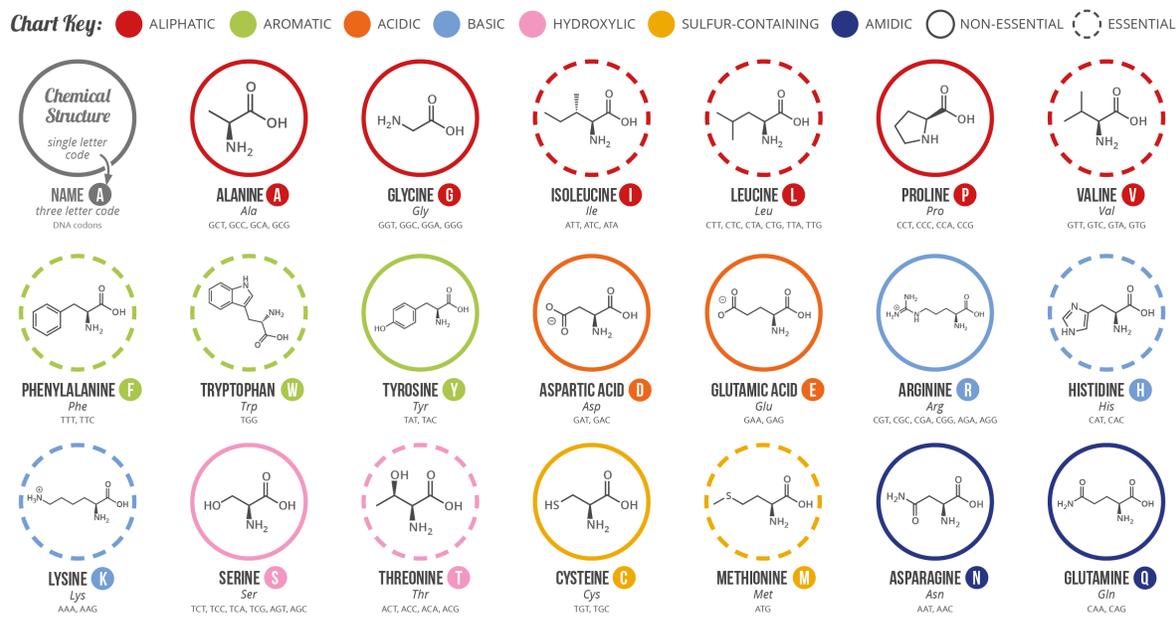
Figure 1.1. The twenty common amino acids, as illustrated by CompoundChem.[3]

Given the significance of proteins, similar importance is therefore given to the process(es) responsible for producing them. The process of translation, referring to the synthesis of functional proteins through the stitching of amino acids directed by mRNA codons, represents one of the most fundamental processes in cellular biology, requires extensive biological machinery and requires subtle yet sophisticated regulation.[2] Such regulation itself is affected spatiotemporally by multiple factors, including various stimuli, gene expression, and more.[3] A majority of resources within cells and organisms is dedicated to the production and modulation of proteins.[4]

To note, the field of protein synthesis is incredibly rich and diverse, with applications spanning medicine, agriculture, enzymology, and more. For the sake of this work, we will primarily focus on a general view of eukaryotic translation (figure 1.2).[4] For more in-depth

information on this process, see the following references.[5–10] The main component responsible for catalyzing this process, the ribosome, consists of a massive ribonucleic acid (RNA) and protein hybrid that acts as a ribozyme, meaning that the RNA components play more important roles than the proteins. Such discoveries constitute historical hallmarks of cellular biology, including a Nobel prize in chemistry for the crystal structure of the ribosome in 2009.[11] Intriguingly, these discoveries appear to support the idea that life evolved through what's termed an "RNA world", a hypothetical stage in Earth's evolutionary history containing RNA molecules that proliferated before the main expression system of the genetic code, deoxyribonucleic acid (DNA), and proteins emerged.[12]

1.1.1 Translation
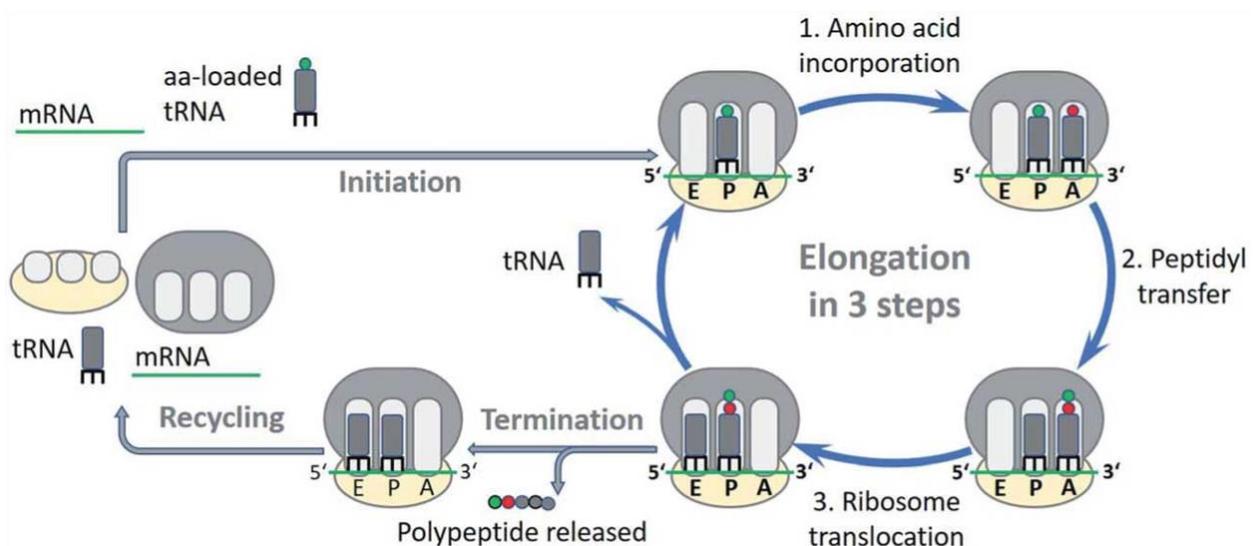
Figure 1.2. Schematic representation of eukaryotic translation.[4]

A series of biomolecules participate in protein synthesis – ribosomes, messenger RNA and transfer RNA. mRNA is a single-stranded form of RNA transcribed from DNA that is post transcriptionally modified through events such as 5'-capping, polyadenylation, and splicing leading to the expression functional genetic code.[13] Transfer RNAs (tRNAs) are small RNAs that

are responsible for shuttling amino acids with the aid of several protein elongation factors such as eIF2 and EF-Tu that catalyze the binding of aminoacyl tRNAs to ribosomes.[14,15] tRNAs are also heavily modified. The individual nucleotides possess additional chemical modifications making it one of the most modified biomolecules in cells.[16]

Translation is divided into three phases: initiation, elongation, and termination.[2] Translation initiation encompasses the process by which the ribosome and its protein factors bind an mRNA at a site called the start codon, which consists of an AUG codon.[5] A majority of protein synthesis is termed cap-dependent and requires the presence of a 5'cap and a sequence of nucleotides dubbed the 5' untranslated region (UTR) directly upstream from the start codon for assembly.[5] A set of proteins called eukaryotic initiation factors (eIFs) then aid in the stabilization of the ribosomal preinitiation complex, consisting of the small ribosomal subunit with Methionine-charged initiator tRNA (Met-tRNA$_i^{Met}$) in the P site, onto the 5' cap, which then relocates to the start codon, when the 60S ribosomal subunits binds and releases the initiation factors to form the fully functional 80S ribosomal complex.[17] While most protein synthesis occurs through this process, in times of cellular stress cap-independent translation can initiate translation without the need for a 5'cap.[18] The best example of this event is through the internal ribosome entry site (IRES) an RNA element commonly found in many RNA viruses.[19]

Upon formation of the 80S ribosomal subunit, the process of elongation works to extend the newly growing peptide.[4,20–22] With the aid of several eukaryotic elongation factors, tRNAs charged with amino acids (aminoacyl tRNAs) are shuttled to the A-site of the ribosome. Upon binding of the tRNA, a verification process is initiated through the base pairing of the active mRNA codon with the anticodon of the tRNA stem loop. Mismatched pairs lead to ejection of the tRNA from the A-site, while correct base-pairing and wobble pairs trigger peptide

4

bond formation between the α-amino group of the A-site aminoacyl tRNA and the 3' ester linkage CCA end of P-site peptidyl-tRNA. The reaction triggers the transfer of the extended peptide to the A-site aminoacyl-tRNA, causing the ejection of P-site tRNA and translocation of A-site peptidyl tRNA to the P-site. The newly emptied A-site is free to bind new aminoacyl-tRNAs and repeat the process.

Upon reaching the end of the mRNA reading frame, the ribosome eventually reaches a stop codon, designated by UAA, UAG, or UGA, prompting eukaryotic release factors (eRFs) to bind the stop codon causing ribosome disassembly and full-length polypeptide release. The proteins are then shuttled to the golgi apparatus for any post-translational modifications necessary for them to assume their proper functions.[2,10]

1.1.2 The proteome and localized protein synthesis
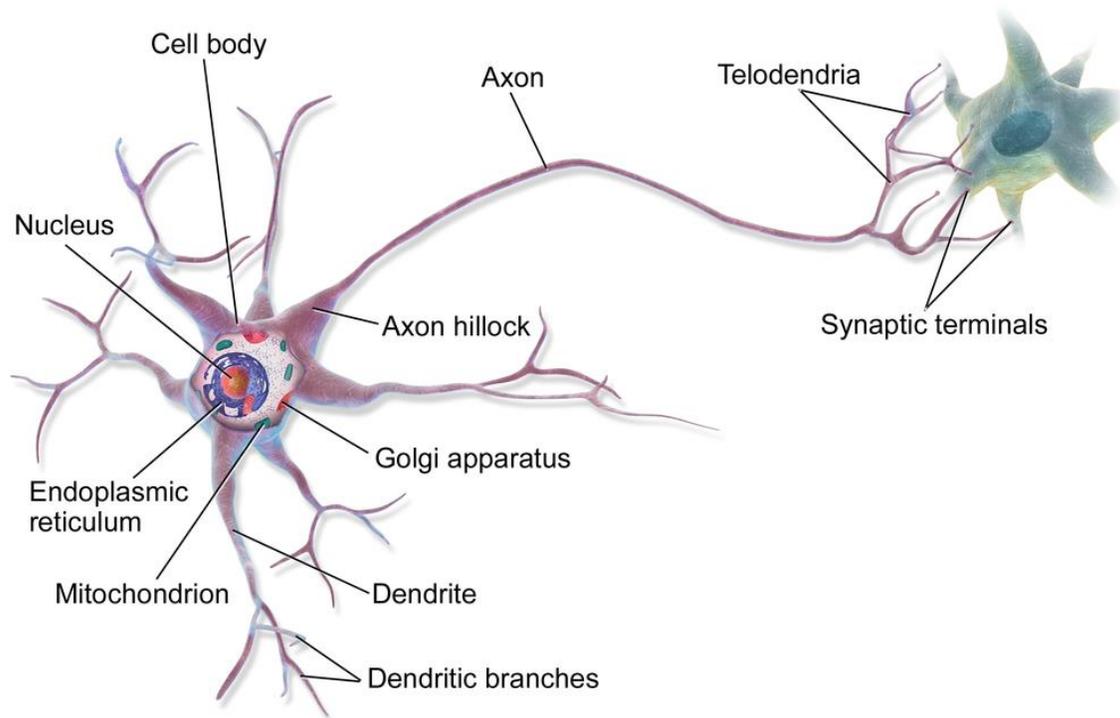
Figure 1.3. Neuron physiology.[23]
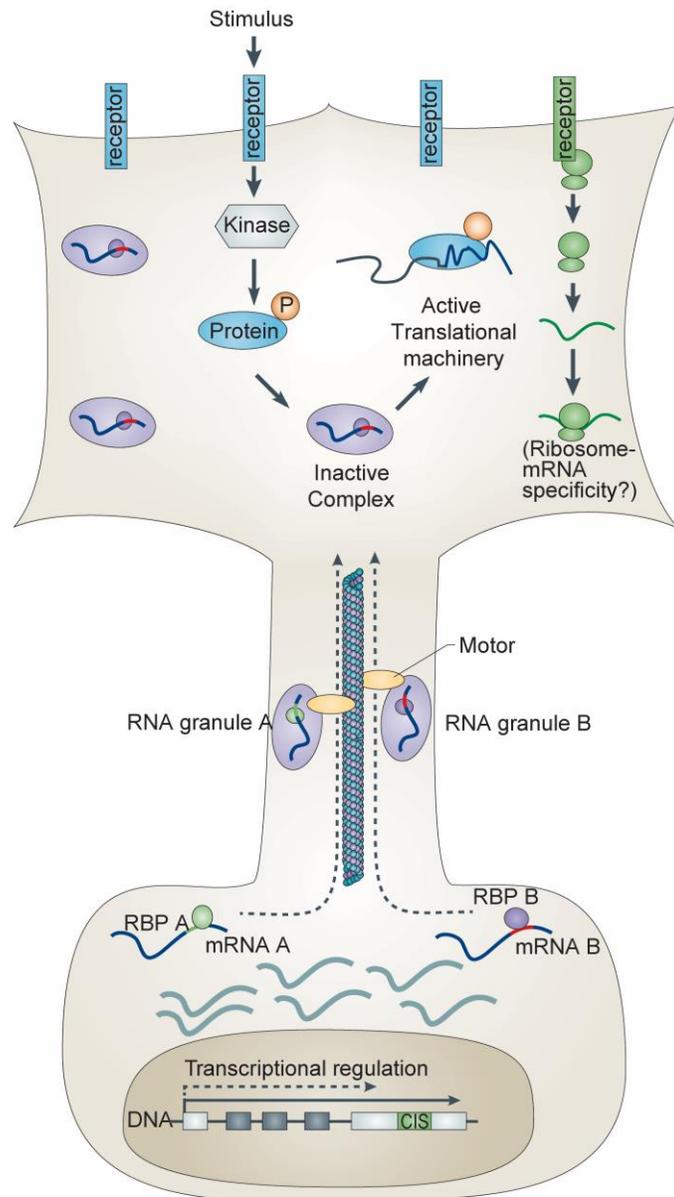
Figure 1.4. Translation of prepackaged RNA granules at axonal and dendritic terminals.[24]

Protein synthesis and processing events result in the formation of active proteins that constitute a cell's proteome, which is defined as the entire set of proteins expressed by a genome within a cell. The flexible nature of the proteome is a result of both *de novo* synthesis and degradation of undesirable proteins, allowing cells to maintain homeostasis and control several important factors such cell proliferation, growth, and development.[25] Additionally, proteomic

changes allow for rapid adaptation to environmental changes over time, such as cell-cell signaling and extracellular stimuli.[26] Given that the vast majority of eukaryotic mRNAs possess half-lives over 2 hours, swift adjustments to the proteins encoded by these mRNAs must be done at the translational level.[27]

Such rapid responses through protein synthesis are especially significant within axons and dendrites of neuronal cells, which receive continuous guidance information from their environment.[24] However, neurons differ from various classical cell types given the unique morphology they possess, exuding branches from the main cell body spanning large areas in an attempt to form circuitry through connection with multiple synaptic targets (figure 1.3). However, in the context of protein synthesis and synaptic response, the expression of proteins necessary for the cellular response to extracellular stimuli cannot be feasibly performed at the cell body due to the extensive distance required to transverse axons and dendrites and would theoretically cause sluggish response times.[27] To circumvent this, neurons prepackage their translational materials, such as ribosomes, mRNA, and assistor proteins into bundles termed "RNA granules". These packages are then shuttled to the sites of synaptic response within axons and dendrites, where they remain translationally dormant until the neuron receives an extracellular signal, causing a secondary messenger cascade that ultimately actives the translational material of interest (figure 1.4).[24] This general observation has been referred to as "local protein synthesis" and provides several advantages over the transport of preexisting proteins from one part of the cell to another. First, translationally silent forms of mRNAs can be stored locally and used to make many copies of a protein when needed, thereby providing economic advantages. Second, the ectopic presence of proteins in other parts of the cell during protein transport is avoided. Third, mRNAs can be targeted to different subcellular localizations

using 'address' information in their untranslated regions (UTRs) without changing the structure

and function of the proteins they encode.[27]

1.1.3 Proteomic detection and puromycin

Given that a neuron's proteome at axonal and dendritic sites potentially correlate to the

various stimuli received by the cell, proteomic analysis of expressed axonal and dendritic

proteins is paramount towards understanding how neurons respond to unique stimuli.[28] Several

methods have been developed in practice, including metabolic labeling through the incorporation

of modified amino acids (ex. Azidohomoalanine), translating ribosome affinity purification

(TRAP) utilizing stalled ribosomes, and puromycin tagging.[28] Specifically, this work will focus

solely on puromycin tagging, which utilizes the antibiotic puromycin, which incorporates into

nascent polypeptides. For the sake of this work, we will focus solely on puromycin treatment.
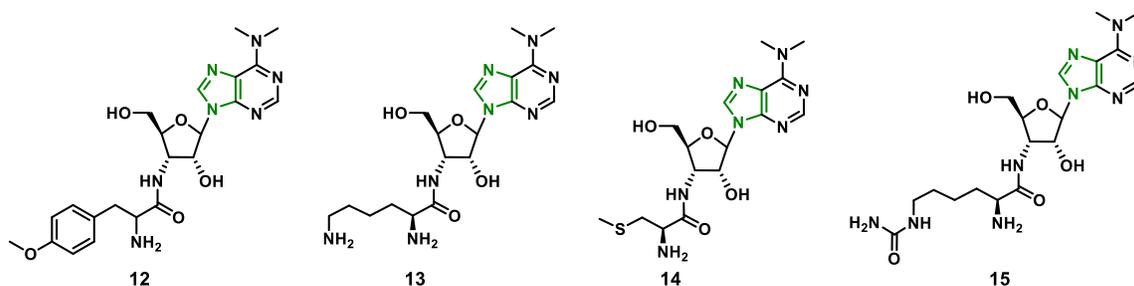
Figure 1.5. Chemical structure of puromycin (**12**) and similar antibiotics (lysylaminoadenosine **13**, Cystocin **14**,[29] and homocitrullylaminoadenosine (**15**)[30] isolated from various strains of *streptomyces* bacteria.

Puromycin is a naturally occurring aminonucleoside antibiotic isolated from

Streptomyces Alboniger that inhibits protein synthesis through incorporation into the C-terminal

end of elongating nascent chains (figure 1.5). The incorporation of the antibiotic results in

preventing elongation and prematurely terminates translation. Due to this unique mechanism of

action, its use is highly varied and invaluable, with examples including a selection marker for

cell lines genetically engineered to express resistance, and as a probe for protein synthesis,

ranging from purified ribosomes and cell-free translation to cultured cells and animals (when treated below the corersponding IC$_{50}$ value).[31]

Structurally, puromycin is comprised of a N^6,N^6 dimethylated adenosine covalently bound to an amino acid, mimicking the final adenosine residue sitting at the 3' end of aminoacylated tRNAs that participate in the delivery of amino acids to elongating ribosomes.[32,33] Importantly, both moieties are able to tolerate chemical substitutions and modifications without exhibiting significant loss of activity.[31,34] Due to this property, a diverse toolbox of puromycin-based reagents with added functionality, such as biotin for affinity purification or fluorophores for fluorescent microscopy detection have been developed.[31] These analogues have been crucial in the understanding and deconvolution of translation-dependent processes. However, it is important to note that most fluorescent analogues produced rely on the photophysical prowess of conventional fluorophores used in molecular biology, such as cyanine and fluorescein dyes. Such dyes are prominent for localization studies yet provide little insight into the environmental activities occurring around the probes. Additionally, many of these analogues mentioned are not compatible with real-time live-cell analysis, given that subsequent treatments are necessary, the additional reagents are toxic, or both. *It is for this reason, we opted to modify the puromycin skeleton to create fluorescent analogues of puromycin that are bright, responsive to the microenvironment, and require no additional treatment to provide real-time information about proteins being expressed in the cellular environment.* We opted to do this using fluorescent nucleosides, which possess a rich history of the aforementioned qualities,[35] and produce emissive puromycin derivatives employing fluorescent adenine surrogates.

1.2 Introduction to nucleic acids and their constituents

The field of fluorescent nucleosides is rich, and one cannot accurately assess the advantages and disadvantages of the field without first examining their inspiration, nucleic acids. Among the plethora of biomolecules that exist in living things, nucleic acids are a family of biomacromolecules that perform key biological functions not limited to catalysis, gene expression, protein synthesis, secondary messengers and more. Within this diverse family of small- and macro- molecules, the oligonucleotides deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are composed of monomeric constituents termed nucleosides that are assembled into the polyelectrolytes through linking their triphosphorylated derivatives called nucleotides.[2] Four nucleosides constitute the building blocks of DNA and RNA and can be structurally identified through the covalent linkage between a sugar residue and a purine or pyrimidine heterocycle. Three nucleosides are shared between DNA and RNA: adenosine (A or dA), guanosine (G or dG) and cytidine (C or dC), after accounting for a primary structural difference in the absence or presence of a key alcohol on the 2' carbon of the ribose sugar ring.[2] The remaining two nucleosides, thymidine (dT) and uridine (U) are native to DNA and RNA respectfully. Polymerase-mediated assembly of these nucleotides into oligonucleotides proceeds through the polymerization of the respective building blocks with phosphodiester bonds linking the 3' and 5' alcohols. For the sake of this thesis, we will primarily focus on ribonucleosides and their analogues.

While DNA is generally known for the storage of genetic information, RNA possesses an impressive variety of functional variability. The source of this variety is thought to reside in its structural flexibility, base-pairing ability, and stability.[36] To begin, RNA can exhibit many secondary structures. Unlike DNA, which usually exists in the duplex form, RNA can be found

in both the single strand and duplex form, as well as other exotic structures such as hairpin loops, symmetric and asymmetric internal loops, three and four stem junctions, etc. which restrict or allow access to certain sites and residues on the RNA motif, regulating processes such as transcription, splicing, catalysis, and others through steric constraints.[37] The primary molecular interactions that cause the spontaneous formation of these structures are base pairing and base stacking, including both the canonical Watson-Crick base pairing[38] as well as non-cannonical base pairs such as Hoogsteen and wobble pairs.[39] A key example of this point is the molecular recognition used in translation, in which the codon inscribed in messenger RNA wobble pairs with the corresponding anticodon in tRNA, bringing an amino acid attached to the CCA end of the tRNA into the ribosome.[2] The wobble pair of the codon-anticodon interaction is weak, allowing for mistakes that could lead to mutations and subsequently introduce the ability for living things to undergo natural selection.

Different types of RNA perform different functions, and their function is inherently tied to their stability.[36] Naked single-stranded RNA is notoriously unstable due to the close proximity of the 2' hydroxyl to the 3' phosphodiester linkage, a useful degratory tool in messenger RNAs (mRNAs) transcribed upon introduction of a stimuli, facilitating quick biochemical responses. For mRNAs encoding proteins vital for cellular function that require longer half-life's, stabilization against degradation is performed by masking the end with the 5' cap and poly A tail, protecting the ends of the oligonucleotide from degratory enzymes.[40]

The different classes of RNA include two examples of classically longer oligonucleotides, ribosomal RNA (rRNA), and transfer RNA (tRNA). Intriguingly, both rRNA and tRNA possess nucleosides containing additional structural modifications such as methylation, formylation, glycosylation, and many other exotic examples, adding an additional

dimension to the relationship between RNA and their respective functions.[41,42] Shorter length RNA oligonucleotides also participate in significant cellular functions and include small nuclear RNA (snRNA) that aid in splicing, small interfering RNAs (siRNAs) that regulate gene expression, and non-coding RNAs (ncRNAs) whose function remains unclear.[43]

The field of RNA therapeutics has evolved tremendously over the past two decades. RNA nucleoside analogues have historically been utilized as antivirals, and whilst small molecule drugs have typically filled this space,[44] RNA oligonucleotide-based therapies are now being investigated as well given modern synthetic methodologies can more readily produce longer oligonucleotides at larger scales.[45] The most relevant example of this technology are the COVID-19 vaccines produced by pharmaceutical companies Pfizer and Moderna employing mRNAs that incorporate modified nucleosides.[46,47] Additionally, the technology is also being expanded to address numerous other diseases including cancer, highlighting the utility of RNA oligonucleotides as a therapeutic modality in addition to its classical biological relevance.

Given the significance of RNA, its functions, uses, modifications, and biological interactions, numerous biophysical assays and techniques have been developed to study these factors. The main concept of these assays focuses on attempting to monitor a unique RNA nucleoside, nucleotide, or oligonucleotide strand as it undergoes either chemical modification (e.g., methylation or depurination) or environmental changes (e.g., protein binding, duplex formation, and folding). The techniques used to monitor these changes include but are not limited to nuclear magnetic resonance (NMR), electron paramagnetic resonance (EPR), X-ray crystallography, UV-Vis spectroscopy and fluorescence, isotopic labelling, and circular dichroism (CD).[35]

Biological interactions and transformations (not limited to nucleosides and oligonucleotides) proceed through various transition states and kinetic reaction rates, highlighting the significance of employing real-time detection methods capable of providing information for these processes as they occur *in vitro*. While this may appear to be obvious at first glance, many techniques simply do not possess the capacity to rapidly measure a system repeatedly and are restricted to providing specific snapshots at a certain point in time. This presents a huge disadvantage in biological contexts given the sheer number of processes occurring and the potential sensitivity to change. Regarding nucleosides, a simple technique to monitor small molecule nucleosides or oligonucleotides in real-time is UV-Vis spectroscopy.[35] The heterocyclic cores of the native nucleosides are comprised of purine and pyrimidines, which absorb UV light at 260 nm. However, many other biologically relevant molecules (e.g., amino acids) absorb in this region as well, convoluting measurements without additional modifications. Therefore, nucleosides have been chemically modified to bypass such obstacles presented when attempting to track and measure the molecules in an experimental framework.

Traditionally researchers have utilized radioactive isotopes for tagging and detecting nucleosides and nucleic acids. ^{14}C, ^{3}H, ^{35}S, and ^{32}P have all been incorporated into specific segments of the (either native or modified) nucleic acid skeleton such as the phosphate backbone, nucleobase, or sugar.[48] There are significant benefits associated with the technique, the most notable being high sensitivity along with retaining native chemical structure and biological activity. The former is a highly desirable feature of any biochemical experiment – any structural deviation from the native nucleosides will affect the outcome depending on how perturbing the modifications are. Therefore, it is highly desirable to use probes that are both "isomorphic" and "isofunctional" – indicating that the probes have similar structural

resemblance and biological activity as the native nucleosides in a particular system. As alluded to previously, radioactive labelling utilizes isotopically labelled nucleosides, which are obviously both isomorphic and isofunctional since the core structure remains the same. However, radioactive labelling also comes with significant drawbacks, such as mediocre half-lifes of commonly utilized isotopes, cost, and inability of real-time analysis. For this reason, the most common biochemical methods utilized in the analysis of nucleosides and nucleotides currently invoke the use of fluorescence spectroscopy.[48,49]

Fluorescence detection through spectroscopy and microscopy is a powerful tool that is the basis of most biomedical and spectroscopic techniques including imaging, flow cytometry, cell sorting, DNA sequencing, single-molecule spectroscopy, anisotropy, fluorescence-lifetime imaging microscopy. It has revolutionized the field of molecular biology in multiple facets and employs not only small-molecule fluorophores but also fluorescent proteins capable of multiplexed imaging.[50] The sensitivity of certain fluorescent probes is so high detection at the single-molecule level can be achieved.[51] The advantages of fluorescence microscopy are unparalleled – high sensitivity, rapidity of measurement, real-time diagnostics, virtually nonexistent cellular invasiveness, and deep penetration into tissues with infrared light have pioneered biological data acquisition both *in vitro* and in animal models.[52] It is of no surprise that the elucidation of processes involving nucleosides and nucleotides has been performed through fluorescence spectroscopy as well. However, while certainly advantageous, there are significant obstacles in utilizing fluorescence spectroscopy with nucleosides. Notably, except for a small number of naturally occurring nucleosides, the canonical purines and pyrimidines are essentially non-emissive under physiologically relevant conditions.[35] Thus, chemists have synthesized isomorphic and isofunctional derivatives of the native nucleosides with the intent to instill

emissive properties for fluorescence-based studies.[53] Fluorescent labelling of nucleosides revolutionized nucleic acid-based experimentation and analysis through the covalent attachment of emissive chromophores.[54] Given that more heavily modified nucleosides with larger structures are more perturbing, chemists have developed and employed increasingly isomorphic fluorescent nucleosides to better parallel biochemical activity of native nucleosides in a particular system.

1.3 Fluorescent nucleoside analogues

1.3.1 Photophysical parameters

A major obstacle in the design of novel chromophores is the ability to accurately predict the photophysical properties of a molecule without experimentation. Although recent enhancements in molecular modeling software have significantly improved in this regard and can accurately calculate orbital energies, the ability to predict emissive properties under diverse conditions remains difficult.[55] Therefore, synthetic chemists have synthesized various compounds theorizing that they may possess desirable photophysical properties. However, in the context of synthetic nucleosides and nucleotides, it is crucial to minimize structural perturbations and subsequently deviations in biological activity in addition to the challenge of instilling emissive properties in the nucleoside scaffold.

The desirable photophysical properties in question are absorption and emission wavelengths (λ), along with molar absorptivity (or extinction coefficient, ε) and emission quantum yield (Φ). Ideally, the use of emission wavelengths in the far-red or infrared region are highly desirable as the light is lower in energy and subsequently produces lower phototoxicity. However, the small aromatic rings that encompass the nucleoside heterocycles makes this a difficult property to instill as red-shifted emission is usually a byproduct of increased delocalization in larger, more conjugated aromatic systems.[56]

Nevertheless, unless the emission of a fluorophore is of significant intensity, the absorption and emission maxima themselves are of little use. Classicly, the "brightness" of a fluorophore is defined as the product of the molar absorptivity (ε) and quantum yield (Φ). Beginning with the molar absorptivity, this value can be described as the number of photons absorbed by a given chromophore per molar concentration. Aside from this value, the actual absorbance of a sample in solution also depends on the concentration and pathlength of the sample being measured, described in the Beer-Lambert law:[56]

$$A = \varepsilon bc$$

Equation 1.1. The Beer-Lambert Equation.

where A is defined as the absorbance of a particular sample, ε is the molar absorptivity (L mol^{-1} cm^{-1}), b is the pathlength of the cuvette used to measure the sample (cm), and c is the concentration (M). The equation forms a linear relationship between A and c given that the path length and concentration remain constant.

To define quantum yield, the process of fluorescence must be analyzed as outlined in figure 1.6.[56,57] Three events take place for any given fluorophore to emit light which occur on time scales separated by orders of magnitude. Vertical excitation of a chromophore to the excited state by a photon of discreet quanta initiates the process and occurs rapidly in the femtosecond timescale (10^{-15}). This excitation follows the Franck-Condon principle, which dictates that vertical excitation occurs faster than the general timescale of nuclear motion.[56] Upon reaching the Franck-Condon state following direct excitation, nuclear motion occurs in the excited state, and the chromophore persists through an excited state lifetime τ_0 which may be defined as the sum of all radiative and non-radiative processes undergone by a molecule as it decays to the ground state.[56] During the excited state lifetime, internal conversion and vibrational relaxation

occur decreasing the energy of the excited chromophore within a particular energy state. This loss of energy can ultimately lead to an excited state that results in fluorescence upon relaxation, or an excited state that produces non-radiative decay, termed a "dark state".[56]
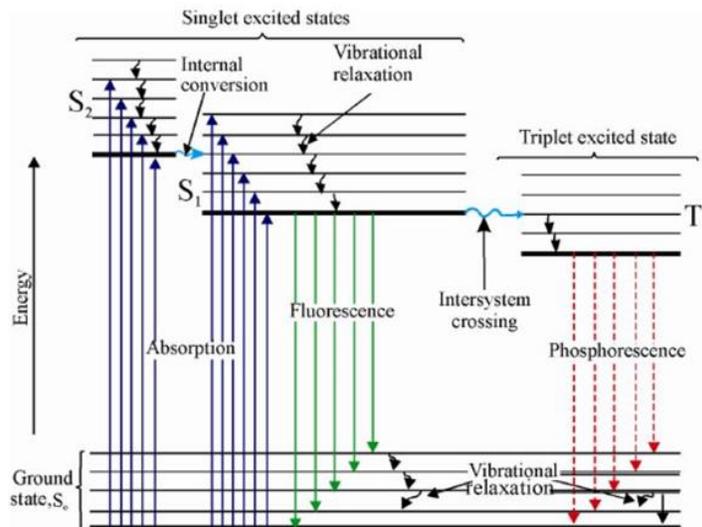
Figure 1.6. Jablonski diagram depicting electronic states and transitions.[57]

The quantum yield (Φ) of a fluorophore can therefore be defined as the ratio between the number of radiative processes within an excited chromophore τ over the total processes occurring (the excited state lifetime, τ_0), or the ratio between the number of photons emitted by a sample and the number of photons absorbed:[56]

$$\Phi = \frac{\tau}{\tau_0}$$

Equation 1.2. Relationship between quantum yield and excited state lifetime.

However, the quantum yield of a fluorophore can also be calculated using another fluorophore with known photophysical parameters:[56]

$$\Phi = \Phi_{STD} \frac{I}{I_{STD}} \frac{OD_{STD}}{OD} \frac{n^2}{n_{STD}^2}$$

Equation 1.3. Calculation of quantum yield through a known standard.

Where Φ_{STD} is the fluorescence quantum yield of the standard, I and I_{STD} are the areas under the emission curve of the sample and standard, OD and OD_{STD} are the optical density at a particular excitation wavelength for the sample and standard, and n and n_{STD} are the solvent refractive indexes of the sample and standard solutions.

Given that we are attempting to synthesize fluorophores with inspiration from non-emissive compounds, one needs to establish the specific photophysical properties necessary for a desired investigation. Ideally, the absorption and emission maxima of a fluorophore for *in vitro* or *in vivo* applications should exist around low energy, near- infrared region of light primarily to prevent photodegradation. "Red-shifted" light is also beneficial *in vivo* for deep-tissue penetration. However, in the context of isomorphic fluorescent nucleosides, red-shifted absorption and emission bands are difficult to produce compared to polycyclic aromatics due to lower conjugation in the purine skeleton. Nevertheless, while infrared-red may be difficult to utilize, an exclusive absorption band above 300 nm suffices as this avoids interference from other photoactive biologics, including the native nucleosides, nucleotides, nucleobases, and certain amino acids such as tryptophan.[35,53,56]

A second highly desirable photophysical property is the spectroscopic responsiveness of a probe towards changes in the microenvironment as a means of obtaining information about the processes occurring within close proximity. Usually, a particular fluorophore may respond to any individual or combination of four major factors: viscosity, pH, polarity, and solvation (also related to polarity). Photophysical sensitivity to these factors presents a unique advantage in biological contexts, as changes in intracellular and extracellular environments are known to exist/arise under unique circumstances. A couple recent example are found in 1. Lysosomes, acidic vesicles in cells responsible for cellular digestion,[58] and 2. liquid-liquid phase separated

droplets (dense solutions of macromolecules forming what resembles liquid drops in cells).[59] Fluorescent nucleosides found in these environments may show spectroscopic changes if they exhibit sensitivity to pH or polarity. This is a unique advantage compared to more classically used fluorophores in molecular biology such as rhodamine, cyanine dyes, fluorescein, etc. given that they are simply strong fluorophores in any solution and provide little information about the environment surrounding them.[60]

The last photophysical property discussed here will be regarding the brightness of the fluorophore(s). Brightness is simply defined as $\Phi\varepsilon$, or the product of the quantum yield and the molar extinction coefficient. Although high brightness values are not necessarily required for experimental analysis, chromophores with higher brightness produce more intense fluorescence are lower concentrations. This is especially useful in a cellular or *in vivo* context, where translocation of small molecules into cells or tissues may impede reaching high enough concentration inside the cell to obtain a fluorescent signal.

1.3.2 Chemical Interactions

To appropriately design isomorphic nucleosides (fluorescent or non-fluorescent), one needs to properly assess the chemical interactions present in the native structures for a faithful analogue to mimic. Two such intermolecular interactions are base-pairing and base-stacking.[61] To address the latter, base-stacking occurs through favorable π-π interactions of the purine and pyrimidine heterocycles, leading to reduced exposure of the more hydrophobic nucleobases to the aqueous environment.[61] Intriguingly, while base-stacking is more prevalent in DNA for duplex stability, the interaction has also been reported to mediate RNA-RNA interactions.[62]
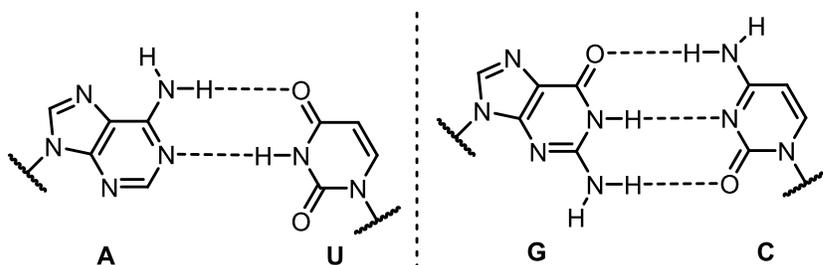
Figure 1.7. Watson-Crick base-pairing of nucleobases in RNA.

While base-stacking is responsible for duplex stability, base-pairing is responsible for the specificity of base interactions and occurs when two separate nucleobases hydrogen-bond with one another in an energetically favorable manor.[38,39] The interaction is the backbone of all nucleic acid secondary and tertiary structures, as well as a key molecular recognition system in the process of protein synthesis, where 3-nucleotide codon sequences encode specific amino acid residues.[2] Of the various types of base-pairing found in nature, Watson-Crick pairing is ubiquitous, and involves a purine nucleoside engaged in hydrogen bonding to a pyrimidine nucleoside.[38] Within DNA and RNA, the pairs observed are very similar – in DNA, dA base pairs with T and dG base pairs with dC, whilst A pairs with U and G pairs with C in RNA. These pairs are termed "complimentary" and hold great importance ensuring proper DNA replication and transcription. The origin for the selectivity of base-pairing resides in chemical functionality found within the specific purines and pyrimidines, leading to energetically favorable hydrogen bond donor-acceptor arrangements. Intriguingly, the hydrogen bond length between the two Watson-Crick pairs is approximately 1.5-2.5 Å, making the G-C and A-U base pairs nearly equidistant.[63] This physical property is the molecular basis of duplex stability. Mismatched base pairs disrupt duplexes through physical strain as the intrapair hydrogen-bond length contrasts that of A-U and G-C.[63]

1.4 Fluorescent nucleoside parameters in the context of puromycin

Given that puromycin is a functionalized nucleoside, it stands to reason that the parameters governing a faithful fluorescent analogue will mirror that of the fluorescent nucleosides incorporated into them. While this is true in some respects, significant constraints are also present within the molecular design described in this work-the primary point being only adenosine analogues are employed, barring the use of fluorescent analogues of other nucleobases. Additionally, while various other amino acids have been reported in other naturally-occuring puromycin-mimics[29,30], O-methyl-L-tyrosine will be the only amino-acid analogue employed given the significance of puromycin over other members of the same structural family. Lastly, the fluorescent analogues must behave in a comparable manner to native puromycin in order to accurately assess the spatiotemporal activity of the antibiotic upon incorporation into proteins.

Whilst significant constraints are present, several aspects of fluorescent nucleoside design are less impactful in the context of puromycin derivatization. Watson-Crick pairing, although crucial in faithful fluorescent nucleosides, is rendered virtually useless within this unique design space. Dimethylation of puromycin's heterocyclic amine completely blocks base-pairing, providing an additional degree of freedom as one can deviate from the classical donor-acceptor systems towards more exotic chemical structures. Modifications of the adenosine nucleobase are also generally well-tolerated as the primary source of puromycin's biochemical activity originates in the α-amino group, and various analogues of puromycin employing nucleobase modifications have been previously studied.[32,34,64,65] With these points considered, the ability for fluorescent nucleosides to act as a benign, faithful probe within the context of puromycin appears to be a reasonable hypothesis.

1.5 References

(1) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. Analyzing Protein Structure and Function. *Molecular Biology of the Cell. 4th edition* **2002**.

(2) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, 7th ed.; W.H. Freeman: Basingstoke, 2012.

(3) A Brief Guide to the Twenty Common Amino Acids. *Compound Interest*, 2014.

(4) Brönstrup, M.; Sasse, F. Natural Products Targeting the Elongation Phase of Eukaryotic Protein Biosynthesis. *Nat. Prod. Rep.* **2020**, *37* (6), 752–762. https://doi.org/10.1039/D0NP00011F.

(5) Jackson, R. J.; Hellen, C. U. T.; Pestova, T. V. The Mechanism of Eukaryotic Translation Initiation and Principles of Its Regulation. *Nat Rev Mol Cell Biol* **2010**, *11* (2), 113–127. https://doi.org/10.1038/nrm2838.

(6) Schuller, A. P.; Green, R. Roadblocks and Resolutions in Eukaryotic Translation. *Nat Rev Mol Cell Biol* **2018**, *19* (8), 526–541. https://doi.org/10.1038/s41580-018-0011-4.

(7) Kapp, L. D.; Lorsch, J. R. The Molecular Mechanics of Eukaryotic Translation. *Annual Review of Biochemistry* **2004**, *73* (1), 657–704. https://doi.org/10.1146/annurev.biochem.73.030403.080419.

(8) Andreev, D. E.; O'Connor, P. B. F.; Loughran, G.; Dmitriev, S. E.; Baranov, P. V.; Shatsky, I. N. Insights into the Mechanisms of Eukaryotic Translation Gained with Ribosome Profiling. *Nucleic Acids Research* **2017**, *45* (2), 513–526. https://doi.org/10.1093/nar/gkw1190.

(9) Sokabe, M.; Fraser, C. S. Toward a Kinetic Understanding of Eukaryotic Translation. *Cold Spring Harb Perspect Biol* **2019**, *11* (2), a032706. https://doi.org/10.1101/cshperspect.a032706.

(10) *Protein Synthesis and Ribosome Structure*, 1st ed.; John Wiley & Sons, Ltd, 2004. https://doi.org/10.1002/3527603433.

(11) The Nobel Prize in Chemistry 2009 https://www.nobelprize.org/prizes/chemistry/2009/summary/ (accessed 2022 -03 -07).

(12) Higgs, P. G.; Lehman, N. The RNA World: Molecular Cooperation at the Origins of Life. *Nat Rev Genet* **2015**, *16* (1), 7–17. https://doi.org/10.1038/nrg3841.

(13) Buccitelli, C.; Selbach, M. MRNAs, Proteins and the Emerging Principles of Gene Expression Control. *Nat Rev Genet* **2020**, *21* (10), 630–644. https://doi.org/10.1038/s41576-020-0258-4.

(14) Harvey, K. L.; Jarocki, V. M.; Charles, I. G.; Djordjevic, S. P. The Diverse Functional Roles of Elongation Factor Tu (EF-Tu) in Microbial Pathogenesis. *Frontiers in Microbiology* **2019**, *10*.

(15) Kaul, G.; Pattan, G.; Rafeequi, T. Eukaryotic Elongation Factor-2 (EEF2): Its Regulation and Peptide Chain Elongation. *Cell Biochem Funct* **2011**, *29* (3), 227–234. https://doi.org/10.1002/cbf.1740.

(16) Jackman, J. E.; Alfonzo, J. D. Transfer RNA Modifications: Nature's Combinatorial Chemistry Playground. *Wiley Interdiscip Rev RNA* **2013**, *4* (1), 35–48. https://doi.org/10.1002/wrna.1144.

(17) Sonenberg, N.; Hinnebusch, A. G. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* **2009**, *136* (4), 731–745. https://doi.org/10.1016/j.cell.2009.01.042.

(18) López-Lastra, M.; Rivas, A.; Barría, M. I. Protein Synthesis in Eukaryotes: The Growing Biological Relevance of Cap-Independent Translation Initiation. *Biological Research* **2005**, *38* (2–3), 121–146. https://doi.org/10.4067/S0716-97602005000200003.

(19) Hellen, C. U. T.; Sarnow, P. Internal Ribosome Entry Sites in Eukaryotic MRNA Molecules. *Genes Dev.* **2001**, *15* (13), 1593–1612. https://doi.org/10.1101/gad.891101.

(20) Dever, T. E.; Dinman, J. D.; Green, R. Translation Elongation and Recoding in Eukaryotes. *Cold Spring Harb Perspect Biol* **2018**, *10* (8), a032649. https://doi.org/10.1101/cshperspect.a032649.

(21) Neelagandan, N.; Lamberti, I.; Carvalho, H. J. F.; Gobet, C.; Naef, F. What Determines Eukaryotic Translation Elongation: Recent Molecular and Quantitative Analyses of Protein Synthesis. *Open Biology 10* (12), 200292. https://doi.org/10.1098/rsob.200292.

(22) Knight, J. R. P.; Garland, G.; Pöyry, T.; Mead, E.; Vlahov, N.; Sfakianos, A.; Grosso, S.; De-Lima-Hedayioglu, F.; Mallucci, G. R.; von der Haar, T.; Smales, C. M.; Sansom, O. J.; Willis, A. E. Control of Translation Elongation in Health and Disease. *Disease Models & Mechanisms* **2020**, *13* (3), dmm043208. https://doi.org/10.1242/dmm.043208.

(23) Blausen_0657_MultipolarNeuron.png(2500×1612) https://upload.wikimedia.org/wikipedia/commons/1/10/Blausen_0657_MultipolarNeuron.png (accessed 2022 -03 -07).

(24) Jung, H.; Yoon, B. C.; Holt, C. E. Axonal MRNA Localization and Local Protein Synthesis in Nervous System Assembly, Maintenance and Repair. *Nat Rev Neurosci* **2012**, *13* (5), 308–324. https://doi.org/10.1038/nrn3210.

(25) Ross, A. B.; Langer, J. D.; Jovanovic, M. Proteome Turnover in the Spotlight: Approaches, Applications, and Perspectives. *Mol Cell Proteomics* **2020**, *20*, 100016. https://doi.org/10.1074/mcp.R120.002190.

(26) Brown, T. A. *Synthesis and Processing of the Proteome*; Wiley-Liss, 2002.

(27) Thomsen, S.; Anders, S.; Janga, S. C.; Huber, W.; Alonso, C. R. Genome-Wide Analysis of MRNA Decay Patterns during Early Drosophiladevelopment. *Genome Biology* **2010**, *11* (9), R93. https://doi.org/10.1186/gb-2010-11-9-r93.

(28) Kim, E.; Jung, H. Local Protein Synthesis in Neuronal Axons: Why and How We Study. *BMB Rep* **2015**, *48* (3), 139–146. https://doi.org/10.5483/BMBRep.2015.48.3.010.

(29) Sohng, J.-K.; Lee, H.-C.; Liou, K.-K.; Lee, E.-B.; Kang, S.-Y.; Woo, J.-S. Cystocin, a Novel Antibiotic, Produced by Streptomyces Sp. GCA0001: Production and Characterization of Cystocin. *Journal of Microbiology and Biotechnology* **2003**, *13* (4), 483–486.

(30) Kredich, N. M.; Guarino, A. J. Homocitrullylaminoadenosine, a Nucleoside Isolated from Cordyceps Militaris. *J. Biol. Chem.* **1961**, *236* (12), 3300–3302.

(31) Aviner, R. The Science of Puromycin: From Studies of Ribosome Function to Applications in Biotechnology. *Comput Struct Biotechnol J* **2020**, *18*, 1074–1083. https://doi.org/10.1016/j.csbj.2020.04.014.

(32) Nathans, D. PUROMYCIN INHIBITION OF PROTEIN SYNTHESIS: INCORPORATION OF PUROMYCIN INTO PEPTIDE CHAINS. *Proc Natl Acad Sci U S A* **1964**, *51* (4), 585–592.

(33) Yarmolinsky, M. B.; Haba, G. L. D. L. INHIBITION BY PUROMYCIN OF AMINO ACID INCORPORATION INTO PROTEIN*. *Proc Natl Acad Sci U S A* **1959**, *45* (12), 1721–1729.

(34) Robins, M. J.; Miles, R. W.; Samano, M. C.; Kaspar, R. L. Syntheses of Puromycin from Adenosine and 7-Deazapuromycin from Tubercidin, and Biological Comparisons of the 7-Aza/Deaza Pair1. *J. Org. Chem.* **2001**, *66* (24), 8204–8210. https://doi.org/10.1021/jo010935d.

(35) Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chem. Rev.* **2010**, *110* (5), 2579–2619. https://doi.org/10.1021/cr900301e.

(36) Ganser, L. R.; Kelly, M. L.; Herschlag, D.; Al-Hashimi, H. M. The Roles of Structural Dynamics in the Cellular Functions of RNAs. *Nat Rev Mol Cell Biol* **2019**, *20* (8), 474–489. https://doi.org/10.1038/s41580-019-0136-0.

(37) Fallmann, J.; Will, S.; Engelhardt, J.; Grüning, B.; Backofen, R.; Stadler, P. F. Recent Advances in RNA Folding. *Journal of Biotechnology* **2017**, *261*, 97–104. https://doi.org/10.1016/j.jbiotec.2017.07.007.

(38) Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171* (4356), 737–738. https://doi.org/10.1038/171737a0.

(39) Westhof, E.; Fritsch, V. RNA Folding: Beyond Watson–Crick Pairs. *Structure* **2000**, *8* (3), R55–R65. https://doi.org/10.1016/S0969-2126(00)00112-X.

(40) Li, Y.; Breaker, R. R. Kinetics of RNA Degradation by Specific Base Catalysis of Transesterification Involving the 2'-Hydroxyl Group. *J. Am. Chem. Soc.* **1999**, *121* (23), 5364–5372. https://doi.org/10.1021/ja990592p.

(41) Taoka, M.; Nobe, Y.; Yamaki, Y.; Sato, K.; Ishikawa, H.; Izumikawa, K.; Yamauchi, Y.; Hirota, K.; Nakayama, H.; Takahashi, N.; Isobe, T. Landscape of the Complete RNA Chemical Modifications in the Human 80S Ribosome. *Nucleic Acids Research* **2018**, *46* (18), 9289–9298. https://doi.org/10.1093/nar/gky811.

(42) Pan, T. Modifications and Functional Genomics of Human Transfer RNA. *Cell Res* **2018**, *28* (4), 395–404. https://doi.org/10.1038/s41422-018-0013-y.

(43) Statello, L.; Guo, C.-J.; Chen, L.-L.; Huarte, M. Gene Regulation by Long Non-Coding RNAs and Its Biological Functions. *Nat Rev Mol Cell Biol* **2021**, *22* (2), 96–118. https://doi.org/10.1038/s41580-020-00315-9.

(44) Seley-Radtke, K. L.; Yates, M. K. The Evolution of Nucleoside Analogue Antivirals: A Review for Chemists and Non-Chemists. Part 1: Early Structural Modifications to the Nucleoside Scaffold. *Antiviral Res* **2018**, *154*, 66–86. https://doi.org/10.1016/j.antiviral.2018.04.004.

(45) Bajan, S.; Hutvagner, G. RNA-Based Therapeutics: From Antisense Oligonucleotides to MiRNAs. *Cells* **2020**, *9* (1), 137. https://doi.org/10.3390/cells9010137.

(46) Xia, X. Detailed Dissection and Critical Evaluation of the Pfizer/BioNTech and Moderna MRNA Vaccines. *Vaccines (Basel)* **2021**, *9* (7), 734. https://doi.org/10.3390/vaccines9070734.

(47) Nance, K. D.; Meier, J. L. Modifications in an Emergency: The Role of N1-Methylpseudouridine in COVID-19 Vaccines. *ACS Cent. Sci.* **2021**, *7* (5), 748–756. https://doi.org/10.1021/acscentsci.1c00197.

(48) James Slater, R. Radioisotopes in Molecular Biology. In *Reviews in Cell Biology and Molecular Medicine*; John Wiley & Sons, Ltd, 2006. https://doi.org/10.1002/3527600906.mcb.200400162.

(49) Sinkeldam, R. W.; Tor, Y. To D or Not to D? *Org. Biomol. Chem.* **2007**, *5* (16), 2523–2528. https://doi.org/10.1039/B707719J.

(50) Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. Creating New Fluorescent Probes for Cell Biology. *Nat Rev Mol Cell Biol* **2002**, *3* (12), 906–918. https://doi.org/10.1038/nrm976.

(51) Joo, C.; Balci, H.; Ishitsuka, Y.; Buranachai, C.; Ha, T. Advances in Single-Molecule Fluorescence Methods for Molecular Biology. *Annual Review of Biochemistry* **2008**, *77* (1), 51–76. https://doi.org/10.1146/annurev.biochem.77.070606.101543.

(52) A. Thaer, A.; Sernetz, M. *Fluorescence Techniques in Cell Biology*; 1973. https://doi.org/10.1007/978-3-642-49204-4.

(53) Tanpure, A. A.; Pawar, M. G.; Srivatsan, S. G. Fluorescent Nucleoside Analogs: Probes for Investigating Nucleic Acid Structure and Function. *Israel Journal of Chemistry* **2013**, *53* (6–7), 366–378. https://doi.org/10.1002/ijch.201300010.

(54) Xu, W.; Chan, K. M.; Kool, E. T. Fluorescent Nucleobases as Tools for Studying DNA and RNA. *Nat Chem* **2017**, *9* (11), 1043–1055. https://doi.org/10.1038/nchem.2859.

(55) Ju, C.-W.; Bai, H.; Li, B.; Liu, R. Machine Learning Enables Highly Accurate Predictions of Photophysical Properties of Organic Fluorescent Materials: Emission Wavelengths and Quantum Yields. *J. Chem. Inf. Model.* **2021**, *61* (3), 1053–1065. https://doi.org/10.1021/acs.jcim.0c01203.

(56) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Springer US, 2006.

(57) Askim, J. R.; Mahmoudi, M.; Suslick, K. S. Optical Sensor Arrays for Chemical Sensing: The Optoelectronic Nose. *Chem. Soc. Rev.* **2013**, *42* (22), 8649–8682. https://doi.org/10.1039/C3CS60179J.

(58) Ballabio, A.; Bonifacino, J. S. Lysosomes as Dynamic Regulators of Cell and Organismal Homeostasis. *Nat Rev Mol Cell Biol* **2020**, *21* (2), 101–118. https://doi.org/10.1038/s41580-019-0185-4.

(59) Wang, B.; Zhang, L.; Dai, T.; Qin, Z.; Lu, H.; Zhang, L.; Zhou, F. Liquid–Liquid Phase Separation in Human Health and Diseases. *Sig Transduct Target Ther* **2021**, *6* (1), 1–16. https://doi.org/10.1038/s41392-021-00678-1.

(60) Wang, S.; Li, B.; Zhang, F. Molecular Fluorophores for Deep-Tissue Bioimaging. *ACS Cent Sci* **2020**, *6* (8), 1302–1316. https://doi.org/10.1021/acscentsci.0c00544.

(61) Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M. D. Base-Stacking and Base-Pairing Contributions into Thermal Stability of the DNA Double Helix. *Nucleic Acids Research* **2006**, *34* (2), 564–574. https://doi.org/10.1093/nar/gkj454.

(62) Schulz, E. C.; Seiler, M.; Zuliani, C.; Voigt, F.; Rybin, V.; Pogenberg, V.; Mücke, N.; Wilmanns, M.; Gibson, T. J.; Barabas, O. Intermolecular Base Stacking Mediates RNA-RNA Interaction in a Crystal Structure of the RNA Chaperone Hfq. *Sci Rep* **2017**, *7* (1), 9903. https://doi.org/10.1038/s41598-017-10085-8.

(63) Dannenberg, J. J. An Introduction to Hydrogen Bonding By George A. Jeffrey (University of Pittsburgh). Oxford University Press: New York and Oxford. 1997. Ix + 303 Pp. $60.00. ISBN 0-19-509549-9. *J. Am. Chem. Soc.* **1998**, *120* (22), 5604–5604. https://doi.org/10.1021/ja9756331.

(64) Ge, J.; Zhang, C.-W.; Ng, X. W.; Peng, B.; Pan, S.; Du, S.; Wang, D.; Li, L.; Lim, K.-L.; Wohland, T.; Yao, S. Q. Puromycin Analogues Capable of Multiplexed Imaging and Profiling of Protein Synthesis and Dynamics in Live Cells and Neurons. *Angewandte Chemie International Edition* **2016**, *55* (16), 4933–4937. https://doi.org/10.1002/anie.201511030.

(65) Chapuis, H.; Strazewski, P. Shorter Puromycin Analog Synthesis by Means of an Efficient Staudinger–Vilarrasa Coupling. *Tetrahedron* **2006**, *62* (51), 12108–12115. https://doi.org/10.1016/j.tet.2006.09.045.

CHAPTER 2: Photophysics of novel dialkylated fluorophores

2.1 Introduction

As mentioned in the previous chapter, the native ribonucleosides possess impractical photophysical features with extremely low quantum yields and brightness coefficients.[1] While higher than the canonical nucleosides, m^{6,6}A (figure 2.1a), a modified adenosine analaogue, still suffers from similar deficient photophysical properties as well with low emission quantum yields (<1 % across various solvents).[2,3] Intriguingly, N^6,N^6-dimethyladenosine was reported by Albinsson to possess dual emission in various solvents with emission maxima at 355 and 568 nm.[2] The red emission band, attributed to the twisted intramolecular charge transfer state (TICT), predominates in aprotic environments but is significantly quenched by protic solvents.[2] In our quest to develop non-perturbing fluorescent nucleosides, we sought to create a small molecule
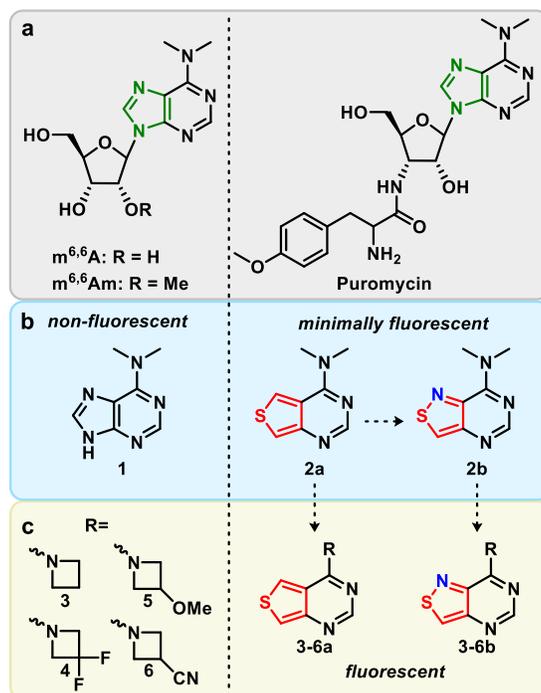
Figure 2.1. Structures of (a) N^6,N^6-dimethyladenosine (m^6,^{6}A), N^6,N^6,2'-O-trimethyladenosine (m^6,^{6}Am), and puromycin. (b) Thieno- and isothiazolo- derived purines bearing dimethylamine substituents. (c) Thieno- and isothiazolo- derived purines bearing azetidine and difluoroazetidine substituents.

fluorescent analogue of N^6,N^6-dimethyladenine (**1**, figure 2.1b) that remains isomorphic and isofunctional, yet exhibits augmented photophysical properties for use as emissive probes of nucleosides/puromycin bioactivity.
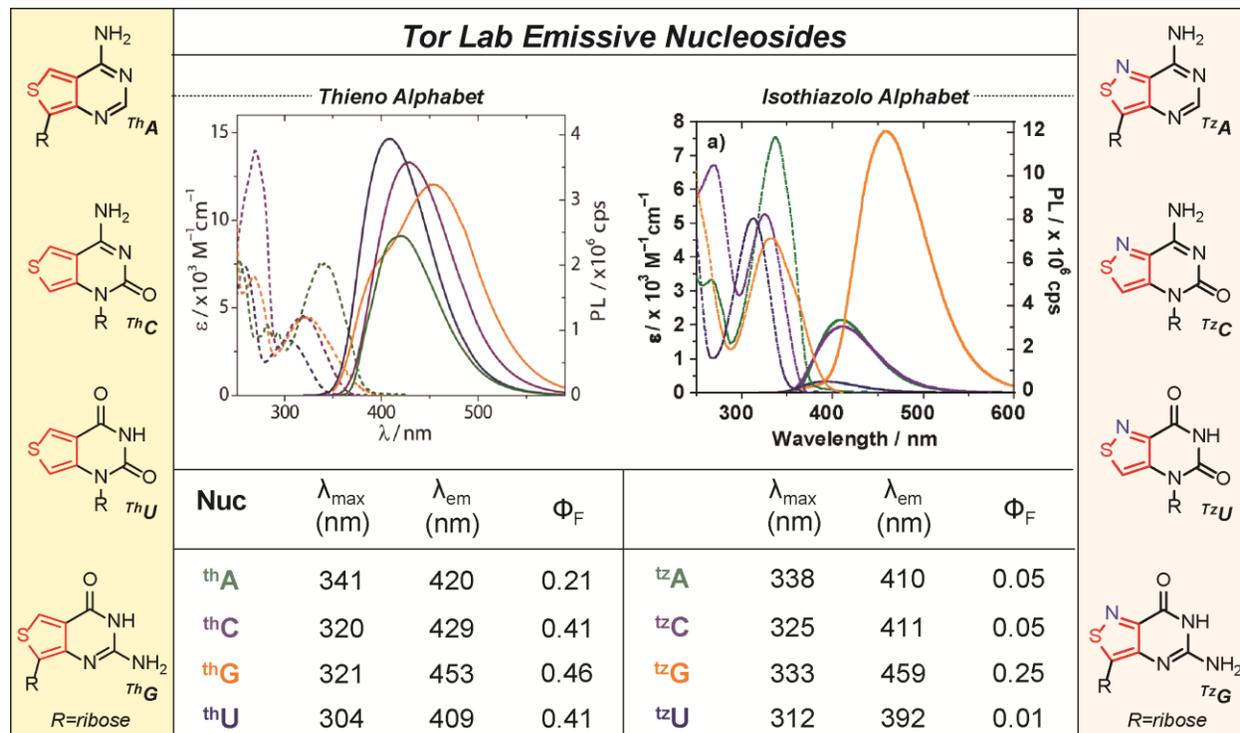
Figure 2.2. Thieno- and Isothiazolo- based purine and pyrimidine analogues developed by Tor.[4-6]

Over the last two decades, we have developed and investigated two emissive RNA alphabets comprising both purine and pyrimidine analogues (figure 2.2).[4–6] The first alphabet, derived from a thieno[3,4-*d*]-pyrimidine heterocyclic nucleus, exhibits markedly improved photophysical properties compared to the native ribonucleosides as evidenced by its employment in the analysis of various biomolecular processes.[5] Its limitation is the lack of a nitrogen atom corresponding to the 7 position of the native purine skeleton. To circumvent this, a second generation alphabet based on an isothiazolo[4,3-*d*]pyrimidine core was synthesized.[4,5] While the restoration of the nitrogen atom into the Hoogsteen face resulted in lower emission quantum yields, the higher degree of isofunctionality facilitated a near-identical behavior of the

fluorescent probes compared to the thieno alphabet when studying N7-dependent processes.[4,6–10] With these fluorescent nucleosides at our disposal, we opted to synthesize and explore the photophysical properties of the N^6,N^6-dimethyladenine derivatives adopting both thieno- and isothiazolo-pyrimidines in place of the imidazole ring of the native purine (**2a** and **2b**, figure 2.1b,c), and assess their potential for generating emissive nucleoside analogs.

Similarly to the impact of dimethylation on the photophysics of native adenosine,[2,3] we observed elevated emission quantum yields for the dimethylated derivatives of the corresponding thieno- and isothiazolo-pyrimidines but they overall retained relatively of low brightness values. To therefore expand the scope of fluorescent demethylated adenines and adenosines, we thus employed modified azetidines in place of the dimethylamine group to produce purines with augmented photophysical properties. Originally developed by Luke Lavis,[11–13] substituting the Me₂N group found in common fluorophores with azetidines was reported to significantly improve the photophysical features of nearly all fluorophore scaffolds analyzed.[11,12] Substituents at the 3-position of the azetidine ring were found to further tune the emission maxima.[11] Considering the fluorescent nucleobase analogues discussed here contain dimethylamine substituents, we opted to synthesize novel fluorophores employing substituted azetidine and compare them to the Me₂N bearing heterocycles. Here we report synthetic approaches to these novel structures and their photophysical properties, where unprecedented quantum yields and brightness coefficients are observed. The most promising combination of a purine surrogate with a substituted azetidine was then implemented in a nucleoside skeleton.

2.2 Synthesis of fluorescent N^6,N^6-dimethyladenine derivatives

We first synthesized the dimethylated thieno and isothiazolo derivatives **2a** and **2b** through a 3-step divergent pathway, starting from either methyl 4-aminothiophene-3-carboxylate

hydrochloride (**7a**) or methyl 4-aminoisothiazole-3-carboxylate hydrochloride (**7b**), respectively.[4] A facile cyclization of the 4-aminoheterocycles with formamidine acetate yielded the inosine analogues **8a** and **8b**, respectively, in good yield. Thionation of these purine analogues with phosphorous pentasulfide or Lawesson's reagent produced the thiopheno (**9a**) and isothiazolo (**9b**) analogues in 83% and 92% yield, respectively. Displacement reactions with methanolic dimethylamine yielded the desired heterocycles **2a** and **2b** in 64% and 69% yield, respectively. X-ray crystallography confirmed the proposed heterocycle structures.
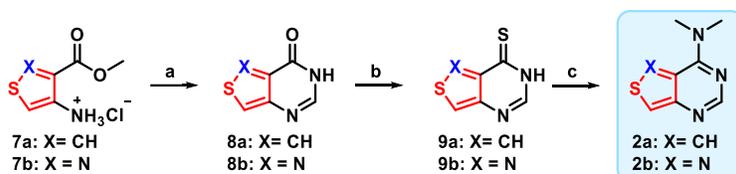
Scheme 2.1. Synthesis of Dimethyladenine Analogues **2a** and **2b**: (a) Formamidine acetate, EtOH, reflux, ON, X=CH: 78%, X=N: 76%. (b) i. X=CH: P_2S_5, Pyridine, 110 °C, 2 h, 83%; X=N: Lawesson's reagent, Pyridine, 110 °C, 2 h, 92%. (c) 2M $HNMe_2$ in MeOH, 60 °C, ON, X=CH: 64%, X=N: 69%

2.3 Photophysical properties of fluorescent N^6,N^6-dimethyladenine derivatives.

The absorption and emission spectra of the two heterocycles **2a** and **2b** are given in figure 2.3 and the fundamental spectroscopic properties are summarized in table 2.1. Both compounds displayed substantial bathochromic shifted maxima in their aqueous ground-state absorption spectra compared to native $m^{6,6}A$ (346 nm and 350 nm, respectively, vs. 275 nm). The corresponding aqueous emission maxima exhibited only one emission band at 411 nm and 412 nm as opposed to the dual emission bands observed for the native dimethylated adenosine. No significant variations in quantum yield were observed when compared to $m^{6,6}A$ (<1%), indicating that the two synthetic dimethylated analogues remained weakly fluorescent in aqueous media. The absorption and emission spectra largely retained similar photophysical properties in D_2O, indicating no major solvent isotope effect. However, an increase in the brightness

coefficient of the thieno analogue **2a** from 5 to 14 due to an increase in quantum yield, along with a corresponding decrease in brightness of the isothiazolo derivative **2b** from 1.6 to 1.1 due to a drop in extinction coefficient, hinted to a potential minor solvent isotopic effect. Notably, the changes observed for **2b** are likely within error of inherently miniscule values, but nevertheless, any solvent isotopic effects are likely due to the presence of the N7 nitrogen augmenting hydrogen bonding through the Hoogsteen face of the heterocycle.
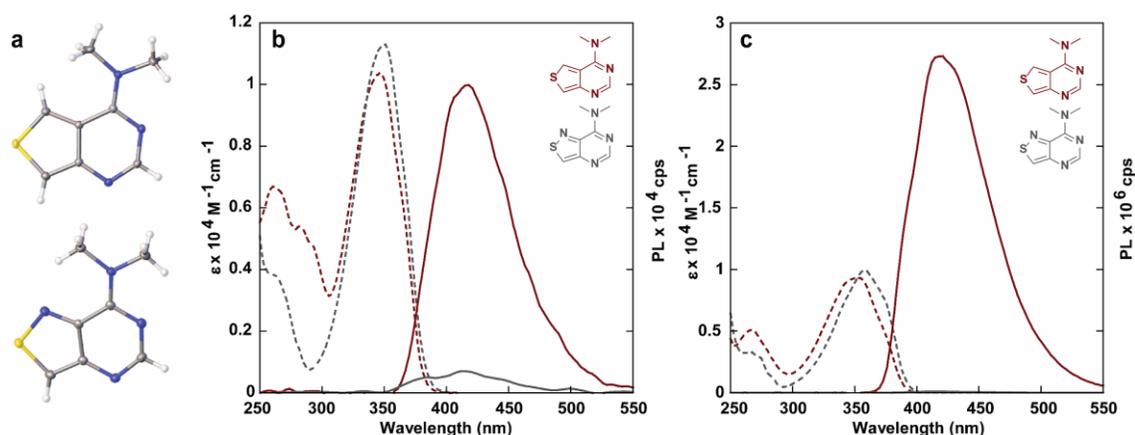
Figure 2.3. (a) Crystal structures of compound **2a** (top) and **2b** (bottom). (b) absorption (dashed line) and emission (solid line) spectra of compounds **2a** and **2b** in water and (c) dioxane.

Unlike the parent **1**, the absorption spectra taken in dioxane were slightly red-shifted by almost 10 nm compared to aqueous solutions (346 to 353 nm, 350 to 358 nm for **2a** and **2b** respectively), and a similar trend was also observed in the emission spectra (411 nm to 416 nm, 412 nm to 418 nm, respectively, excitation at 350 nm). Intriguingly, while the quantum yield of analogue **2b** increased ten-fold, that of corresponding analogue **2a** increased nearly a thousand-fold and sharply rose to 47%. Additionally, exciting **2a** in 10 nm increments between 260 nm and 380 nm in water and dioxane indicated no changes to the emission band after correcting for emission intensity. Measuring the excitation spectra of the emission bands between 360 and 560 nm in water and dioxane indicated no distinct changes after correcting for intensity, suggesting

Table 2.1. Photophysical Properties of Thieno [3,4-*d*] and Isothiazolo [4,3-*d*] Nucleobase analogues **2a** and **2b**

	Solvent	λ_{abs} (ε)[a]	λ_{em} (Φ)[a]	$\Phi\varepsilon$	Stokes Shift[a]	Polarity Sensitivity[b]
1[c]	water	275 (n/a)	355 (0.00035) 568 (0.00026)	n/a	n/a	n/a
	dioxane	275 (n/a)	333 (0.00051) 493 (0.00062)	n/a	n/a	
	D$_2$O	275 (n/a)	355 (0.00040) 568 (0.00034	n/a	n/a	
2a	water	346 (1.03)	411 (0.0005)	5	3.80	-0.69
	dioxane	353 (0.94)	416 (0.47)	4392	4.33	
	D$_2$O	346 (0.98)	416 (0.0015)	14	4.77	
2b	water	350 (1.13)	412 (0.0001)	1.5	n/a	n/a
	dioxane	358 (0.92)	418 (0.0010)	11	n/a	
	D$_2$O	349 (1.11)	417 (0.0001)	1.1	n/a	

[a] λabs, ε, λem and Stokes shift are reported in nm, 10^4 M^{-1} cm^{-1}, nm and 10^3 cm^{-1} respectively. All the photophysical values reflect the average over three independent measurements. [b] Sensitivity to solvent polarity reported in cm^{-1}/(kcal mol^{-1}) is equal to the slope of the linear fit in figure 2.29.[c] Compound is reported as the nucleoside.[10] For full error analysis see table 2.9.

that all emissive transitions observed are a result of excitation through the low energy absorption band (figure 2.10c). However, when analyzing its sensitivity to polarity (Figure 2.4a), we found water to be an effective quencher (with as little as 5% in dioxane enough to drastically diminish the emission), a phenomenon noted in the behavior of (m^{6,6}A) and its TICT state fluorescence quenching. Furthermore, substitution of the carbon atom at the 7-position of compound **2a** (following native purine numbering) with a nitrogen atom (compound **2b**) results in a loss of emission in dioxane.

Fluorophores bearing dialkylamine substituents are known to non-radiatively decay through rotational relaxation around the carbon-nitrogen bond. To investigate the prevalence of this phenomenon in the analogues synthesized, we collected absorption and emission spectra in methanol/glycerol mixtures of varying viscosity (Figure 2.4b). Intriguingly, due to the analogues' high sensitivity to microenvironmental polarity it appears that the effects of polarity

outweigh the impact of viscosity in low % glycerol solutions (Figure 2.21), despite the relatively small polarity difference between the two ($\Delta E_T(30) = 1.6$ kcal mol^{-1}).[14] A drop in emission intensity is observed from 0 to 40% glycerol, yet beyond that it seems the effect of viscosity dominates, and the emission intensity steadily rises in solutions containing 40–100% glycerol. To circumvent or at least minimize the inevitable effects of polarity changes in solvent mixtures, we attempted to assess the effects of viscosity in a single solvent system, measuring the emission intensity in ethylene glycol whilst varying the temperature (figure 2.23a). Notably, while the optical density remained unchanged, the emission intensity dropped slightly with increasing temperature, as expected due to the expected inverse relationship between emission quantum yield and temperature for Me$_2$N-containing chromophores.
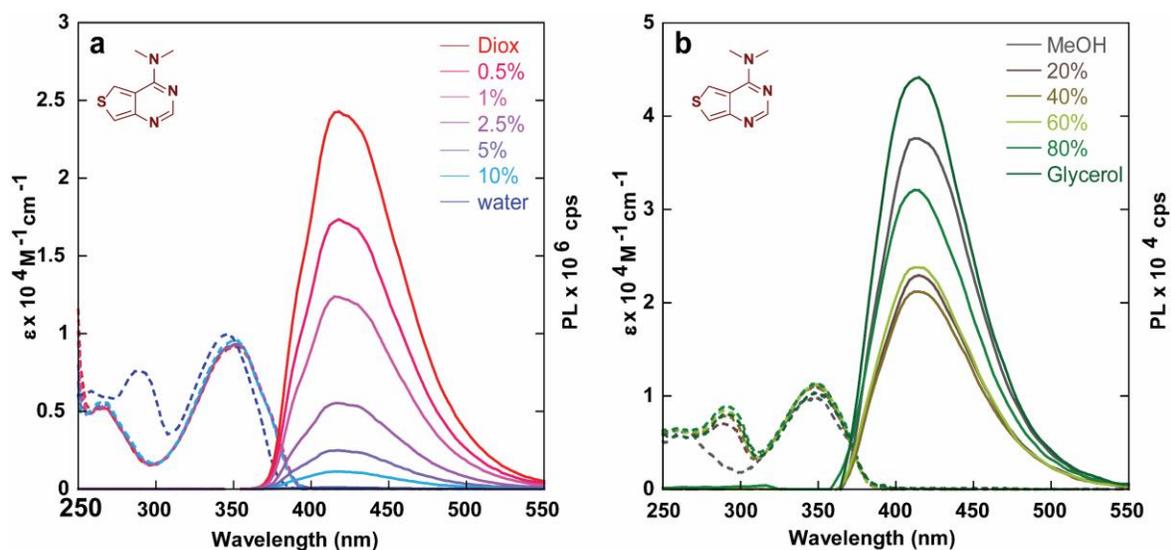
Figure 2.4. Absorption (dashed line) and emission (solid line) of compound **2a** in (a) water/dioxane mixtures and (b) MeOH/glycerol mixtures.

2.4 Synthesis of fluorescent azetidine-modified adenine analogues

Ultimately, the dimethylated derivatives **2a** and **2b** exhibited less than desirable photophysical characteristics in aqueous solution due to their low emission quantum efficiencies, with only the thieno analogue **2a** displaying strong fluorescence in relatively apolar media (e.g.,

dioxane), albeit with unusually high sensitivity to water content. We have therefore opted to employ the methodology pioneered by Lavis and substitute the dimethylamine groups on these fluorophores for azetidines. The strained azetidine ring reduces rotation around the C–N bond,[15] frequently limiting the non-radiative decay pathway through a TICT pathway.[15] Additionally, substituents at the azetidine's 3 position were reported to shift the emission maxima across the visible, thus providing a molecular handle for tuning the photophysical features.[11] Initially, we opted to introduce the unsubstituted azetidine as well as the similarly sized 3,3-difluoroazetidine. Both the azetidine-bearing thieno and isothiazolo purine derivatives **3a–6a** and **3b–6b**, respectively, were synthesized through a substitution of the sulfur atom of the thionated intermediates **9a** and **9b** in methanol (scheme 2.2), similar to the pathway used for the heterocycles **2a** and **2b** (scheme 2.1). Since the commercial azetidines are provided as their hydrochloride salts, excess DIPEA was added to release the free cyclic secondary amines (scheme 2.2).

(a) H$_2$RCl, DIPEA, MeOH, 60 °C, ON
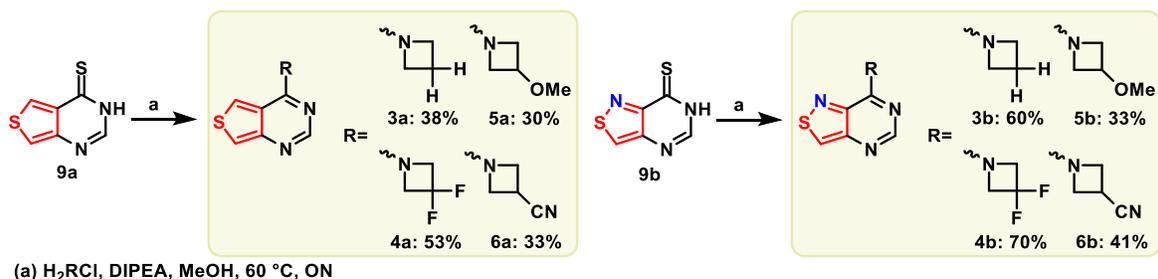
Scheme 2.2. Synthesis of azetidine-modified analogues **3–6a/b** (a) H$_2$RCl, DIPEA, MeOH, 60 °C, ON

2.5 Photophysical properties of fluorescent azetidine-modified adenine analogues

The absorption and emission spectra of heterocycles **2–6a** and **2–6b** are shown in figure 2.5 and their fundamental spectroscopic features are summarized in Table 2. Compared to the "parent" dimethylamino derivatives **2a** and **2b**, no significant changes in the absorption spectra were observed for the thieno- or isothiazolo- analogues bearing the azetidino substituents

36

in aqueous solution, within a 10 nm range. A similar trend was observed in deuterium oxide. The emission maxima also exhibited minimal changes, yet the fluorescence quantum yields differed drastically, with compound **4a** exhibiting a quantum yield of 64% in aqueous solution, nearly a thousand-fold higher than that of compound **2a** (<0.1 %, table 2.2). Intriguingly, the quantum yields of the isothiazolo analogues **3b** and **4b** also increased compared to analogue **2b** (0.3 % and 6% vs. 0.1), but not to the same degree observed in analogues **3a** and **4a** (6% and 64%, respectively). Plotting the quantum yields values for **3–6b** in water against the respective Hammett inductive substituent constants (σ_1)[16] revealed an excellent correlation (figure 2.6). In D_2O, a hypsochromic shift of 10 nm is observed in the emission maxima of **3a** compared to **2a**, which was not observed for **4a** and **2a** (406 and 418 vs. 416 nm), and an opposing trend is observed for the isothiazolo analogues, with azetidino derivative **3b** possessing a similar emission maxima to **2b** and the difluoro derivative **4b** displaying a hypsochromic shift (415 and 403 vs. 417 nm, respectively), indicating the potential presence of a minor solvent isotope effect. The fluorescence quantum yield in D2O generally remain the same, although a slight decrease is observed for the difluoroazetidino-modified heterocycle **4a** (0.57 in D_2O vs. 0.64 in H_2O).
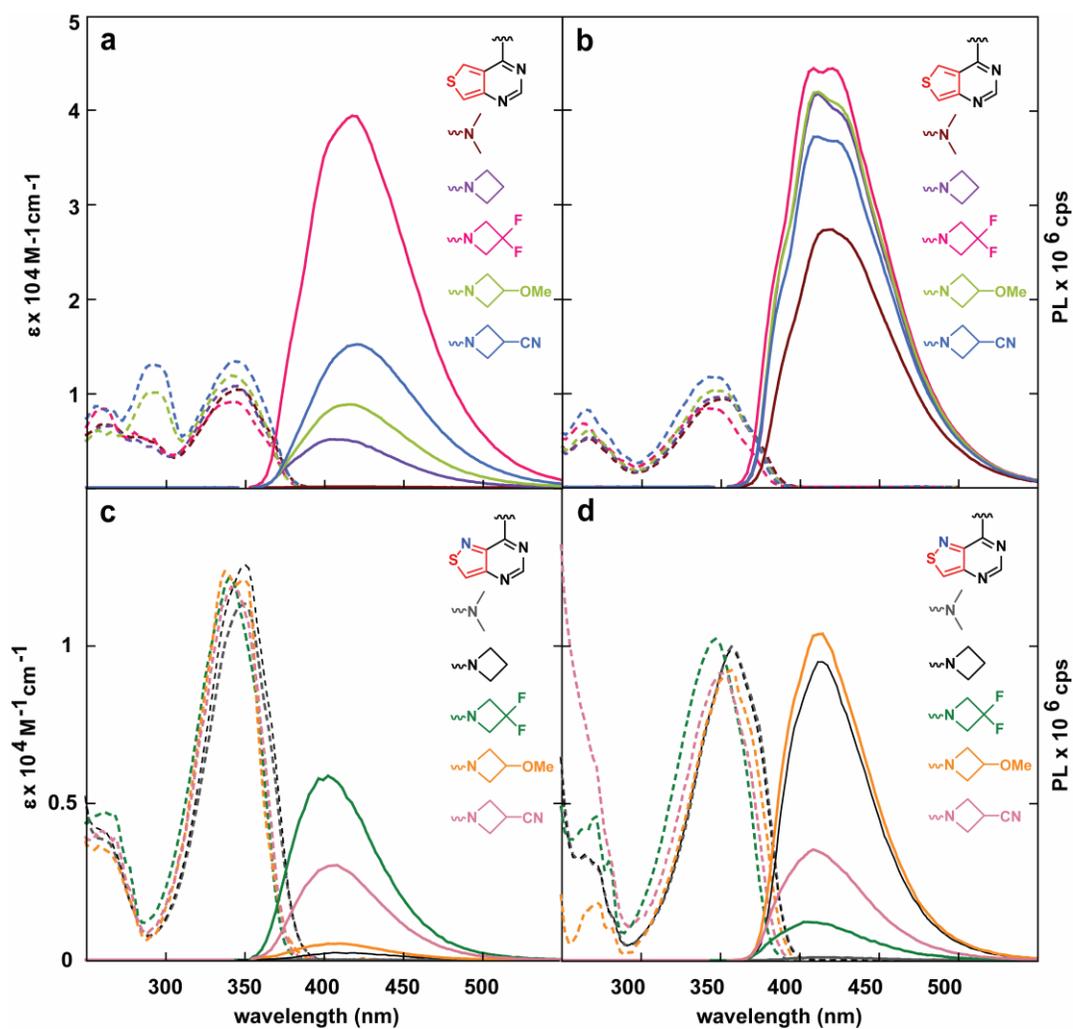
Figure 2.5. Absorption (dashed line) and emission (solid line) spectra of analogues **2–6a** in (a) water (b) dioxane. (c) Absorption (dashed line) and emission (solid line) spectra of analogues **2–6b** in water and (d) dioxane.
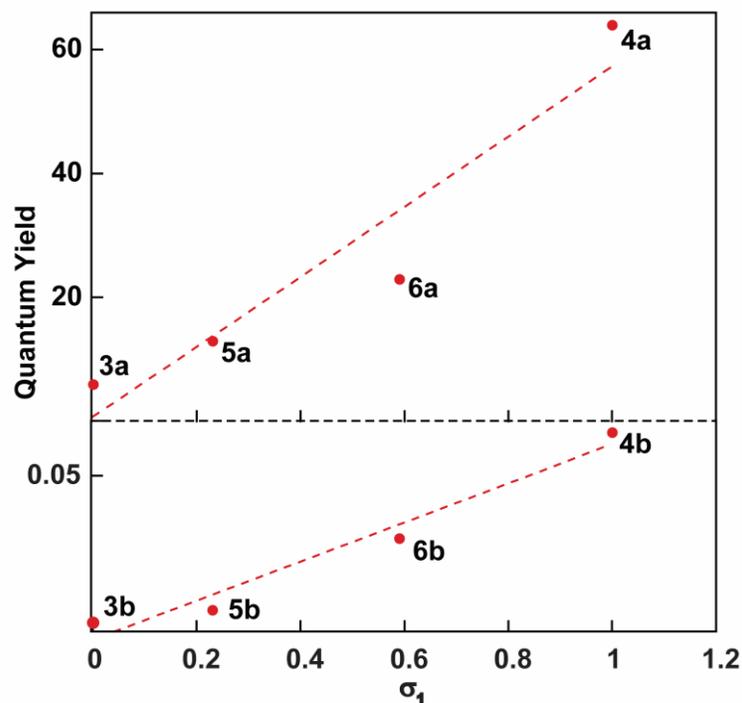
Figure 2.6. Quantum yield vs Hammett inductive substituent constants for compounds **3–6a** (bottom) and compounds **3–6b** (top).

The absorption maxima measured in dioxane show a hypsochromic shift for compounds **4–6a** and **4–6b** containing modified azetidines compared to **2a/3a** and **2b/3b**. While compounds **2–6a** exhibited more classically shaped emission bands in water, the emission spectra in dioxane are highly irregular, likely indicating multiple partially overlapping emission bands (as the chromophores cannot tautomerize). Additionally, compounds **3–6a** and **3–6b** show significant increases in their emission quantum efficiency (0.64–0.77 and 0.02–0.14 respectively). Intriguingly, the emission quantum efficiencies of the isothiazolo- analogues **3–6b** increased in dioxane compared to **2b** but to a lesser extent than those seen in the thieno- derivatives. Taken together, the azetidine substitutions endows the respective heterocycles with unique photophysical properties, with the unsubstituted azetidine modification slightly improving

Table 2.2: Photophysical Properties of Thieno [3,4-*d*] and Isothiazolo [4,3-*d*] Nucleobase analogues **2-6a/ b**

	Solvent	λ_{abs} (ϵ)[a]	λ_{em} (Φ)[a]	$\Phi\epsilon$	Stokes Shift[a]	Polarity Sensitivity[b]
2a	water	346 (1.03)	411 (0.0005)	5	3.80	-0.69
	dioxane	353 (0.94)	416 (0.47)	4392	4.33	
	D$_2$O	346 (0.98)	416 (0.0015)	14	4.77	
3a	water	344 (1.11)	407 (0.06)	715	4.40	53
	dioxane	351 (1.27)	412 (0.75)	9484	4.23	
	D$_2$O	345 (1.00)	406 (0.07)	687	4.41	
4a	water	344 (0.90)	419 (0.64)	5766	5.14	81
	dioxane	343 (0.85)	419 (0.77)	6587	5.26	
	D$_2$O	343 (0.84)	418 (0.57)	4826	5.29	
5a	water	344 (1.12)	415 (0.13)	1421	4.99	80
	dioxane	349 (1.03)	411 (0.70)	7218	4.28	
	D$_2$O	344 (1.11)	416 (0.15)	1728	5.00	
6a	water	345 (1.30)	421 (0.23)	2951	5.17	110
	dioxane	345 (1.19)	409 (0.64)	7579	4.52	
	D$_2$O	344 (1.27)	416 (0.33)	4164	5.03	
2b	water	350 (1.17)	412 (0.0001)	1.6	n/a	n/a
	dioxane	358 (1.03)	418 (0.0010)	14	n/a	
	D$_2$O	349 (1.11)	417 (0.0001)	1.1	n/a	
3b	water	352 (1.22)	416 (0.0030)	33	4.38	-17
	dioxane	358 (1.05)	415 (0.13)	1192	3.84	
	D$_2$O	349 (1.23)	415 (0.0030)	41	4.51	
4b	water	341 (1.21)	403 (0.064)	770	4.54	11
	dioxane	348 (1.03)	407 (0.020)	173	4.19	
	D$_2$O	341 (1.22)	403 (0.073)	884	4.49	
5b	water	346 (1.11)	405 (0.007)	77	4.12	-2
	dioxane	356 (0.96)	413 (0.14)	821	3.84	
	D$_2$O	346 (1.06)	408 (0.010)	102	4.47	
6b	water	344 (1.11)	406 (0.03)	340	4.45	34
	dioxane	352 (0.94)	410 (0.06)	607	4.07	
	D$_2$O	343 (1.07)	406 (0.03)	293	4.49	

[a] λ_{abs}, ϵ, λ_{em} and Stokes shift are reported in nm, 10^4 M^{-1} cm^{-1}, nm and 10^3 cm^{-1} respectively. All the photophysical values reflect the average over three independent measurements. [b] Sensitivity to solvent polarity reported in cm^{-1}/ (kcal mol^{-1}) is equal to the slope of the linear fit in Figure 2.29. For full error analysis see table 2.9.

quantum efficiency in water yet dramatically in dioxane, and the difluoroazetidine significantly

elevating quantum efficiency in both water and dioxane. As previously observed, however, the

presence of the "N7" atom seems to dramatically enhance the propensity for the fluorophores to

non-radiatively decay. Whether that is through increased relaxation through a "dark" state or through solvent-mediated non-radiative relaxation of the excited state remains unclear.

Emission spectra of compounds **3–6b** and **4–6b** in water or dioxane, taken via excitation every 10 nm of their absorption band, show little variation in emission maxima, yet stark changes in emission intensity, suggesting minimal impact on their Franck-Condon states (figure 2.10–2.18). Analogue **4a**, however, shows a slight shift in the emission band in water, indicating the potential presence of two emissive states. This, in conjunction with excitation spectra of compounds **3–6a/b** taken at varying emission wavelengths show stark similarities to thieno-derivative **2a** (Figure 2.10). Intriguingly, only derivative **4a** shows variation in the excitation spectrum in water, with a particular excitation band at 280 nm increasing in intensity as the emission approaches longer wavelengths.

The sensitivity to polarity of compounds **3–6a/b** in water–dioxane mixtures indicated that both azetidine-modified analogues **3a** and **3b** seemed to mirror the behavior of predecessor **2a**, in that upon addition of water the fluorescence intensity of the main emission band decreases dramatically (Figure 2.7). Intriguingly, while isothiazolo analogue **4b** exhibits increasing fluorescence intensity in water/dioxane mixtures with more water content due to its corresponding quantum efficiency in the two solvents, the thieno derivative **4a** appears sporadic. A similar trend was observed in **6a** as well, the two bases with the highest inductive Hammett constants. In water/dioxane mixtures with low water content, the fluorescence intensity initially increases plateauing at 20 or 40%. The compound then exhibits lower fluorescence intensity in solutions of higher water content, and the emission curve approaches a more "Gaussian" form (figure 2.7b, figure 2.27). Since electron-withdrawing substituents comprise the structural difference between compounds **3–6a**, this suggests that the stark variations observed in the

41

water-dioxane mixtures are likely attributed to their inductive impact on the charge transfer band of the heterocycle, a property previously introduced in Figure 2.5. However, it is important to note that the inductive effect is not one from a substituent directly bonded to the π-system and signifies a rare occurrence in which a functional group separated from the heterocycle by sigma
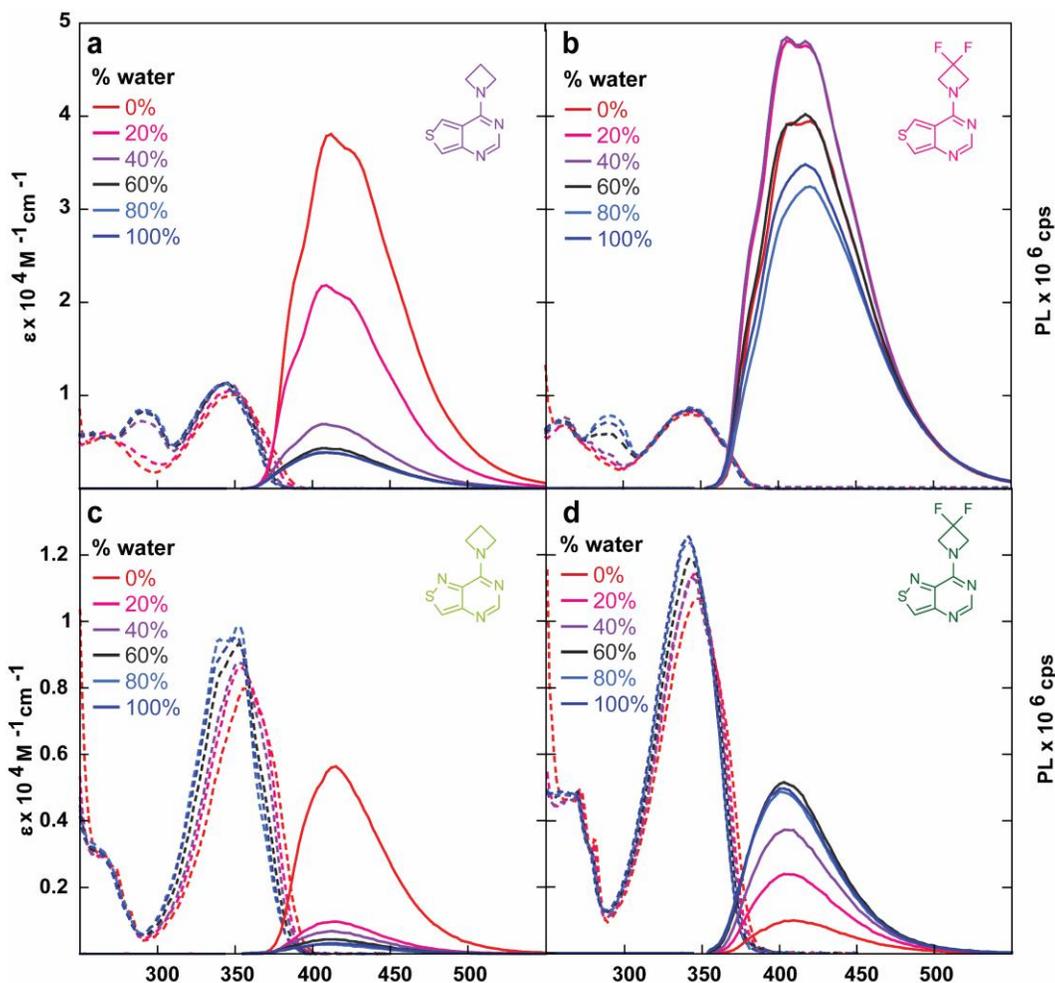
Figure 2.7. Absorption (dashed line) and emission (solid line) spectra of compounds (a) **3a**, (b) **4a**, (c) **3b**, and (d) **4b** in water/dioxane mixtures.

bonds is able to affect the photophysical properties of the chromophore. This inductive effect was also observed in Lavis's original application of 3-modified azetidines.[11] However, the effects were observed on emission maxima rather than quantum yield as discussed here.

42

Analysis of compounds' **3–6a/b** responsiveness to viscosity in ethylene glycol by varying the temperature indicated similar behavior to compound **2a**, with no differences in the absorption spectra yet slight decreases in emission intensity (Figure 2.23–2.26). Emission spectra taken in methanol/glycerol mixtures indicates bathochromic and hypsochromic shifts in mixtures with higher glycerol contents for compounds **4a** and **3b**, respectively (Figure 2.8). Given that
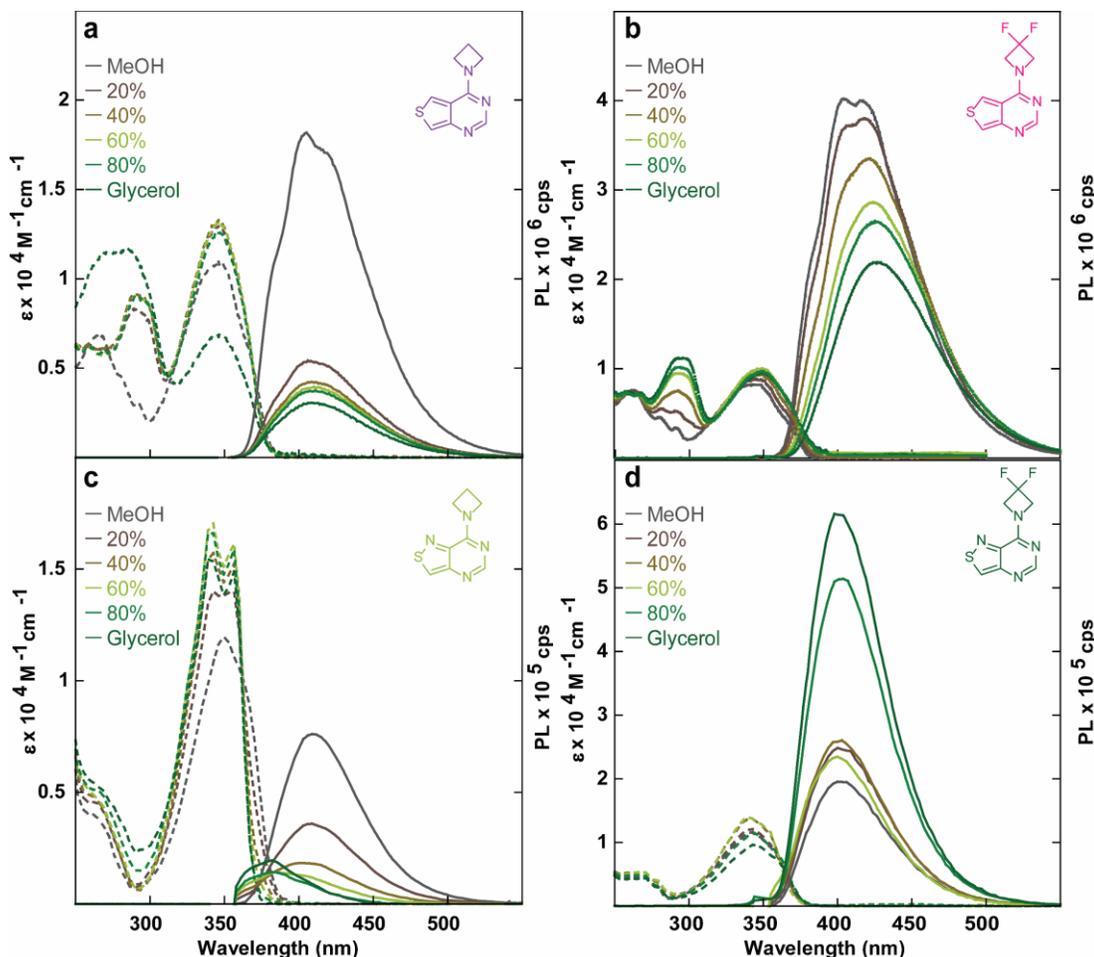
Figure 2.8. Absorption (dashed line) and emission (solid line) spectra of compounds (a) **3a**, (b) **4a**, (c), **3b**, and (d) **4b** in MeOH/glycerol mixtures.

structural rigidification through azetidine modifications are reported to diminish bond rotation compared to dialkylamino groups in both the ground and excited states, the diminution in fluorescence was expected, yet the shifts in emission maxima indicate potential solvent-

dependent interactions with the excited state, despite very similar, though not identical, polarity for methanol and glycerol.

Inspired by the photophysical properties of compound **4a**, we synthesized the ribonucleoside analogue **11** as a potential fluorescent surrogate of m^{6,6}a. Deprotection of benzoate protected thioribonucleoside analogue **10**[5] with 1:1 satd methylamine: satd. ammonium hydroxide, followed by S$_n$Ar with difluoroazetidine yielded the desired nucleoside **11** in 61% yield over two steps.
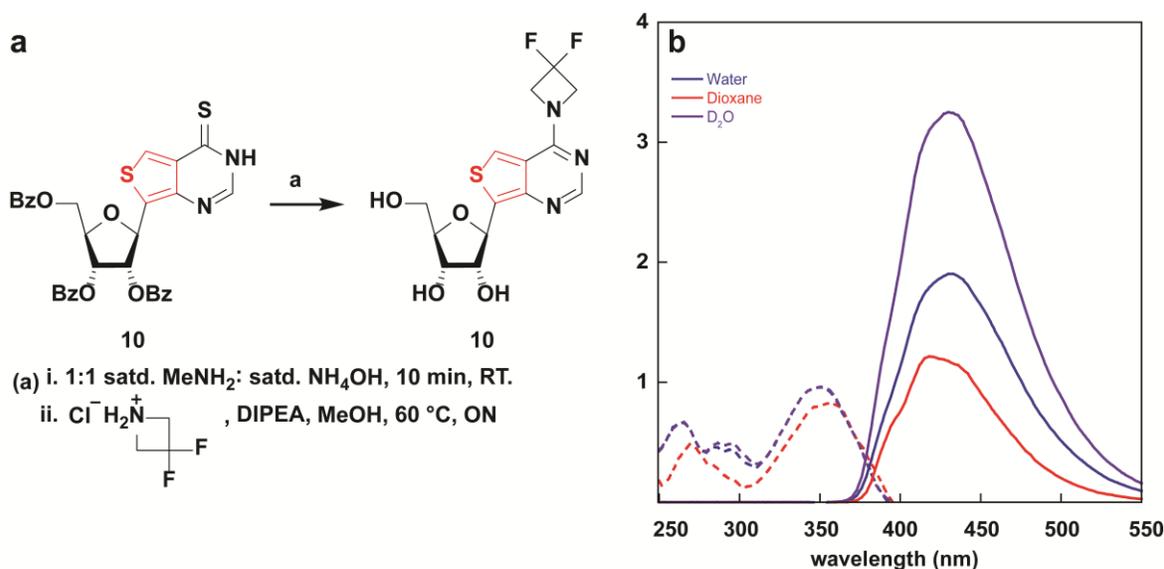
Figure 2.9. Synthesis of nucleoside **11**. (b) Absorption (dashed line) and emission (solid line) of nucleoside **11** in water.

The photophysical parameters displayed by nucleoside **11** (figure 2.9) differ from the ones possessed by its aglycon counterpart **4a**. Incorporating the ribose moiety resulted in approximately 10 nm bathochromic shifts in the absorbance maxima in water, dioxane, and D$_2$O compared to the nucleobase (352 vs 344 nm, 357 vs 343, 352 vs 343 nm, respectively). Similar shifts are visible in the emission maxima (435 vs 419 nm, 420 vs 419 nm, 431 vs 418 nm respectively) (table 2.2, table 2.3).

44

Table 2.3. Photophysical Properties of Nucleoside **11**

	Solvent	λ_{abs} (ε)[a]	λ_{em} (Φ)[a]	$\Phi\varepsilon$	Stokes Shift[a]
11	water	352 (0.94)	435 (0.30)	2838	5.41±0.19
	dioxane	357 (0.85)	420 (0.17)	1444	4.21±0.01
	D_2O	352 (0.95)	431 (0.47)	4486	5.23±0.05

[a] λ_{abs}, ε, λ_{em} and Stokes shift are reported in nm, 10^4 M^{-1} cm^{-1}, nm and 10^3 cm^{-1} respectively. All the photophysical values reflect the average over three independent measurements. For full error analysis see table 2.9.

Intriguingly, a two-fold and nearly a five-fold drop in quantum yield between **4a** and **11** in water and dioxane, respectively, was observed (0.64 vs. 0.30 and 0.77 vs. 0.17, respectively). The emission quantum yield exhibits a significantly smaller drop in D_2O (0.57 vs 0.47), indicating solvent-assisted quenching pathways play critical roles in defining the excited state landscape and processes of such chromophores. The substantial differences observed between H_2O and D_2O align well with a recent universal model proposed for the quenching effect of water, suggesting energy transfer to vibrational overtones and high energy stretching modes that exist in water and alcohols but become much less significant in the corresponding deuterated counterparts.[17] This observation explain the dichotomy frequently observed (although not well documented) between the photophysical features of the nucleobases compared to the corresponding modified nucleosides, suggesting a critical role for the sugar moiety that is frequently considered to be non-chromophoric and photophysically benign. The proximity of the sugar hydroxyl and their solvation may therefore impact the excited state features and should be further considered in designing fluorescent nucleosides and other emissive probes.

2.6 Summary

A new set of N^6,N^6 dialkylated fluorescent adenine analogues **3–6a/b** incorporating modified azetidines has been developed. Given the intense fluorescence of difluoroazetidine-modified nucleobase **4a**, nucleoside **11** was developed and exhibited significant solvatochromic differences in photophysical properties compared to **4a** which can be utilized as a tool for potentially studying distinct environments. The azetidine modification drastically enhanced the emission quantum yield of the fluorophores, turning otherwise non-fluorescent heterocycles into useful fluorophores in aqueous solutions. While the azetidine modification has been explored to improve small molecule fluorophores commonly used in molecular biology, this tactic shows promise for the field of fluorescent modified nucleosides to produce novel chromophores with higher sensitivities. Additionally, the rigorously assessed solvent and isotopic effects on the emission quantum yield of the new nucleobases/nucleoside should alert the community to the intricacies of fluorophore design and implementation in this and related fields.

2.7 Acknowledgements

2.8 References

(1) Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chem. Rev.* **2010**, *110* (5), 2579–2619. https://doi.org/10.1021/cr900301e.

(2) Albinsson, B. Dual Fluorescence from N6,N6-Dimethyladenosine. *J. Am. Chem. Soc.* **1997**, *119* (27), 6369–6375. https://doi.org/10.1021/ja9705203.

(3) Ballini, J.-P.; Daniels, M.; Vigny, P. Synchrotron-Excited Time-Resolved Fluorescence Spectroscopy of Adenosine, Protonated Adenosine and 6N, 6N-Dimethyladenosine in Aqueous Solution at Room Temperature. *Eur Biophys J* **1988**, *16* (3). https://doi.org/10.1007/BF00261899.

(4) Rovira, A. R.; Fin, A.; Tor, Y. Chemical Mutagenesis of an Emissive RNA Alphabet. *J. Am. Chem. Soc.* **2015**, *137* (46), 14602–14605. https://doi.org/10.1021/jacs.5b10420.

(5) Shin, D.; Sinkeldam, R. W.; Tor, Y. Emissive RNA Alphabet. *J. Am. Chem. Soc.* **2011**, *133* (38), 14912–14915. https://doi.org/10.1021/ja206095a.

(6) Rovira, A. R.; Fin, A.; Tor, Y. Expanding a Fluorescent RNA Alphabet: Synthesis, Photophysics and Utility of Isothiazole-Derived Purine Nucleoside Surrogates. *Chem. Sci.* **2017**, *8* (4), 2983–2993. https://doi.org/10.1039/C6SC05354H.

(7) Hallé, F.; Fin, A.; Rovira, A. R.; Tor, Y. Emissive Synthetic Cofactors: Enzymatic Interconversions of TzA Analogues of ATP, NAD+, NADH, NADP+, and NADPH. *Angewandte Chemie International Edition* **2018**, *57* (4), 1087–1090. https://doi.org/10.1002/anie.201711935.

(8) Li, Y.; Ludford III, P. T.; Fin, A.; Rovira, A. R.; Tor, Y. Enzymatic Syntheses and Applications of Fluorescent Cyclic Dinucleotides. *Chemistry – A European Journal* **2020**, *26* (27), 6076–6084. https://doi.org/10.1002/chem.202001194.

(9) Li, Y.; Fin, A.; Rovira, A. R.; Su, Y.; Dippel, A. B.; Valderrama, J. A.; Riestra, A. M.; Nizet, V.; Hammond, M. C.; Tor, Y. Tuning the Innate Immune Response to Cyclic Dinucleotides by Using Atomic Mutagenesis. *ChemBioChem* **2020**, *21* (18), 2595–2598. https://doi.org/10.1002/cbic.202000162.

(10)   Bucardo, M. S.; Wu, Y.; Ludford, P. T.; Li, Y.; Fin, A.; Tor, Y. Real-Time Monitoring of Human Guanine Deaminase Activity by an Emissive Guanine Analog. *ACS Chem. Biol.* **2021**, *16* (7), 1208–1214. https://doi.org/10.1021/acschembio.1c00232.

(11)   Grimm, J. B.; Muthusamy, A. K.; Liang, Y.; Brown, T. A.; Lemon, W. C.; Patel, R.; Lu, R.; Macklin, J. J.; Keller, P. J.; Ji, N.; Lavis, L. D. A General Method to Fine-Tune Fluorophores for Live-Cell and in Vivo Imaging. *Nat Methods* **2017**, *14* (10), 987–994. https://doi.org/10.1038/nmeth.4403.

(12)   Grimm, J. B.; English, B. P.; Chen, J.; Slaughter, J. P.; Zhang, Z.; Revyakin, A.; Patel, R.; Macklin, J. J.; Normanno, D.; Singer, R. H.; Lionnet, T.; Lavis, L. D. A General Method to Improve Fluorophores for Live-Cell and Single-Molecule Microscopy. *Nature Methods* **2015**, *12* (3), 244–250. https://doi.org/10.1038/nmeth.3256.

(13)   Grimm, J. B.; Xie, L.; Casler, J. C.; Patel, R.; Tkachuk, A. N.; Falco, N.; Choi, H.; Lippincott-Schwartz, J.; Brown, T. A.; Glick, B. S.; Liu, Z.; Lavis, L. D. A General Method to Improve Fluorophores Using Deuterated Auxochromes. *JACS Au* **2021**, *1* (5), 690–696. https://doi.org/10.1021/jacsau.1c00006.

(14)   Reichardt, C. Solvatochromic Dyes as Solvent Polarity Indicators. *Chem. Rev.* **1994**, *94* (8), 2319–2358. https://doi.org/10.1021/cr00032a005.

(15)   Grabowski, Z. R.; Rotkiewicz, K.; Rettig, W. Structural Changes Accompanying Intramolecular Electron Transfer: Focus on Twisted Intramolecular Charge-Transfer States and Structures. *Chem. Rev.* **2003**, *103* (10), 3899–4032. https://doi.org/10.1021/cr940745l.

(16)   Hansch, Corwin.; Leo, A.; Taft, R. W. A Survey of Hammett Substituent Constants and Resonance and Field Parameters. *Chem. Rev.* **1991**, *91* (2), 165–195. https://doi.org/10.1021/cr00002a004.

(17)   Maillard, J.; Klehs, K.; Rumble, C.; Vauthey, E.; Heilemann, M.; Fürstenberg, A. Universal Quenching of Common Fluorescent Probes by Water and Alcohols. *Chem. Sci.* **2021**, *12* (4), 1352–1362. https://doi.org/10.1039/D0SC05431C.

2.9 Supplementary Information

2.9.1 Materials and Methods

Reagents were purchased from Sigma-Aldrich, TCI, Spectrum, Acros, Fisher Scientific, and VWR and were used without further purification unless otherwise specified. Methyl 4-amino thiophene 3-carboxylate hydrochloride was purchased from Matrix Scientific. Azetidine hydrochloride and difluoroazetidine hydrochloride were purchased from Synthonix. Solvents were purchased from Sigma-Aldrich and Fisher Scientific, and dried by standard techniques. NMR solvents were purchased from Cambridge Isotope Laboratories (Andover, MA). All reactions were monitored with analytical TLC (Merck Kieselgel 60 F254). Column chromatography was carried out with Teledyne ISCO Combiflash Rf with silica gel particle size 40-63 μm. NMR spectra were obtained on Jeol ECA 500 MHz and Varian VX 500 MHz at room temperature, except for compounds **6a** (taken at 80 °C) and **11** (taken at 60 °C) . Mass spectra were obtained on an Agilent 6230 HR-ESI-TOF MS at the Molecular Mass Spectrometry Facility at the UCSD Chemistry and Biochemistry Department.

2.9.2 Synthetic procedures

Starting precursors **7b** and **10** were synthesized based on previously published procedures.[1,2]

2.9.2.1 Synthesis of dimethyladenine analogue 2a
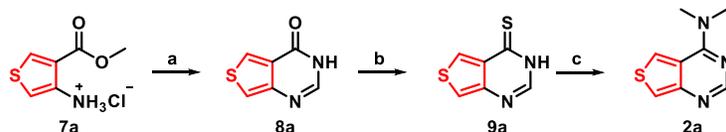
Scheme 2.3. Synthesis of thieno- derivative **2a**. Formamidine acetate, EtOH, reflux, ON, 78%. (b) P_2S_5, Pyridine, 110 °C, 2 h, 83%. (c) 2M $HNMe_2$ in MeOH, 60 °C, ON, 64%.

Thieno[3,4-*d*]pyrimidin-4(3*H*)-one (8a)

To a 1 L round-bottom flask was added methyl 4-amino thiophene 3-carboxylate hydrochloride **7a** (50 g, 256 mmol) and formamidine acetate (80.7 g, 775 mmol, 3 eq). EtOH (500 ml) was added, and the solution was refluxed at 105 °C overnight. A precipitate gradually forms and following reaction completion the solution is cooled and filtered. The solid is washed with water and EtOH, then dried over P_2O_5 to yield the desired product as an off-greenish fluffy solid (30.7 g, 78%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.66 (s, 1H), 8.47 (d, *J* = 3.2 Hz, 1H), 7.80 – 7.75 (m, 2H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 158.26, 148.96, 144.13, 128.32, 126.84, 118.73. ESI-HRMS calculated for $C_6H_5N_2OS$ [M+H]$^+$ 153.0117, found 153.0114.

Thieno[3,4-*d*]pyrimidine-4(*3H*)-thione (9a).

To a flame-dried 100 ml round bottom flask purged with argon was added compound **8a** (15 g, 98.7 mmol) and phosphorous pentasulfide (65.8 g, 296 mmol, 3 eq). Anhydrous pyridine (750 mL) was added and the heterogenous solution was stirred at room temperature for 10 minutes, then heated to 105 °C and stirred for 3 hours. If the reaction did not reach completion within the allotted time, half an equivalent of phosphorous pentasulfide was added every half hour until the starting material was fully consumed. Once finished, the reaction was cooled to room temperature and evaporated to dryness. 750 ml of water was added to the crude to form an orange suspension. A solution of 40% NaOH in water was then added dropwise until the pH was 12-13 and the suspension cleared. 37% HCl was then added dropwise until the solution was neutral and a yellowish precipitate formed that was filtered off and dried over P_2O_5 for 2 hours (14.5g, 83%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 13.20 (s, 2H), 8.55 (d, *J* = 3.3 Hz, 2H), 7.91 (d,

J = 3.3 Hz, 1H), 7.82 (s, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 182.92, 143.85, 141.92, 133.31, 130.74, 120.12. ESI-HRMS calculated for $C_6H_4N_2S_2$ [M+H]$^+$ 168.9889, found 168.9887.

N,N-dimethylthieno[3,4-*d*]pyrimidin-4-amine (2a)

The thionated substate **9a** (0.15 g, 0.9 mmol, 1 eq) was placed in a heavy wall cylindrical pressure vessel along with a solution of 2M dimethylamine in MeOH (7 ml, 14 mmol, 15 eq). The vessel was sealed, and the suspension was stirred at 65 °C overnight. The solution gradually turns clear, and following reaction completion the vessel was then cooled, and the solution was evaporated to dryness. Column chromatography of the crude residue with a gradient of 0-5% MeOH in DCM yielded a cream-colored solid (0.1 g, 64%). ^{1}H NMR (500 MHz, Chloroform-d) δ 8.46 (d, J = 3.0 Hz, 1H), 8.34 – 8.31 (m, 1H), 8.29 (d, J = 3.4 Hz, 1H), 3.84 (d, J = 3.6 Hz, 3H), 3.64 (d, J = 3.6 Hz, 3H). ^{13}C NMR (126 MHz, Chloroform-d) δ 158.51, 149.04, 138.52, 127.44, 118.32, 112.39, 42.22, 41.66. ESI-HRMS calculated for $C_8H_{10}N_3S$ [M+H]$^+$ 180.0590, found 180.0591.

2.9.2.2 Synthesis of azetidine-analogues 3a–6a
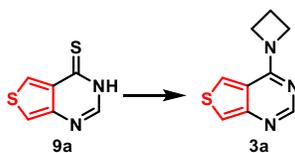
4-(azetidin-1-yl)thieno[3,4-*d*]pyrimidine (3a)

Compound **9a** (100 mg, 0.6 mmol, 1 eq) was placed in a 25 ml round-bottom flask and MeOH (5 ml) was added creating a suspension. Azetidine hydrochloride (0.94 g, 10 mmol, 17 eq) was added followed by DIPEA (1.9 g, 15 mmol, 2.6 ml, 25 eq). The solution was allowed to stir at 70 °C overnight. The reaction was cooled to room temperature and evaporated to dryness.

The crude was dissolved in DCM and extracted with saturated sodium bicarbonate followed by brine. The organic phase was dried with sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude product with a gradient of 0-5% MeOH in DCM yielded the desired compound **3a** as a yellowish solid (40 mg, 38%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.25 (d, J = 3.1 Hz, 1H), 8.02 (s, 1H), 7.68 (d, J = 3.1 Hz, 1H), 4.63 (s, 2H), 4.19 (s, 2H), 2.41 (q, J = 7.6 Hz, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 156.84, 154.22, 151.06, 123.18, 119.47, 114.66, 54.17 – 52.15 (m), 16.92. ESI-HRMS calculated for $C_9H_{10}N_3S$ [M+H]$^+$ 192.0590, found 192.0591.
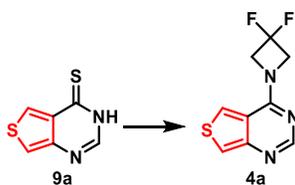
4-(3,3-difluoroazetidin-1-yl)thieno[3,4-d]pyrimidine (4a)

Compound **9a** (0.11 g, 0.7 mmol) was placed in a 25 ml round-bottom flask and MeOH (5 ml) was added creating a suspension. 2,2-Difluoroazetidine hydrochloride (1.3 g, 10 mmol, 20 eq) was added followed by DIPEA (1.9 g, 2.6 ml, 15 mmol, 23 eq). The solution was allowed to stir at 70 °C overnight. The reaction was cooled to room temperature and evaporated to dryness. The crude was dissolved in DCM and extracted with saturated sodium bicarbonate followed by brine. The organic phase was dried with sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude product with a gradient of 0-4% MeOH in DCM yielded the desired compound **4a** as a yellowish solid (78 mg, 53%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.33 (d, J = 3.1 Hz, 1H), 8.17 (s, 1H), 7.86 (d, J = 3.2 Hz, 1H), 4.97 – 4.77 (m, 4H). ^{13}C NMR

(126 MHz, DMSO-d_6) δ 157.30 (t, J = 4.3 Hz), 153.66, 150.94, 123.26, 119.43, 119.37, 117.27, 116.00, 115.11. ESI-HRMS calculated for $C_9H_8F_2N_3S$ [M+H]$^+$ 228.0402, found 228.0403.
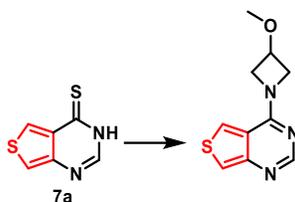
4-(3-methoxyazetidin-1-yl)thieno[3,4-*d*]pyrimidine (5a)

Compound **7a** (100 mg, 0.6 mmol, 1 eq) was placed in a 25 ml round-bottom flask and MeOH (5 ml) was added creating a suspension. 3-methoxyazetidine hydrochloride (1.5 g, 12 mmol, 20 eq) was added followed by DIPEA (2.3 g, 18 mmol, 3.1 ml, 30 eq). The solution was allowed to stir at 70 °C overnight. The reaction was cooled to room temperature and evaporated to dryness. The crude was dissolved in DCM and extracted with saturated sodium bicarbonate followed by brine. The organic phase was dried with sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude product with a gradient of 0-5% MeOH in DCM yielded the desired compound **5a** as a yellowish solid (40 mg, 30%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.30 (d, J = 3.1 Hz, 1H), 8.05 (s, 1H), 7.73 (d, J = 3.1 Hz, 1H), 4.81 (s, 1H), 4.38 (m, 3H), 3.99 (s, 1H), 3.26 (s, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 156.99, 154.11, 151.03, 123.24, 119.51, 114.99, 70.22, 56.03. ESI-HRMS calculated for $C_{10}H_{12}N_3OS$ [M+H]$_+$ 222.0696, found 222.0695.

53
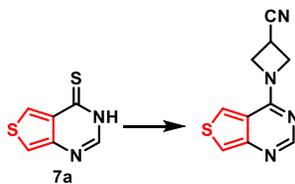
1-(thieno[3,4-*d*]pyrimidin-4-yl)azetidine-3-carbonitrile (6a)

Compound **7a** (100 mg, 0.6 mmol, 1 eq) was placed in a 25 ml round-bottom flask and MeOH (5 ml) was added creating a suspension. 3-azetidine carbonitrile hydrochloride (1.41 g, 12 mmol, 20 eq) was added followed by DIPEA (2.3 g, 18 mmol, 3.1 ml, 30 eq). The solution was allowed to stir at 70 °C overnight. The reaction was cooled to room temperature and evaporated to dryness. The crude was dissolved in DCM and extracted with saturated sodium bicarbonate followed by brine. The organic phase was dried with sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude product with a gradient of 0-5% MeOH in DCM yielded the desired compound **6a** as a yellowish solid (43 mg, 33%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.24 (d, J = 3.1 Hz, 1H), 8.11 (s, 1H), 7.75 (d, J = 3.1 Hz, 1H), 4.72 (t, J = 9.1 Hz, 2H), 4.64 – 4.54 (m, 2H), 4.04 – 3.94 (m, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 157.29, 153.88, 151.05, 123.23, 121.07, 115.61, 18.93. ESI-HRMS calculated for $C_{10}H_9N_4S$ [M+H]$^+$ 219.0542, found 219.0539.

2.9.2.3 Synthesis of isothiazolo[4,3-*d*] based dimethyladenine analogue 2b
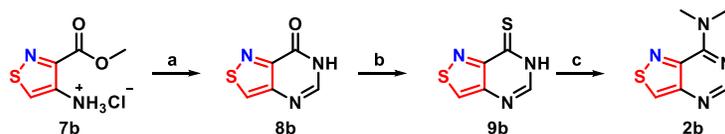
Scheme 2.4. Synthesis of isothiazolo-derivative **2b**. (a) Formamidine acetate, EtOH, reflux, ON, 76%. (b) Lawesson's reagent, Pyridine, 110 °C, 2 h, 92%. (c) 2M HNMe$_2$ in MeOH, 60 °C, ON, 69%.

54

Isothiazolo[4,3-*d*]pyrimidin-7(6*H*)-one (8b)

To a 250 ml round-bottom flask was added **7b** (1.2 g, 6.4 mmol) and formamidine acetate (2 g, 19.2 mmol, 3 eq). 200 proof ethanol (14 ml) was added, and the solution was refluxed at 105 °C overnight. A precipitate gradually forms and following reaction completion the solution is cooled and filtered. The precipitate was filtered off and washed with cold EtOH yielding a yellowish solid that was dried over P_2O_5 (0.74 g, 76%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.15 (s, 1H), 9.47 (s, 1H), 7.98 (s, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 156.44, 151.68, 148.89, 146.38, 145.27. ESI-HRMS calculated for $C_5H_3N_3OS$ [M+Na]$^+$ 175.9889, found 175.9891.

Isothiazolo[4,3-*d*]pyrimidine-7(6*H*)-thione (9b)

To a 100 ml round-bottom flask that was flame-dried and argon purged was added compound **8b** (0.88 g, 5.7 mmol) and anhydrous pyridine (25 ml), Lawesson's reagent (4.6 g, 11.5 mmol, 2 eq) was added and the solution was stirred at 110 °C for 2 hours. Upon reaction completion the solution was cooled in an ice bath and concentrated to dryness. The crude was resuspended in water (6 ml) and placed in an ice bath. 40% NaOH in water was added dropwise until the solution was sufficiently clear and the suspension had cleared. After stirring an additional 5 minutes, the solution was solution by adding 37% HCl dropwise. The resulting precipitate was filtered off and washed with water, yielding the desired product **9b** as a desired solid (0.89 g, 92%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.56 (s, 1H), 7.99 (s, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 181.16, 155.59, 148.43, 144.12, 143.32. ESI-HRMS calculated for $C_9H_8F_2N_3S$ [M-H]$^-$ 167.9696, found 167.9696.

N,N-dimethylisothiazolo[4,3-d]pyrimidin-7-amine (2b)

The thionated substrate **9b** (0.11 g, 0.6 mmol) was placed in a 50 ml round-bottom flask. A solution of 2M HNMe₂ in MeOH (15 ml, 30 mmol, 50 eq) was added and the reaction was stirred at 60 °C overnight. The solution was then evaporated to dryness and the crude was subjected to column chromatography using a gradient of 0-5% MeOH in DCM to yield an orange-colored solid (80 mg, 69%). ^{1}H NMR (300 MHz, DMSO-d_6) δ 9.47 (s, 0H), 8.26 (s, 0H), 3.91 (s, 1H), 3.29 (s, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 155.32 (d, $J = 1.9$ Hz), 154.55, 154.28, 151.32 (d, $J = 2.3$ Hz), 146.62, 143.87 – 143.67 (m). ESI-HRMS calculated for $C_7H_9N_4S$ [M+H]$^+$ 181.0543, found 181.0542.

2.9.2.4 Synthesis of azetidine analogues 3b–6b
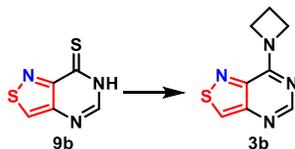
7-(azetidin-1-yl)isothiazolo[4,3-d]pyrimidine (3b)

The thionated substrate **9b** (0.1 g, 0.6 mmol) was placed in a 25 ml round-bottom flask and MeOH (5 ml) was added to create a suspension. Azetidine hydrochloride (1.1 g, 12 mmol, 20 eq) was added followed by DIPEA (2.3 g, 18 mmol, 3.1 ml, 30 eq). The solution was allowed to stir at 60 °C overnight. Following reaction completion, the solution was cooled to room temperature and evaporated to dryness. The crude was then extracted between DCM and saturated sodium bicarbonate solution. The aqueous phase was washed 3 times with DCM and the organic phases were combined, dried with sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude residue using a gradient of 0-3% MeOH in DCM

yielded the desired product **3b** as a yellowish solid (68 mg, 60%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.40 (s, 1H), 8.18 (s, 1H), 4.75 (t, J = 7.7 Hz, 3H), 4.23 (t, J = 7.7 Hz, 3H), 2.44 (dd, J = 15.6, 7.9 Hz, 5H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 155.00, 154.46, 149.86, 146.15, 143.68, 54.57, 50.81, 17.43. ESI-HRMS calculated for $C_7H_9N_4S$ [M+H]$^+$ 193.0542, found 193.0541.
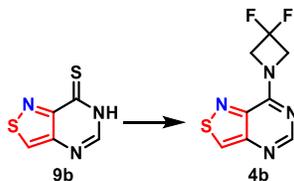
7-(3,3-difluoroazetidin-1-yl)isothiazolo[4,3-*d*]pyrimidine (4b)

Compound **9b** (0.1 g 0.6 mmol) was placed in a 25 ml round-bottom flask and MeOH (5 ml) was added to create a suspension. 2,2-difluoroazetidine hydrochloride (1.5 g, 12 mmol, 20 eq) was added followed by DIPEA (2.3 g, 18 mmol, 3.1 ml, 30 eq). The reaction was allowed to stir at 60 °C overnight. Following reaction completion, the solution was evaporated to dryness. Following reaction completion, the solution was cooled to room temperature and evaporated to dryness. The crude was then extracted between DCM and saturated sodium bicarbonate solution. The aqueous phase was washed 3 times with DCM and the organic phases were combined, dried with sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude residue using a gradient of 0-3% MeOH in DCM yielded the desired product **4b** as a yellowish solid (94 mg, 70%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.57 (s, 1H), 8.33 (s, 1H), 5.15 (s, 3H), 4.68 (s, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 154.92 (t, J = 4.5 Hz), 154.56, 149.84, 145.82, 145.27, 119.58, 117.42, 115.26. ^{19}F NMR (471 MHz, DMSO-d_6) δ -99.57 (p, J = 12.5 Hz). ESI-HRMS calculated for $C_8H_6F_2N_4S$ [M+H]$^+$ 229.0354, found 229.0355.
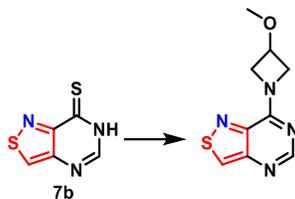
4-(3-methoxyazetidin-1-yl)thieno[3,4-*d*]pyrimidine (5b)

Compound **7b** (61 mg, 0.36 mmol, 1 eq) was placed in a 25 ml round-bottom flask and MeOH (3 ml) was added creating a suspension. 3-methoxyazetidine hydrochloride (0.9 g, 7.3 mmol, 20 eq) was added followed by DIPEA (1.4 g, 11 mmol, 1.9 ml, 30 eq). The solution was allowed to stir at 70 °C overnight. The reaction was cooled to room temperature and evaporated to dryness. The crude was dissolved in DCM and extracted with saturated sodium bicarbonate followed by brine. The organic phase was dried with sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude product with a gradient of 0-5% MeOH in DCM yielded the desired compound **5b** as a yellowish solid (27 mg, 33%) following recrystallization from DCM/hexanes. ^{1}H NMR (500 MHz, DMSO- d_6) δ 9.44 (s, 1H), 8.21 (s, 1H), 4.96 (dd, J = 10.4, 6.1 Hz, 1H), 4.51 (d, J = 11.0 Hz, 1H), 4.45 – 4.37 (m, 2H), 4.03 (d, J = 10.6 Hz, 1H), 3.26 (s, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 154.92, 154.64, 149.87, 146.11, 144.12, 70.55, 61.28, 57.69, 56.04. ESI-HRMS calculated for $C_9H_{11}N_4OS$ [M+H]$^+$ 223.0648, found 223.0649.
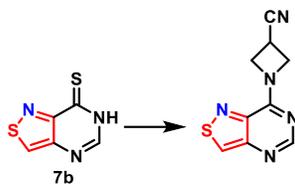
1-(isothiazolo[4,3-d]pyrimidin-7-yl)azetidine-3-carbonitrile (6b)

Compound **7b** (81 mg, 0.5 mmol, 1 eq) was placed in a 25 ml round-bottom flask and MeOH (4 ml) was added creating a suspension. 3-azetidine carbonitrile hydrochloride (1.14 g, 10 mmol, 20 eq) was added followed by DIPEA (1.86 g, 14 mmol, 2.5 ml, 30 eq). The solution was allowed to stir at 70 °C overnight. The reaction was cooled to room temperature and evaporated to dryness. The crude was dissolved in DCM and extracted with saturated sodium bicarbonate followed by brine. The organic phase was dried with sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude product with a gradient of 0-5% MeOH in DCM yielded the desired compound **6b** as a yellowish solid (43 mg, 41%) following recrystallization from Et$_2$O/hexanes. ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.51 (s, 1H), 8.27 (s, 1H), 5.07 – 4.85 (m, 2H), 4.55 – 4.36 (m, 2H), 4.02 (tt, J = 9.0, 6.1 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 154.73 (s), 154.69 (s), 149.82 (s), 145.72 (s), 144.70 (s), 120.95 (s), 57.30 (s), 53.83 (s), 19.38 (s). ESI-HRMS calculated for C$_9$H$_8$N$_5$S [M+H]$^+$ 218.0495, found 218.0497.

2.9.2.5 Synthesis of nucleoside 11
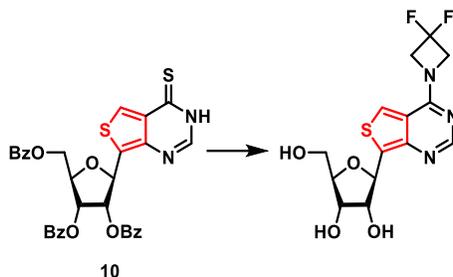
(*2R*,*5R*)-2-(4-(3,3-difluoroazetidin-1-yl)thieno[3,4-*d*]pyrimidin-7-yl)-5-(hydroxymethyl)

tetrahydrofuran-3,4-diol (11)

Compound **10** (150 mg, 0.24 mmol, 1 eq) was placed in a 25 ml round-bottom flask. 5 ml of a 1:1 mixture of satd. NH_4OH: satd. $MeNH_2$ was added. The solution was allowed to stir at room temperature for 15 minutes, then evaporated to dryness. The reaction was cooled to room temperature and evaporated to dryness. The crude riboside was coevaporated with MeOH three times followed by DCM twice. Once dried, the crude was dissolved in MeOH and 3,3-dilfuoroazetidine HCl (632 mg, 4.9 mmol, 20 eq) was added followed by DIPEA (950 mg, 7.35 mmol, 1.3 ml, 30 eq). The reaction was allowed to stir at 60 °C overnight. The solution was evaporated to dryness and the crude was resuspended in MeOH. 1.5 g K_2CO_3 was added, and the heterogeneous mixture was stirred for 15 minutes, then filtered. The filtrate was evaporated and loaded onto silica. Column chromatography using a gradient of 0-7% MeOH in DCM yielded the desired product as a gold solid (65 mg, extracted with saturated sodium bicarbonate followed by brine. The organic phase was dried with sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude product with a gradient of 0-5% MeOH in DCM yielded the desired compound **5b** as a yellowish solid (54 mg, 61%). The compound was then triturated with ethanol yielding a white solid (34 mg, 24%) suitable in purity for photophysical studies. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.23 (s, 1H), 8.15 (s, 1H), 5.35 (d, *J* = 6.5 Hz, 1H), 5.05 (s, 1H),

4.87 (t, J = 12.2 Hz, 4H), 4.75 (s, 1H), 4.13 (dd, J = 6.5, 5.1 Hz, 1H), 4.01 – 3.92 (m, 1H), 3.85 (q, J = 4.1 Hz, 1H), 3.56 (dd, J = 11.7, 4.0 Hz, 1H), 3.50 (dd, J = 11.7, 4.5 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 157.43 (t, J = 4.6 Hz), 153.24, 147.22, 133.32, 121.52, 119.95, 117.35, 85.74, 77.98, 77.86, 72.29, 62.79. ESI-HRMS calculated for $C_{14}H_{16}F_2N_3O_4S$ [M+H]$^+$ 360.0824, found 360.0828.

2.9.3 Structural analysis

2.9.3.1 Experimental summary

The single crystal X-ray diffraction studies were carried out on a Bruker APEX II Ultra CCD diffractometer equipped with Cu K$_\alpha$ radiation (λ = 1.54178). A 0.170 x 0.105 x 0.025 mm orange crystal was mounted on a Cryoloop with Paratone oil. Data were collected in a nitrogen gas stream at 100(2) K using ϕ and ϖ scans. Crystal-to-detector distance was 45 mm using exposure 1.0s s with a scan width of 0.75°. Data collection was 100.0% complete to 25.242° in θ.

A total of 6956 reflections were collected covering the indices, -22<=h<=23, -5<=k<4, -17<=l<=17. 2074 reflections were found to be symmetry independent, with a R$_{int}$ of 0.0663. Indexing and unit cell refinement indicated a Primitive, Orthorhombic lattice. The space group was found to be Pca2$_1$. The data were integrated using the Bruker SAINT Software program and scaled using the SADABS software program. Solution by direct methods (SHELXT) produced a complete phasing model consistent with the proposed structure.

All nonhydrogen atoms were refined anisotropically by full-matrix least-squares (SHELXL-2014). All carbon bonded hydrogen atoms were placed using a riding model. Their positions were constrained relative to their parent atom using the appropriate HFIX command in SHELXL-2014. All other hydrogen atoms (H-bonding) were located in the difference map (N-H, O-H). There positions were refined using "riding" model. Crystallographic data are summarized in table 2.4-2.7.

2.9.3.2 X-ray crystal structures

Table 2.4. Crystal data and structure refinement for compound 2a

Report date	2018-11-13
Identification code	tor_kh1027_3
Empirical formula	C8 H9 N3 S
Molecular formula	C8 H9 N3 S
Formula weight	179.24
Temperature	100.0 K
Wavelength	0.71073 Å
Crystal system	Triclinic
Space group	P-1
Unit cell dimensions	a = 7.1286(4) Å α= 92.0730(10)°.
	b = 7.4500(4) Å β= 103.7770(10)°.
	c = 7.8894(4) Å g = 102.2880(10)°.
Volume	395.94(4) Å3
Z	2
Density (calculated)	1.503 Mg/m3
Absorption coefficient	t0.348 mm-1
F(000)	188
Crystal size	0.18 x 0.04 x 0.04 mm3
Crystal color, habit	colorless plank
Theta range for data collection	2.669 to 26.368°.
Index ranges	-8<=h<=8, -9<=k<=9, -8<=l<=9
Reflections collected	4560
Independent reflections	1625 [R(int) = 0.0281]
Completeness to theta = 25.242°	100.0 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.5628 and 0.5205
Refinement method	Full-matrix least-squares on F2
Data / restraints / parameters	1625 / 0 / 111
Goodness-of-fit on F2	1.052
Final R indices [I>2sigma(I)]	R1 = 0.0293, wR2 = 0.0782
R indices (all data)	R1 = 0.0308, wR2 = 0.0797
Largest diff. peak and hole	0.519 and -0.195 e.Å-

63

Table 2.5. Crystal data and structure refinement for compound **2b.**

Report date	2019-08-08
Identification code	tor_mmetza
Empirical formula	C7 H8 N4 S
Molecular formula	C7 H8 N4 S
Formula weight	180.23
Temperature	100.0 K
Wavelength	0.71073 Å
Crystal system	Triclinic
Space group	P-1
Unit cell dimensions	a = 7.6743(3) Å $\quad \alpha= 95.9810(10)°$.
	b = 7.7814(3) Å $\quad \beta= 98.6310(10)°$.
	c = 13.3599(6) Å $\quad \gamma = 91.5710(10)°$.
Volume	783.73(6) Å3
Z	4
Density (calculated)	1.527 Mg/m3
Absorption coefficient	0.355 mm-1
F(000)	376
Crystal size	0.125 x 0.08 x 0.06 mm3
Crystal color, habit	light yellow block
Theta range for data collection	1.551 to 26.369°.
Index ranges	-9<=h<=9, -9<=k<=9, -16<=l<=16
Reflections collected	14003
Independent reflections	3217 [R(int) = 0.0400]
Completeness to theta = 25.242°	100.0 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.7454 and 0.6839
Refinement method	Full-matrix least-squares on F2
Data / restraints / parameters	3217 / 0 / 221
Goodness-of-fit on F2	1.051
Final R indices [I>2sigma(I)]	R1 = 0.0329, wR2 = 0.0814
R indices (all data)	R1 = 0.0370, wR2 = 0.0841
Largest diff. peak and hole	0.399 and -0.266 e.Å-3

Table 2.6. Crystal data and structure refinement for compound **5a**.

Report date	2022-01-10
Identification code	tor157
Empirical formula	C10 H11 N3 O S
Molecular formula	C10 H11 N3 O S
Formula weight	221.28
Temperature	100.0 K
Wavelength	0.71073 Å
Crystal system	Orthorhombic
Space group	Pca2$_1$
Unit cell dimensions	a = 18.4848(10) Å $\quad$ α= 90°.
	b = 3.9998(2) Å $\quad$ β= 90°.
	c = 13.4934(8) Å $\quad$ γ = 90°.
Volume	997.64(9) Å3
Z	4
Density (calculated)	1.473 Mg/m3
Absorption coefficient	0.299 mm-1
F(000)	464
Crystal size	0.175 x 0.12 x 0.1 mm3
Crystal color, habit	light yellow plank
Theta range for data collection	2.204 to 26.736°.
Index ranges	-22<=h<=23, -5<=k<=4, -17<=l<=17
Reflections collected	6956
Independent reflections	2074 [R(int) = 0.0663]
Completeness to theta = 25.242°	100.0 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.6947 and 0.6593
Refinement method	Full-matrix least-squares on F2
Data / restraints / parameters	2074 / 1 / 137
Goodness-of-fit on F2	1.063
Final R indices [I>2sigma(I)]	R1 = 0.0350, wR2 = 0.0901
R indices (all data)	R1 = 0.0362, wR2 = 0.0913
Absolute structure parameter	0.04(7)
Largest diff. peak and hole	0.240 and -0.199 e.Å-
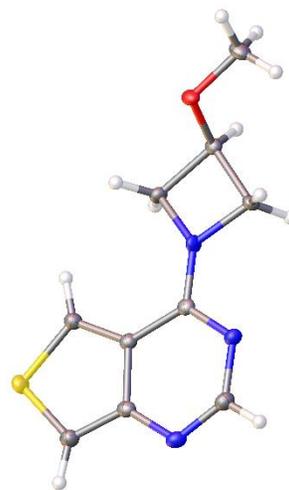
Table 2.7. Crystal data and structure refinement for compound **6a**

Report date	2022-01-10
Identification code	tor158
Empirical formula	C10 H8 N4 S
Molecular formula	C10 H8 N4 S
Formula weight	216.26
Temperature	100.0 K
Wavelength	0.71073 Å
Crystal system	Triclinic
Space group	P-1
Unit cell dimensions	a = 7.6117(7) Å α= 78.693(2)°.
	b = 8.1553(8) Å β= 73.317(2)°.
	c = 8.2163(7) Å γ = 84.818(2)°.
Volume	478.77(8) Å3
Z	2
Density (calculated)	1.500 Mg/m3
Absorption coefficient	0.305 mm-1
F(000)	224
Crystal size	0.14 x 0.125 x 0.1 mm3
Crystal color, habit	colorless block
Theta range for data collection	2.549 to 26.024°.
Index ranges	-9<=h<=9, -10<=k<=10, -10<=l<=10
Reflections collected	5661
Independent reflections	1884 [R(int) = 0.0316]
Completeness to theta = 25.242°	100.0 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.5626 and 0.5337
Refinement method	Full-matrix least-squares on F2
Data / restraints / parameters	1884 / 0 / 136
Goodness-of-fit on F2	1.046
Final R indices [I>2sigma(I)]	R1 = 0.0311, wR2 = 0.0817
R indices (all data)	R1 = 0.0333, wR2 = 0.0834
Largest diff. peak and hole	0.371 and -0.187 e.Å-3
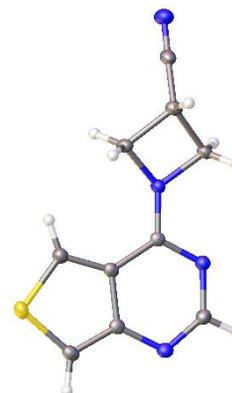
2.9.4 Absorption and emission spectroscopy assays and data

Table 2.8. Photophysical and chemical properties of thieno[3,4-*d*]pyrimidine and isothiazolo[4,3-*d*]pyrimidine nucleobase and nucleoside analogs.

	solvent	λ_{abs} $(\varepsilon)^a$	λ_{em} $(\Phi)^a$	$\Phi\varepsilon$	Stokes shift^a	polarity sensitivity^b
2a	water	346 (1.03±0.01)	411 (0.0005±0.0002)	5	3.80±0.13	-0.69±5.5
	dioxane	353 (0.94±0.02)	416 (0.47±0.01)	4392	4.33±0.06	
	D$_2$O	346 (0.98±0.01)	416 (0.0015±0.0001)	14	4.77±0.10	
3a	water	344 (1.11±0.04)	407 (0.06±0.01)	715	4.40±0.09	53±9
	dioxane	351 (1.27±0.01)	412 (0.75±0.01)	9484	4.23±0.02	
	D$_2$O	345 (1.00±0.01)	406 (0.07±0.01)	687	4.41±0.12	
4a	water	344 (0.90±0.01)	419 (0.64±0.03)	5766	5.14±0.03	81±17
	dioxane	343 (0.85±0.02)	419 (0.77±0.04)	6587	5.26±0.08	
	D$_2$O	343 (0.84±0.02)	418 (0.57±0.01)	4826	5.29±0.05	
5a	water	344 (1.12±0.01)	415 (0.13±0.01)	1421	4.99±0.14	80±19
	dioxane	349 (1.03±0.02)	411 (0.70±0.01)	7218	4.28±0.05	
	D$_2$O	344 (1.11±0.01)	416 (0.15±0.01)	1728	5.00±0.10	
6a	water	345 (1.30±0.01)	421 (0.23±0.01)	2951	5.17±0.04	110±23
	dioxane	345 (1.19±0.01)	409 (0.64±0.01)	7579	4.52±0.12	
	D$_2$O	344 (1.27±0.01)	416 (0.33±0.01)	4164	5.03±0.07	
2b	water	350 (1.17±0.26)	412 (0.0001±0.0001)	1.6	n/a	n/a
	dioxane	358 (1.03±0.02)	418 (0.0010±0.0001)	14	n/a	
	D$_2$O	349 (1.11±0.02)	417 (0.0001±0.0001)	1.1	n/a	
3b	water	352 (1.22±0.05)	416 (0.0030±0.0001)	33	4.38±0.10	-17±4
	dioxane	358 (1.05±0.49)	415 (0.13±0.01)	1192	3.84±0.13	
	D$_2$O	349 (1.23±0.01)	415 (0.0030±0.0001)	41	4.51±0.05	
4b	water	341 (1.21±0.01)	403 (0.064±0.001)	770	4.54±0.06	11±8
	dioxane	348 (1.03±0.01)	407 (0.020±0.001)	173	4.19±0.08	
	D$_2$O	341 (1.22±0.04)	403 (0.073±0.001)	884	4.49±0.06	
5b	water	346 (1.11±0.02)	405 (0.007±0.001)	77	4.12±0.05	-2±8
	dioxane	356 (0.96±0.01)	413 (0.14±0.01)	821	3.84±0.12	
	D$_2$O	346 (1.06±0.01)	408 (0.010±0.01)	102	4.47±0.05	

Table 2.9. Photophysical and chemical properties of thieno[3,4-*d*]pyrimidine and isothiazolo[4,3-*d*]pyrimidine nucleobase and nucleoside analogs, continued.

	solvent	λ_{abs} (ε)[a]	λ_{em} (Φ)[a]	$\Phi\varepsilon$	Stokes shift[a]	polarity sensitivity[b]
6b	water	344 (1.11±0.01)	406 (0.03±0.01)	340	4.45±0.09	34±4
	dioxane	352 (0.94±0.01)	410 (0.06±0.01)	607	4.07±0.04	
	D$_2$O	343 (1.07±0.01)	406 (0.03±0.01)	293	4.49±0.12	
11	water	352 (0.94±0.03)	435 (0.30±0.04)	2838	5.41±0.19	
	dioxane	357 (0.94±0.01)	420 (0.17±0.01)	1444	4.21±0.01	
	D$_2$O	352 (0.95±0.01)	44 (0.47±0.01)	4486	5.23±0.05	

[a] λabs, ε, λem and Stokes shift are reported in nm, 10^3 M^{-1} cm^{-1}, nm and 10^3 cm^{-1} respectively. All the photophysical values reflect the average over three independent measurements. [b] Sensitivity to solvent polarity reported in cm^{-1}/ (kcal mol^{-1}) is equal to the slope of the linear fit in figure 2.29.

2.9.4.1 General

Spectroscopic grade DMSO, dioxane, ethylene glycol, glycerol, and methanol were obtained from Sigma-Aldrich and aqueous solutions were prepared with milliq water. All measurements were carried out in a 1 cm four-sided quartz cuvette from Helma.

Absorption spectra were measured on a Shimadzu UV-2450 spectrophotometer setting the slit at 1 nm and using a resolution of 0.5 nm. All the spectra were corrected for the blank. Emission and excitation spectra were measured on a Horiba Fluoromax-4 equipped with a cuvette holder with a stirring system setting the excitation slit at 1 nm and the emission slit at 3 nm, the resolution at 1 nm and the integration time 0.1 s. Emission spectra were measured exciting at 350 nm and excitation spectra were taken at 420 nm unless otherwise specified. Both instruments were equipped with a thermostat-controlled ethylene glycol-water bath fitted to specially designed cuvette holder and the temperature was kept at 25.00 ± 0.10 °C.

The compounds were dissolved in DMSO to prepare highly concentrated stock solutions (10X): **2a** (31.55 mM), **2b** (23.77 mM), **3a** (34.03 mM), **3b** (24.61 mM), **4a** (25.77 mM), **4b** (27.19 mM), **5a** (26.69 mM), **5b** (30.62 mM), **6a** (25.92 mM), **6b** (32.94 mM), and **11** (31.84

mM). For the determination of extinction coefficient and quantum yield, these 10X stock solutions were used. For these experiments, aliquots (10 μl) of the 10X solution were diluted with air-saturated solvents (2.990 ml). The solutions were then diluted by a factor of 2 in the corresponding solvent and remeasured. This dilution was repeated four times total, with the most dilute solution being used to measure the emission spectra and subsequent quantum yield. The absorbance values were corrected for the blank and the extinction coefficient was extrapolated for each wavelength using a linear best fit. For all other experiments, 1X concentrated stock solutions were utilized by dilution of the 10X stock solutions into spectroscopic grade DMSO. In a typical experiment, aliquots (10 μl) of the 1X solution were diluted with air-saturated solvents (2.990 ml). The solution were mixed with a pipette for 10 seconds and placed in the cuvette holder at 25.00 ± 0.10 °C for 3 minutes before spectra were recorded. All sample contain 0.3 v/v % of DMSO. For solutions in higher viscosity solvents (ethylene glycol, glycerol) the samples were checked under a 360 nm UV lamp to ensure the solutions were properly mixed.

2.9.4.2 Fluorescence quantum yield

The samples' concentration were adjusted to have an optical density lower than 0.1 at the excitation wavelength (λ_{ex}). The fluorescence quantum yield (Φ) were evaluated based on an external standard, quinine sulfate in 0.5 M H_2SO_4 (0.546, λ_{ex} 350 nm) by using equation 1.3:

$$\Phi = \Phi_{STD} \frac{I}{I_{STD}} \frac{OD_{STD}}{OD} \frac{n^2}{n^2_{STD}}$$

Where Φ_{STD} is the fluorescence quantum yield of the standard, I and I_{STD} are the integrated area of the emission band of the sample and the standard respectively, OD and OD_{STD} are the optical

density at the excitation wavelength for the sample and the standard respectively and n and n_{STD} are the solvent refractive index of the sample and the standard solutions respectively.

2.9.4.3 Absorption, emission, and excitation spectra of compounds 2–6a/b

Absorption, emission spectra and the corresponding excitation spectra were also recorded in water and dioxane varying the excitation wavelength or the emission wavelength of interest. Solutions were prepared as described above. The emission spectra were recorded upon excitation every 10 nm covering the main absorption band between 260 and 380 nm. The emission spectra were corrected for either the corresponding optical density of the sample or normalized to unit in intensity at the emission maxima. The excitation spectra were recorded by fixing the emission wavelength every 10 nm between 380 and 500 nm. The excitation spectra were then normalized to unit in intensity at the excitation maxima.
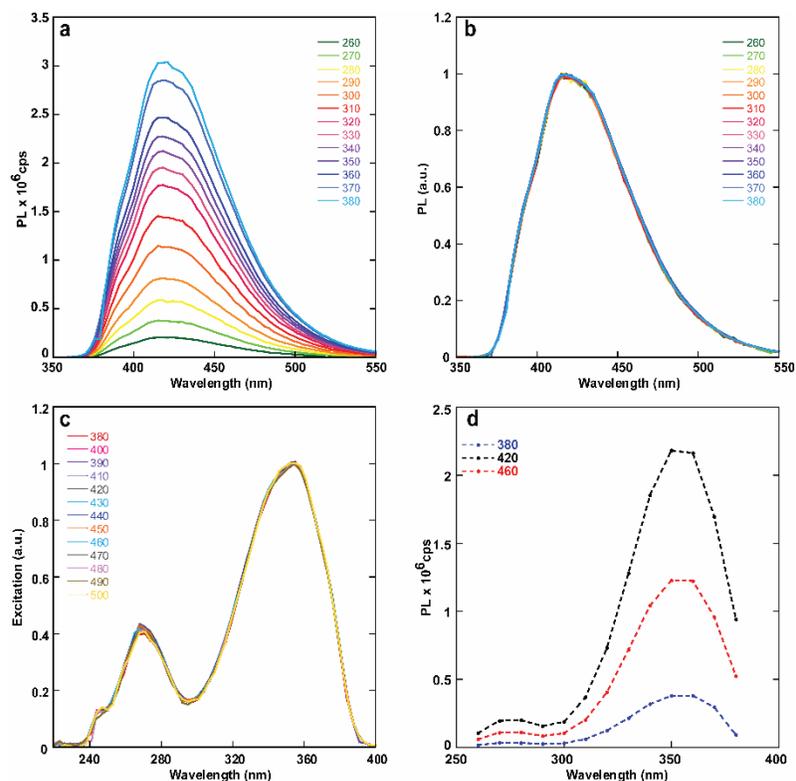
Figure 2.10. Emission spectra in dioxane of compound **2a** recorded upon excitation every 10 nm between 260-380 nm. The spectra were normalized for (a) the corresponding absorbance intensity or (b) to unit in intensity. (c) Excitation spectra of compound **2a** in dioxane recorded at select emission wavelengths, covering the emission band. (d) Extrapolated excitation spectra by plotting the emission spectra intensities at selected wavelengths upon excitation at different wavelengths of compound **2a** in dioxane.
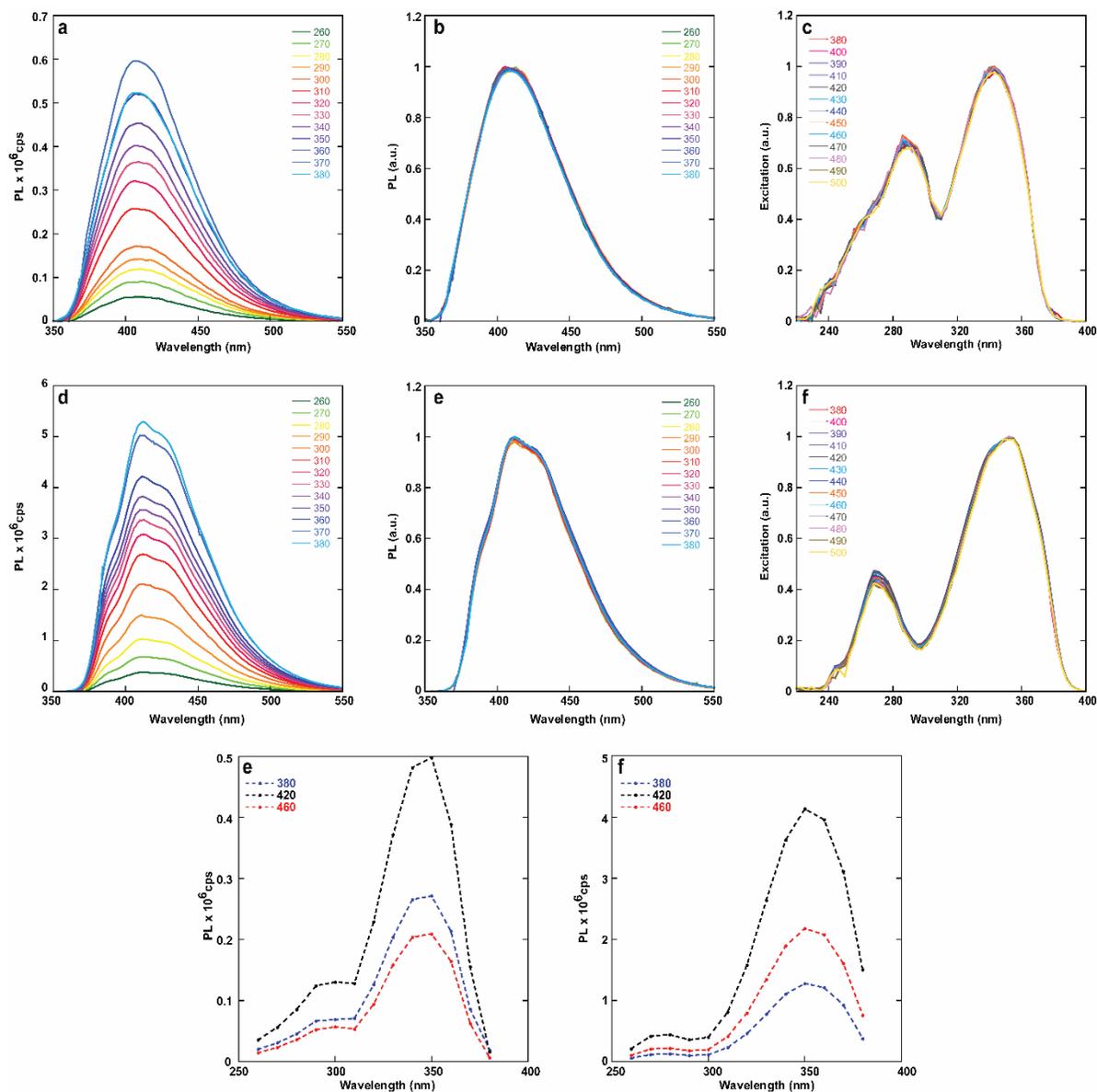
Figure 2.11. Spectra of compound **3a** recorded upon excitation every 10 nm between 260-380 nm in water (a-c) and dioxane (d-f). The spectra were recorded in and normalized for (a,d) the corresponding absorbance intensity or (b,e) to unit in intensity. (c) Excitation spectra of compound **3a** in water and (f) dioxane recorded at select emission wavelengths, covering the emission band. (e) Extrapolated excitation spectra by plotting the emission spectra intensities at selected wavelengths upon excitation at different wavelengths of compound **3a** in water and (f) dioxane.
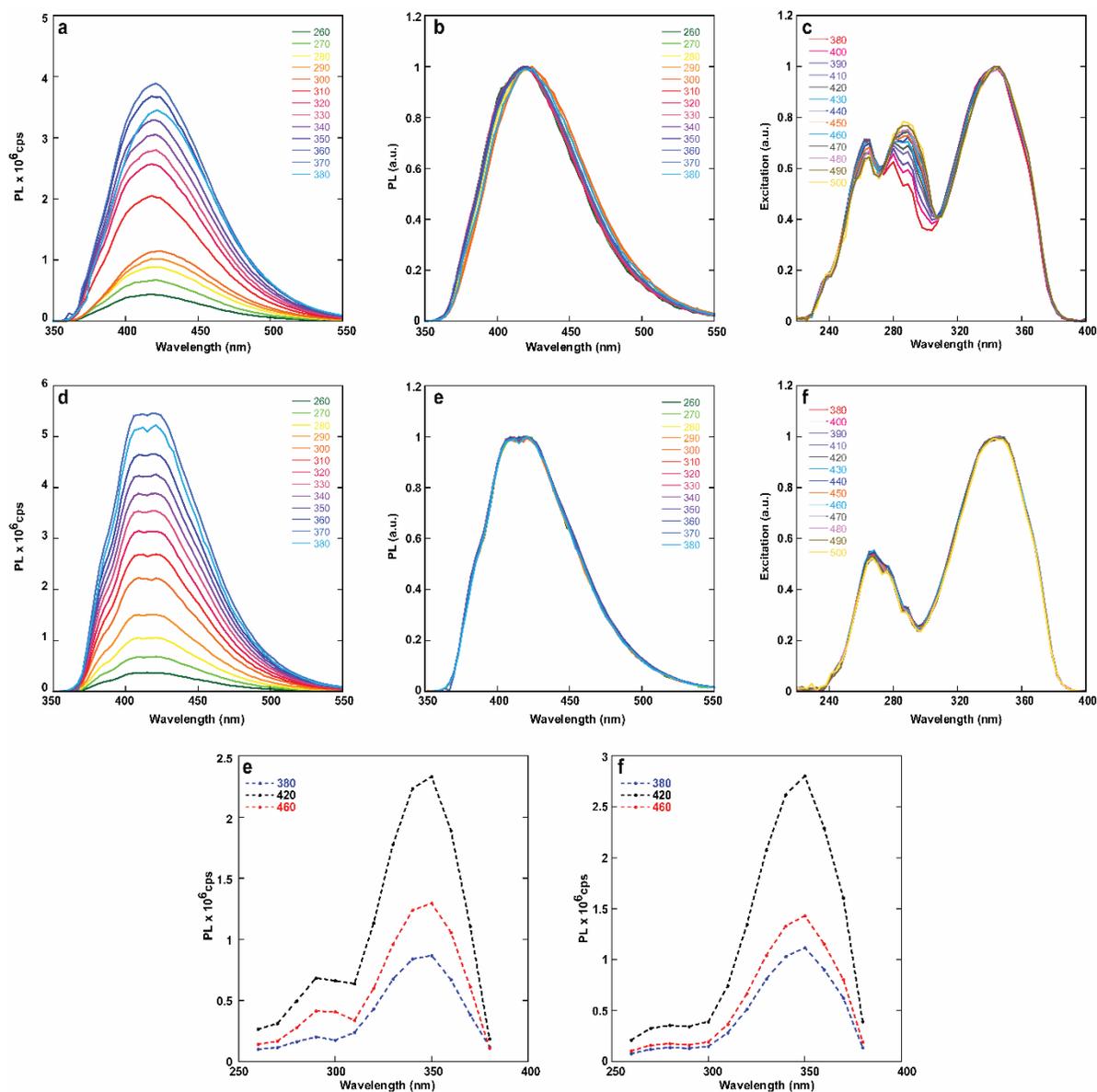
Figure 2.12. Spectra of compound **4a** recorded upon excitation every 10 nm between 260-380 nm in water (a-c) and dioxane (d-f). The spectra were recorded in and normalized for (a,d) the corresponding absorbance intensity or (b,e) to unit in intensity. (c) Excitation spectra of compound **4a** in water and (f) dioxane recorded at select emission wavelengths, covering the emission band. (e) Extrapolated excitation spectra by plotting the emission spectra intensities at selected wavelengths upon excitation at different wavelengths of compound **4a** in water and (f) dioxane.
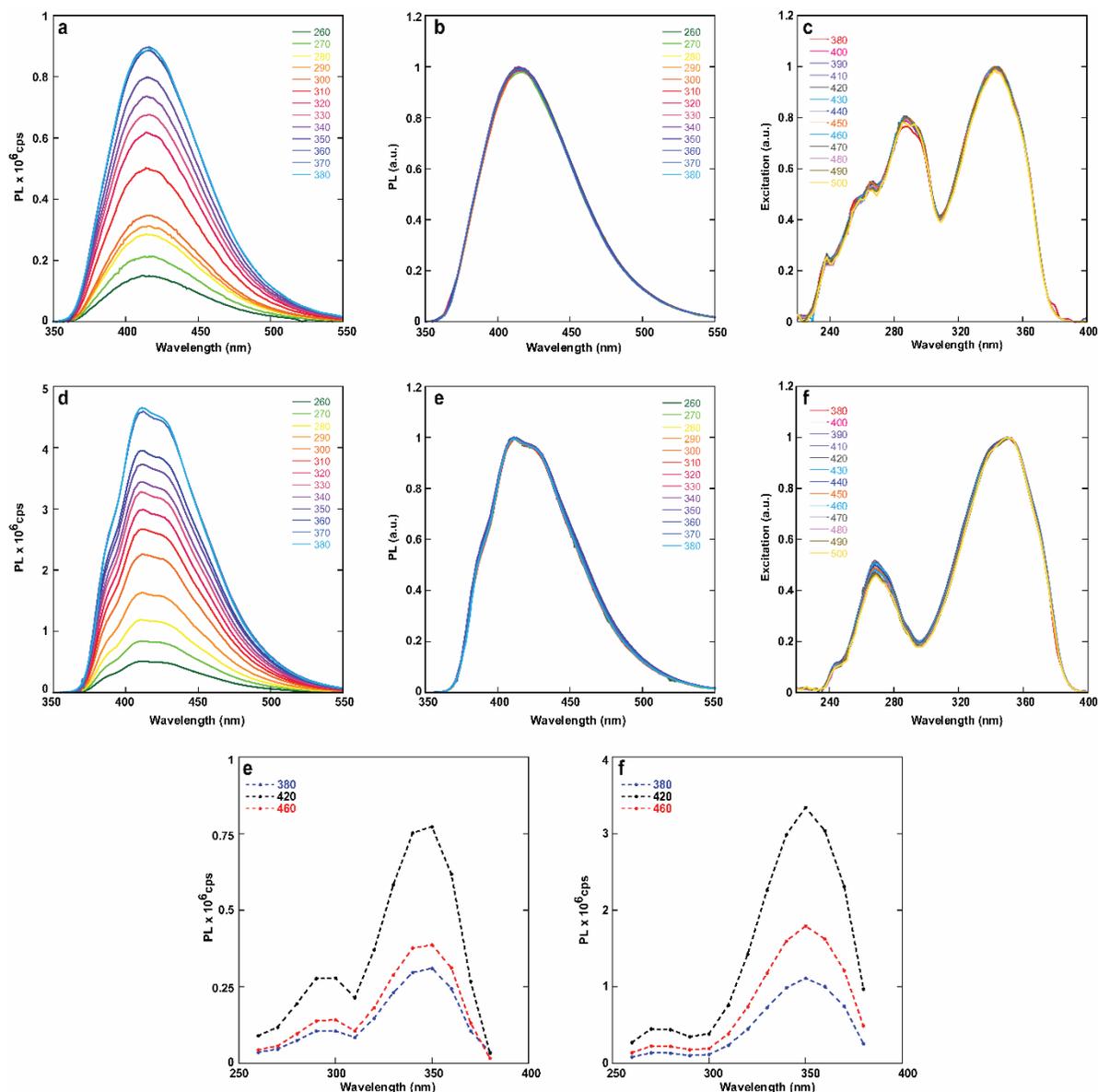
Figure 2.13. Spectra of compound **5a** recorded upon excitation every 10 nm between 260-380 nm in water (a-c) and dioxane (d-f). The spectra were recorded in and normalized for (a,d) the corresponding absorbance intensity or (b,e) to unit in intensity. (c) Excitation spectra of compound **5a** in water and (f) dioxane recorded at select emission wavelengths, covering the emission band. (e) Extrapolated excitation spectra by plotting the emission spectra intensities at selected wavelengths upon excitation at different wavelengths of compound **5a** in water and (f) dioxane.

74
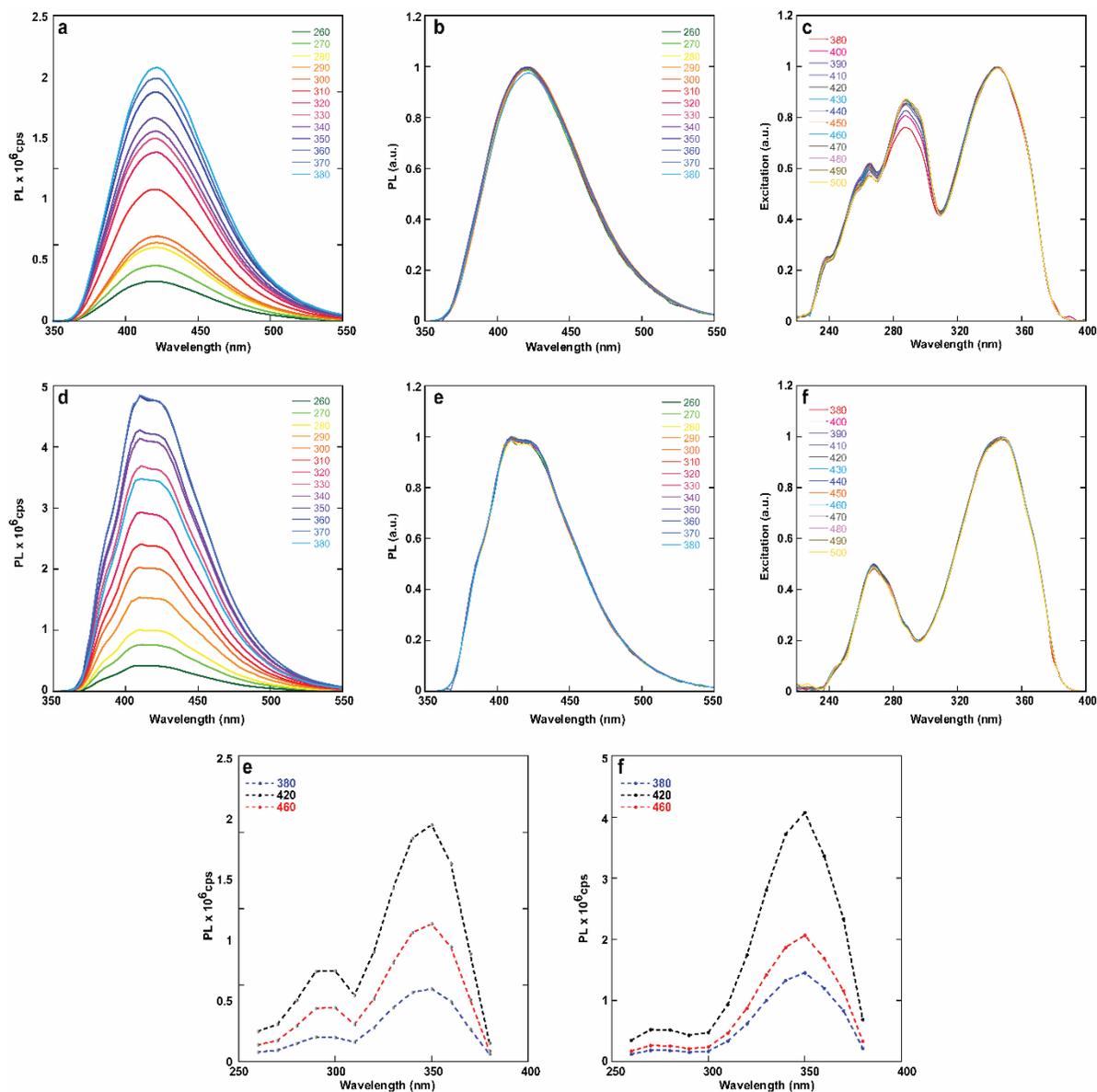
Figure 2.14. Spectra of compound **6a** recorded upon excitation every 10 nm between 260-380 nm in water (a-c) and dioxane (d-f). The spectra were recorded in and normalized for (a,d) the corresponding absorbance intensity or (b,e) to unit in intensity. (c) Excitation spectra of compound **6a** in water and (f) dioxane recorded at select emission wavelengths, covering the emission band. (e) Extrapolated excitation spectra by plotting the emission spectra intensities at selected wavelengths upon excitation at different wavelengths of compound **6a** in water and (f) dioxane.
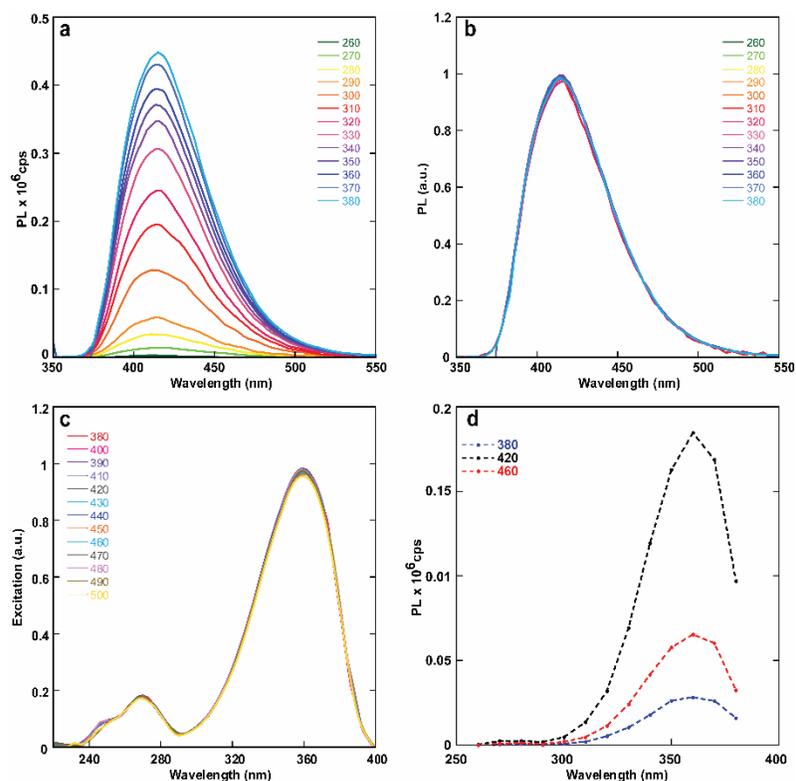
Figure 2.15. Emission spectra in dioxane of compound **3b** recorded upon excitation every 10 nm between 260-380 nm. The spectra were normalized for (a) the corresponding absorbance intensity or (b) to unit in intensity. (c) Excitation spectra of compound **3b** in dioxane recorded at select emission wavelengths, covering the emission band. (d) Extrapolated excitation spectra by plotting the emission spectra intensities at selected wavelengths upon excitation at different wavelengths of compound **3b** in dioxane.
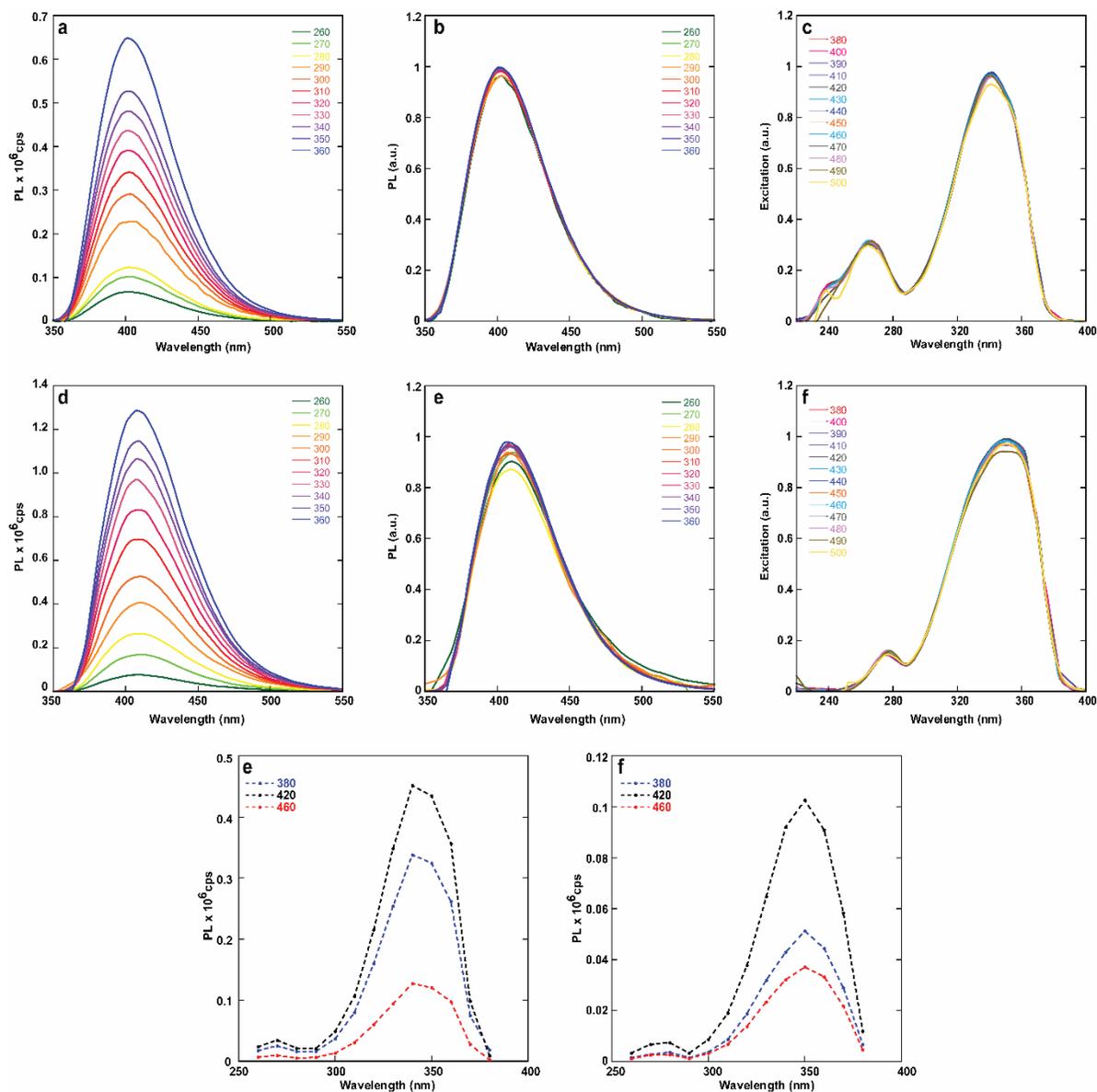
Figure 2.16. Emission spectra in dioxane of compound **4b** recorded upon excitation every 10 nm between 260-380 nm. The spectra were normalized for (a) the corresponding absorbance intensity or (b) to unit in intensity. (c) Excitation spectra of compound **4b** in dioxane recorded at select emission wavelengths, covering the emission band. (d) Extrapolated excitation spectra by plotting the emission spectra intensities at selected wavelengths upon excitation at different wavelengths of compound **4b** in dioxane.
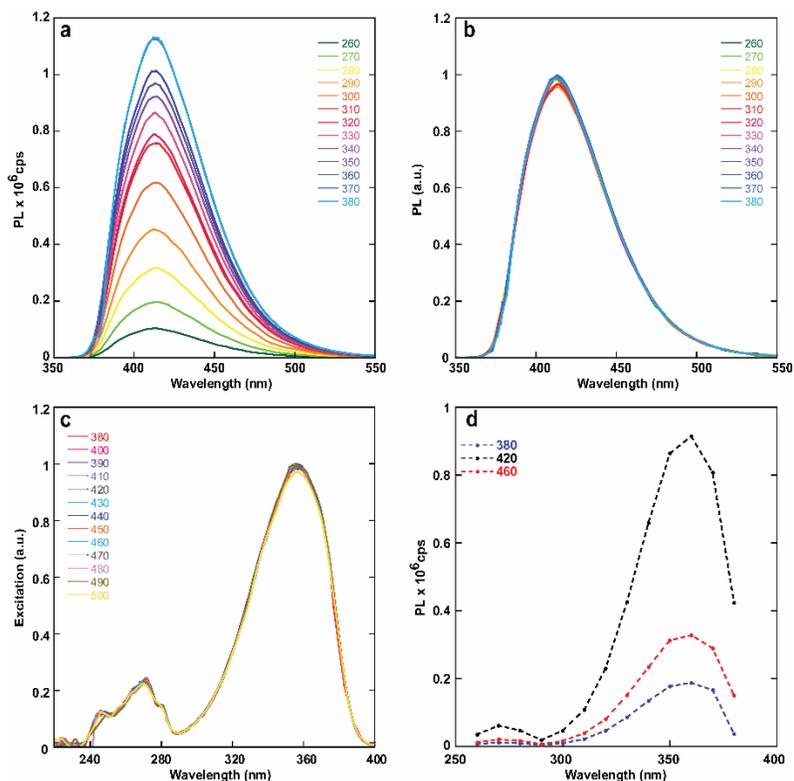
Figure 2.17. Emission spectra in dioxane of compound **5b** recorded upon excitation every 10 nm between 260-380 nm. The spectra were normalized for (a) the corresponding absorbance intensity or (b) to unit in intensity. (c) Excitation spectra of compound **5b** in dioxane recorded at select emission wavelengths, covering the emission band. (d) Extrapolated excitation spectra by plotting the emission spectra intensities at selected wavelengths upon excitation at different wavelengths of compound **5b** in dioxane.
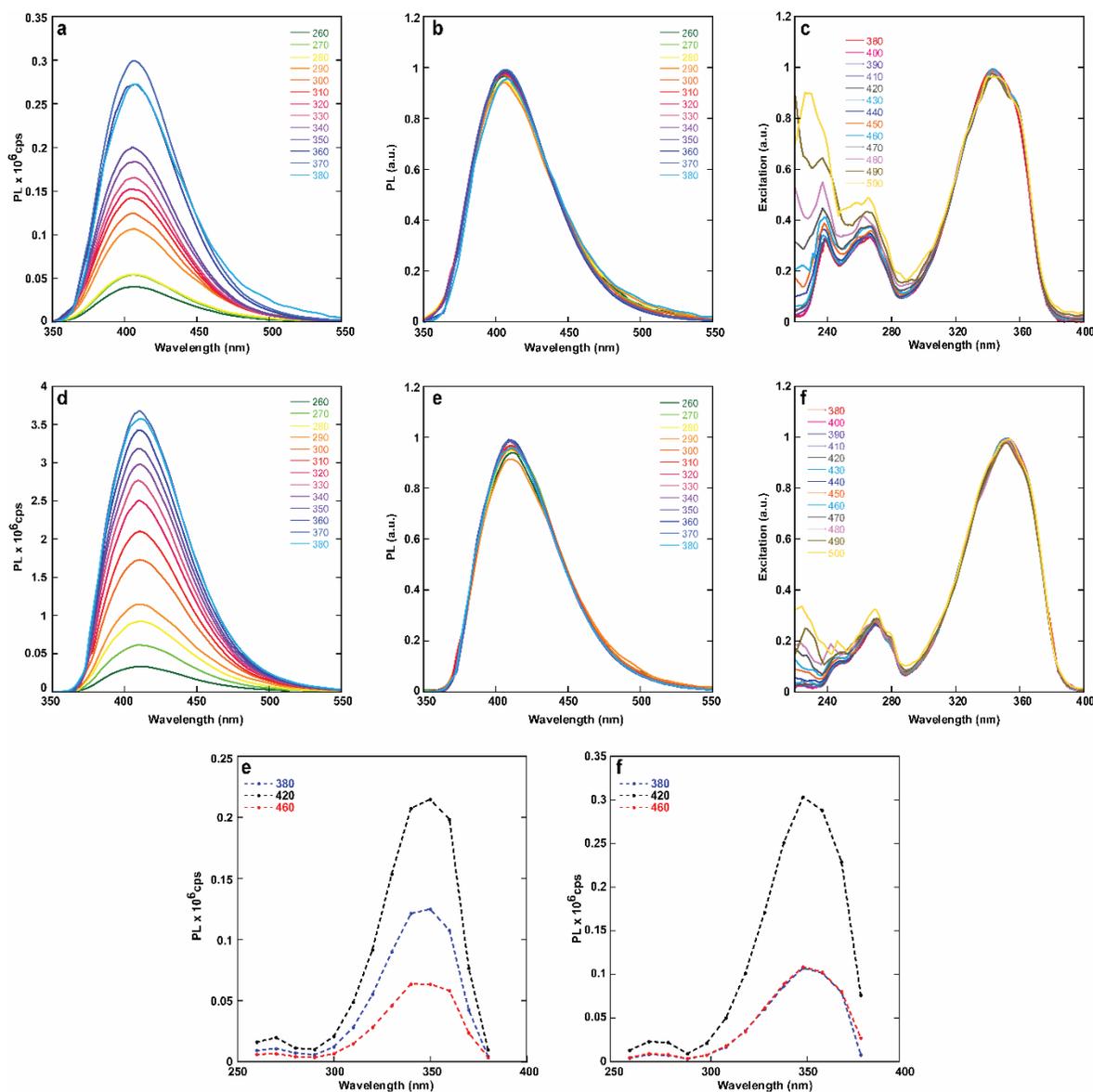
Figure 2.18. Emission spectra in dioxane of compound **6b** recorded upon excitation every 10 nm between 260-380 nm. The spectra were normalized for (a) the corresponding absorbance intensity or (b) to unit in intensity. (c) Excitation spectra of compound **6b** in dioxane recorded at select emission wavelengths, covering the emission band. (d) Extrapolated excitation spectra by plotting the emission spectra intensities at selected wavelengths upon excitation at different wavelengths of compound **6b** in dioxane.

2.9.4.4 Sensitivity to viscosity

Stock solutions (100 ml) of 20%, 40%, 60%, and 80% glycerol in methanol were prepared by mixing spectroscopic grade glycerol and methanol in a volumetric flask. The solutions were mixed for several days to allow for proper mixing. 1X stock solutions of the nucleobases in DMSO (10 µl) were then diluted in 2990 µl of MeOH, glycerol, or MeOH-glycerol mixtures and measured at 25 °C. For 80% glycerol in methanol and pure glycerol solutions, the samples were gently heated to induce mixing and allowed to cool for 30 minutes before taking spectra. The sample viscosity values were determined using the following equation:[3,4]

$$\ln \eta_{mix} = \sum_{i=1}^{2} w_i \cdot \ln \eta_i$$

Equation 2.1. Determination of mixed-solvent viscosities through individual viscosity values.

Where η_{mix} and η_i stand for the viscosity of the mixture and the viscosity of component i respectively. w_i stands for the weight fraction of component i. Values for pure MeOH and glycerol at 25 °C were taken from the literature.[5–7] Viscosity values for the methanol-glycerol mixtures are reported in table 2.10.

Table 2.10. Calculated viscosity values for MeOH-glycerol mixtures.

Glycerol % in MeOH	η [cP]	Log (η)
0	0.63	-0.20
20	4.74	0.67
40	24.8	1.4
60	98.9	2.0
80	319.7	2.5
100	875.00	2.9
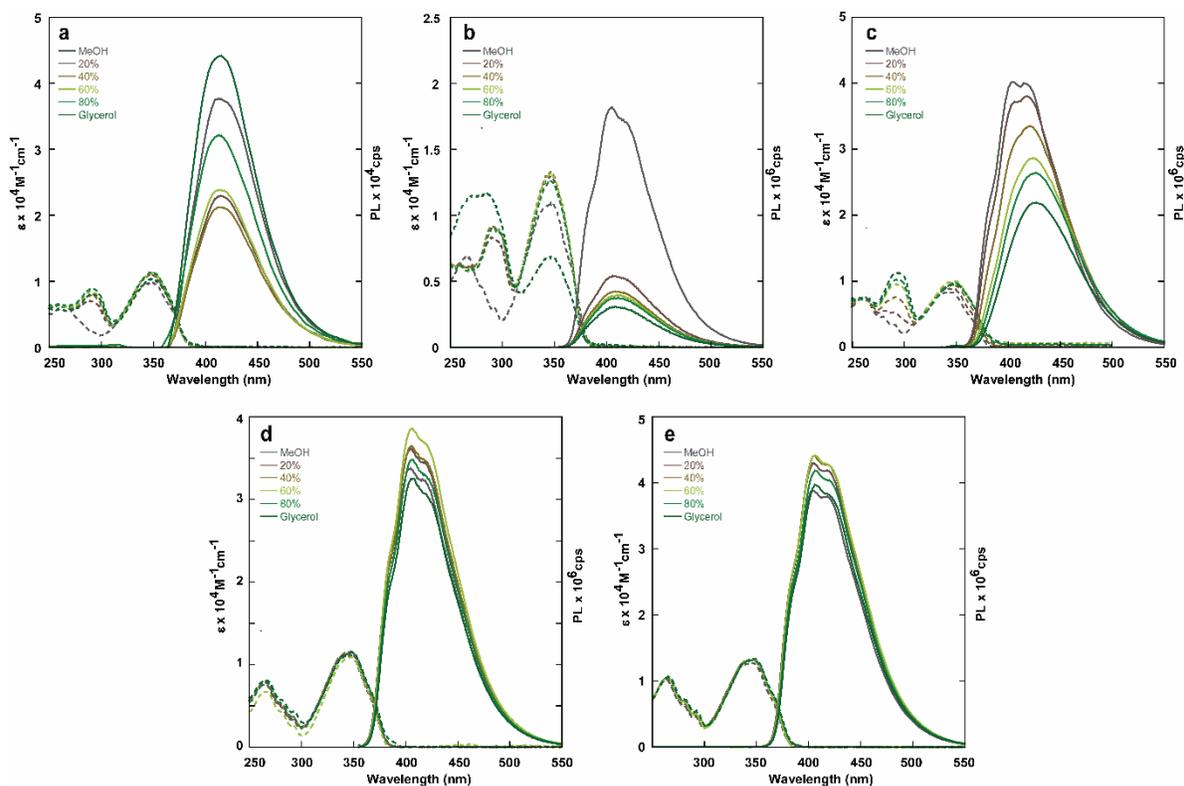
Figure 2.19. Absorption (dashed lines) and emission (solid lines) traces in mixtures of methanol and glycerol at 25 °C for **2a**, (b) **3a**, (c) **4a**, (d) **5a**, and (e) **6a**. The emission spectra were normalized to 0.1 intensity at the excitation wavelength.
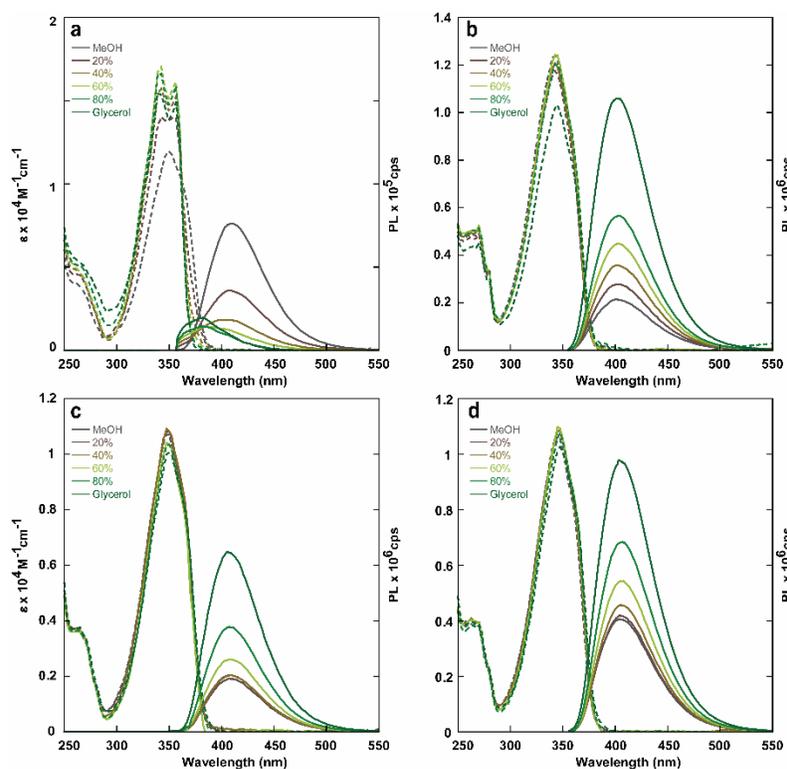
Figure 2.20. Absorption (dashed lines) and emission (solid lines) traces in mixtures of methanol and glycerol at 25 °C for **3b**, (b) **4b**, (c) **5b**, and (d) **6b**. The emission spectra were normalized to 0.1 intensity at the excitation wavelength.
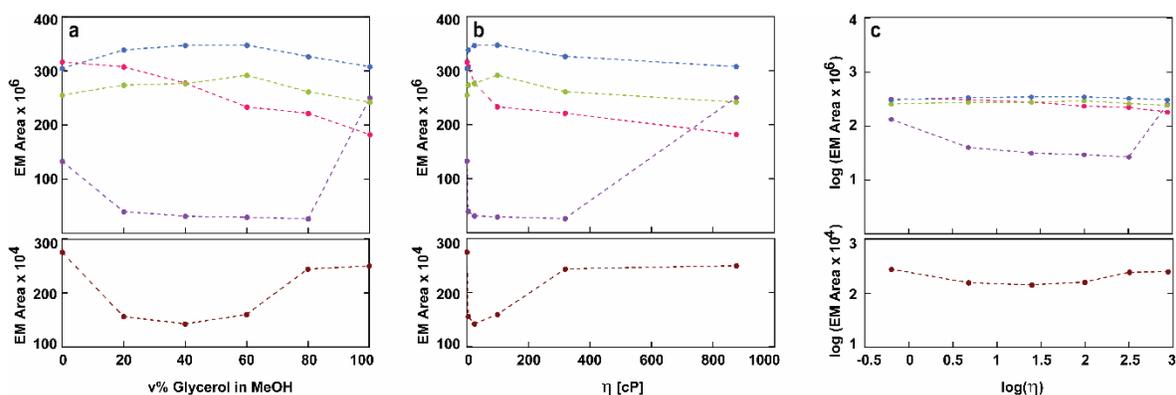
Figure 2.21. Area under the emission band of **2a** (red), **3a** (purple), **4a** (pink), **5a** (light green) and **6a** (blue) plotted against (a) v% glycerol in MeOH and (b) viscosity coefficients for MeOH-glycerol mixtures. (c) logarithm of the area under the emission bands of specified compounds plotted against the logarithm of the viscosity coefficients.

82
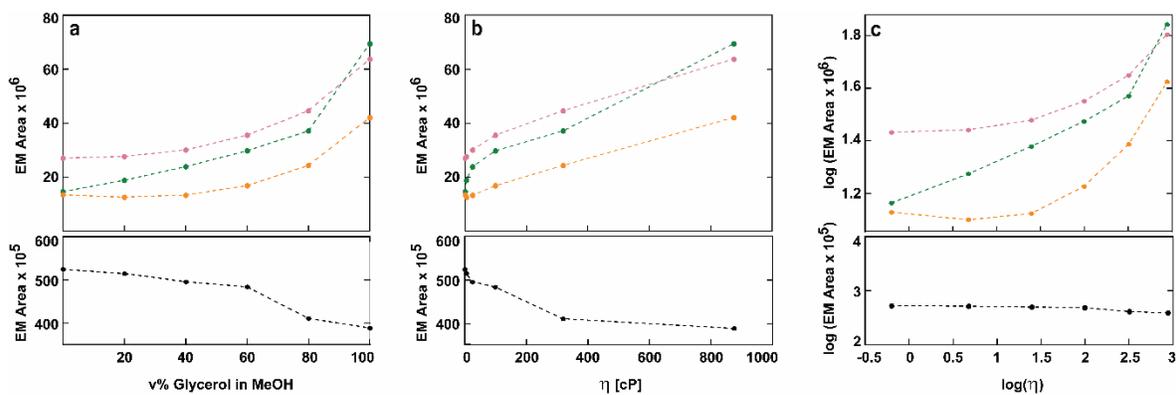
Figure 2.22. Area under the emission band of **3b** (black), **4b** (dark green), **5b** (orange) and **6b** (magenta) plotted against (a) v% glycerol in MeOH and (b) viscosity coefficients for MeOH-glycerol mixtures. (c) logarithm of the area under the emission bands of specified compounds plotted against the logarithm of the viscosity coefficients.

2.9.4.5 Sensitivity to temperature-induced viscosity

For further analysis of viscosity-dependent fluorescence of compounds **2-6a** and **3-6b**, 10 µl of 1X nucleobase stock solutions in DMSO were diluted in 2990 µl of spectroscopic grade ethylene glycol. The solutions were allowed to equilibrate to the appropriate temperature for 30 minutes before measuring the absorption, emission, and excitation spectra of the samples.
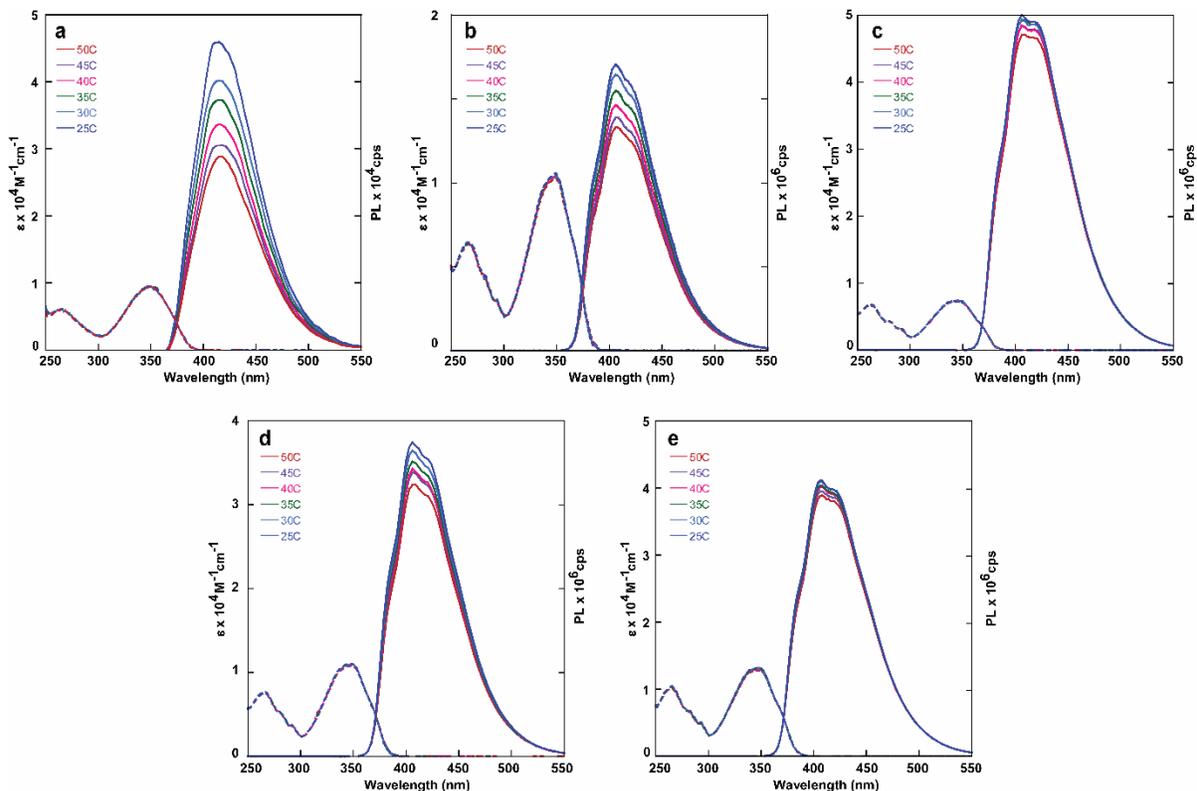
Figure 2.23. Absorption (dashed lines) and emission (solid lines) traces in ethylene glycol for (a) **2a**, (b) **3a**, (c) **4a**, (d) **5a**, and (e) **6a** between 25 and 50 °C. The emission spectra were normalized to 0.1 intensity at the excitation wavelength.
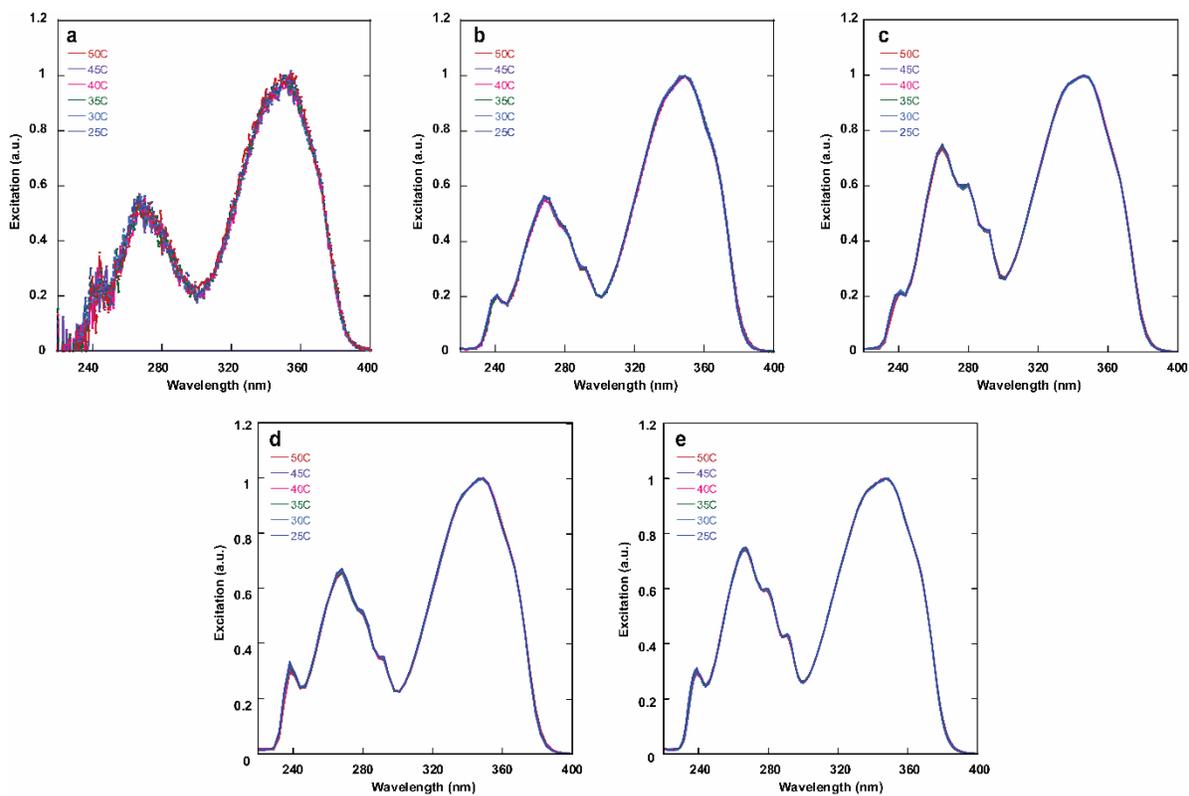
Figure 2.24. Excitation spectra in ethylene glycol for (a) **2a**, (b) **3a**, (c) **4a**, (d) **5a**, and (e) **6a** between 25 and 50 °C. The emission spectra were normalized to the maximum value.
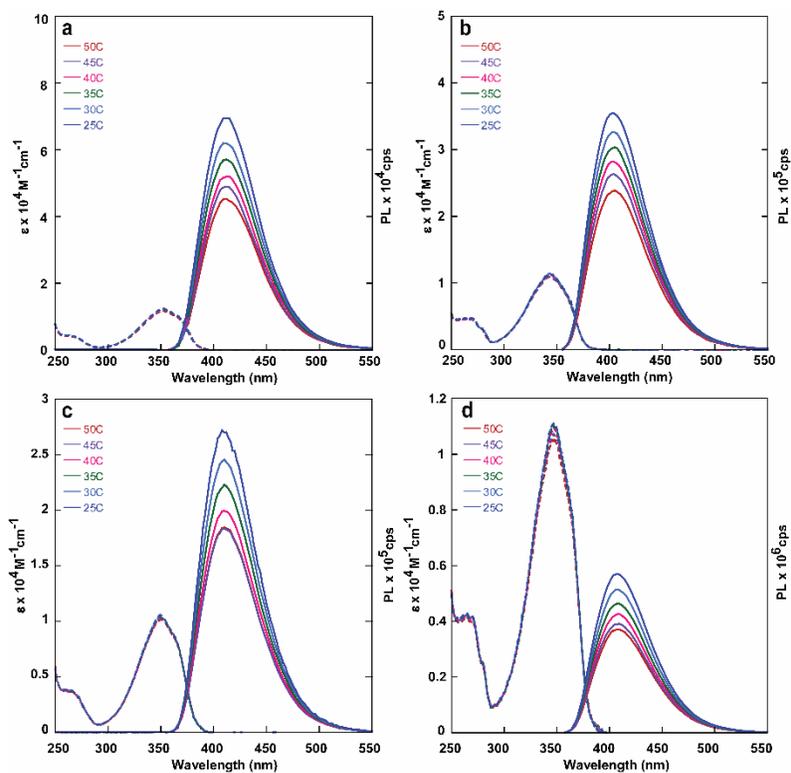
Figure 2.25. Absorption (dashed lines) and emission (solid lines) traces in ethylene glycol for (a) **3b**, (b) **4b**, (c) **5b**, and (d) **6b** between 25 and 50 °C. The emission spectra were normalized to 0.1 intensity at the excitation wavelength.
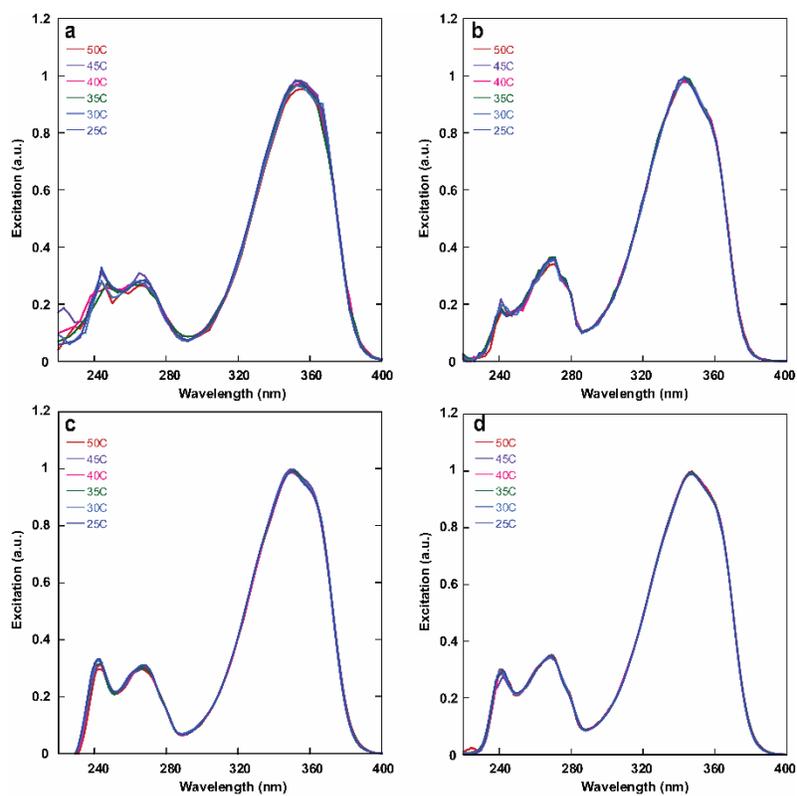
Figure 2.26. Excitation spectra in ethylene glycol for (a) **3b**, (b) **4b**, (c) **5b**, and (d) **6b** between 25 and 50 °C. The excitation spectra were normalized to the maximum value.

2.9.4.6 Sensitivity to Polarity

Experiments evaluating the effect of polarity were performed in water, dioxane and their mixtures (20, 40, 60 and 80 v/v % water in dioxane). Solutions were prepared in advance and allowed to mix overnight. The sample $E_T(30)$ values were determined by dissolving a small amount of Reichardt's dye in the mixture of the same solvent used to dilute the nucleoside's DMSO sample. The observed long wavelength absorption maximum (λ_{abs}^{max}) was converted to the $E_T(30)$ values with the following equation.

$$E_T(30) = \frac{28591}{\lambda_{abs}^{max}}$$

Table 2.11. E$_T$(30) experimental values for water/dioxane mixtures

Water % in dioxane	Reported E$_T$(30)[a] (Kcal mol^{-1})	Experimental E$_T$(30) (Kcal mol^{-1})
0	36.4	36.5
20	48.3	48.3
40	51.6	52.4
60	55.0	55.5
80	57.5	58.4
100	63.1	-

[a] Literature values.[8] Due to solubility limitation the water E$_T$(30) value was not calculated.
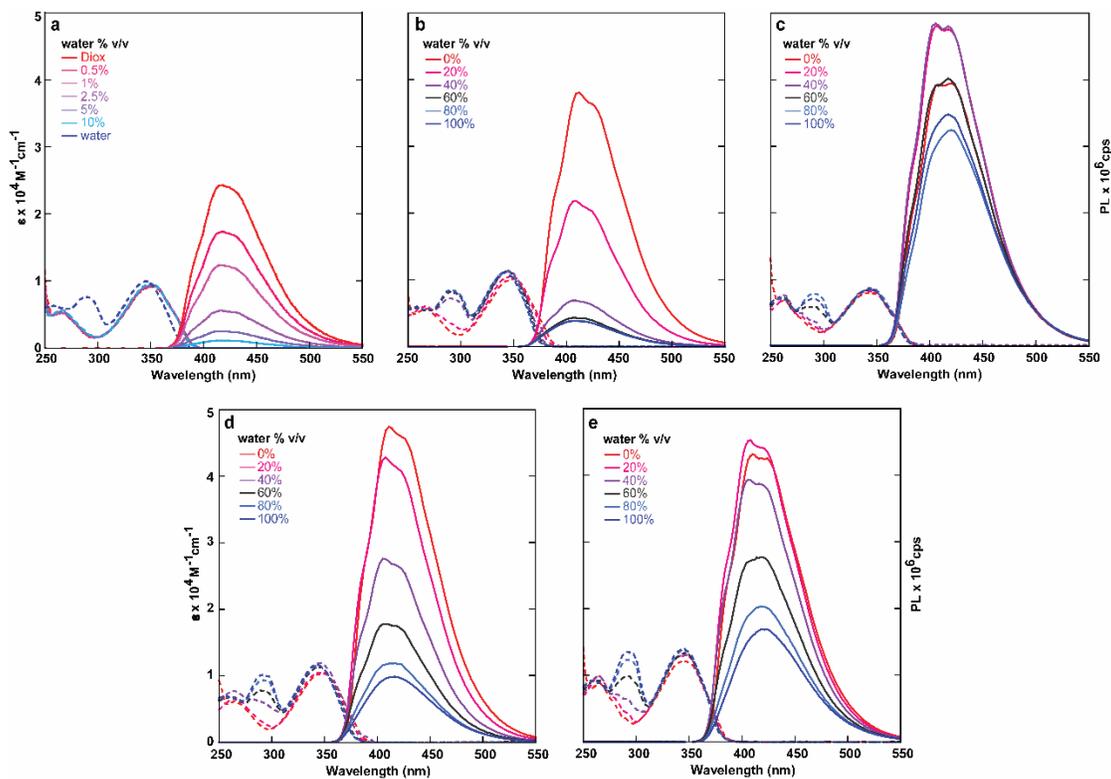
Figure 2.27. Absorption (dashed lines) and emission (solid lines) traces in water, dioxane and mixture thereof for (a) **2a**, (b) **3a**, (c) **4a**, (d) **5a** and (e) **6a**. The emission spectra were normalized to 0.1 intensity at the excitation wavelength.
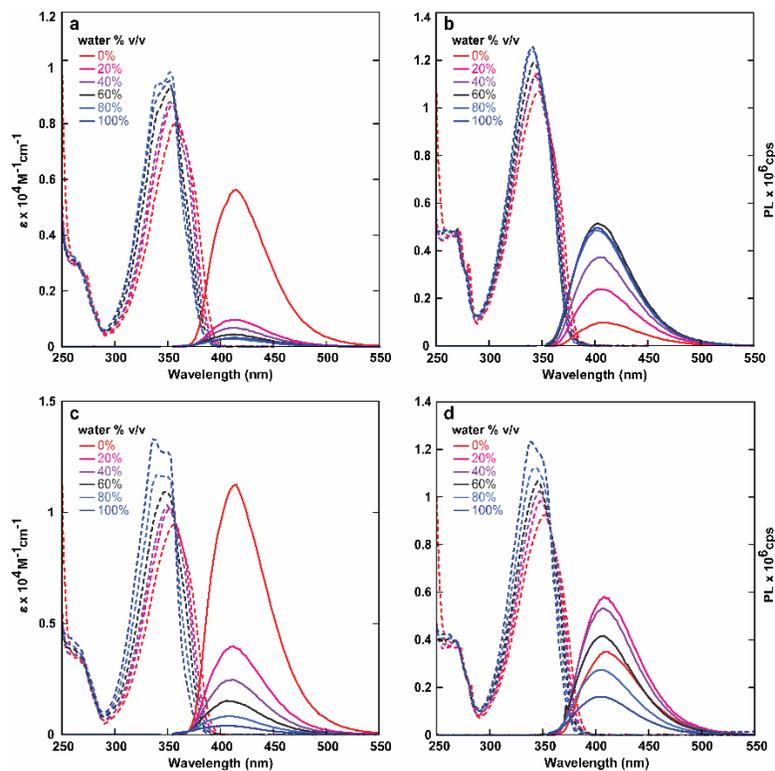
Figure 2.28. Absorption (dashed lines) and emission (solid lines) traces in water, dioxane and mixture thereof for (a) **3b**, (b) **4b**, (c) **5b**, and (d) **6b**. The emission spectra were normalized to 0.1 intensity at the excitation wavelength.
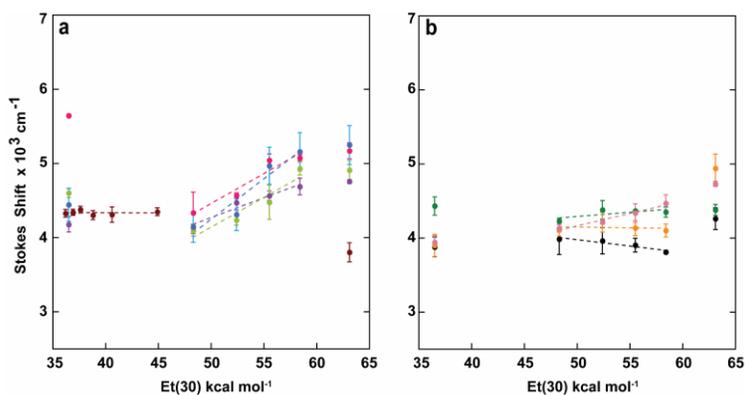
Figure 2.29. Stokes shift vs. solvent polarity correlation ($E_T(30)$) of water/dioxane mixtures for (a) **2a** (red), **3a** (purple), **4a** (pink), **5a** (light green), and **6a** (blue) as well as (b) **3b** (black), **4b** (dark green), **5b** (orange), and **6b** (magenta).

2.9.5 HPLC analysis of compounds 2–6a/b

To validate the photophysical experiments described above, the stock solutions were checked for purity using HPLC before running photophysical experiments. 20 μl of 1X stock solutions (compounds **2a**, **3**–**6a/b**) or 10X stock solutions (compounds **2b** and **11**) were diluted with 0.1% TFA in water (200 μl). HPLC analysis was carried out with an Agilent 1260 series system with a Sepax Bio-C18, 5μm, 300 Å, 10 × 250 mm column. 0.1% TFA stock solutions were prepared by dissolving 1 ml of TFA (Sigma-Aldrich, 99%) in 999 ml MilliQ water and filtered using Millipore type GNWP 0.2 μm filters before use. Each injection (100 μl) was subjected to a gradient (20 minutes, from 0.5 to 40% acetonitrile 0.1% TFA in water 0.1% TFA) followed by a flush (10 minutes). A flow rate of 3 mL / minute was used and the run was carried out at 25.00 ± 0.10 °C. Each run was monitored at 260, 280 and 380 nm with calibrated references at 600 nm and slit set at 1 nm.

2.9.6 Supplementary figures

2.9.6.1 HPLC Traces of stock solutions for compounds 2–6a/b and 11
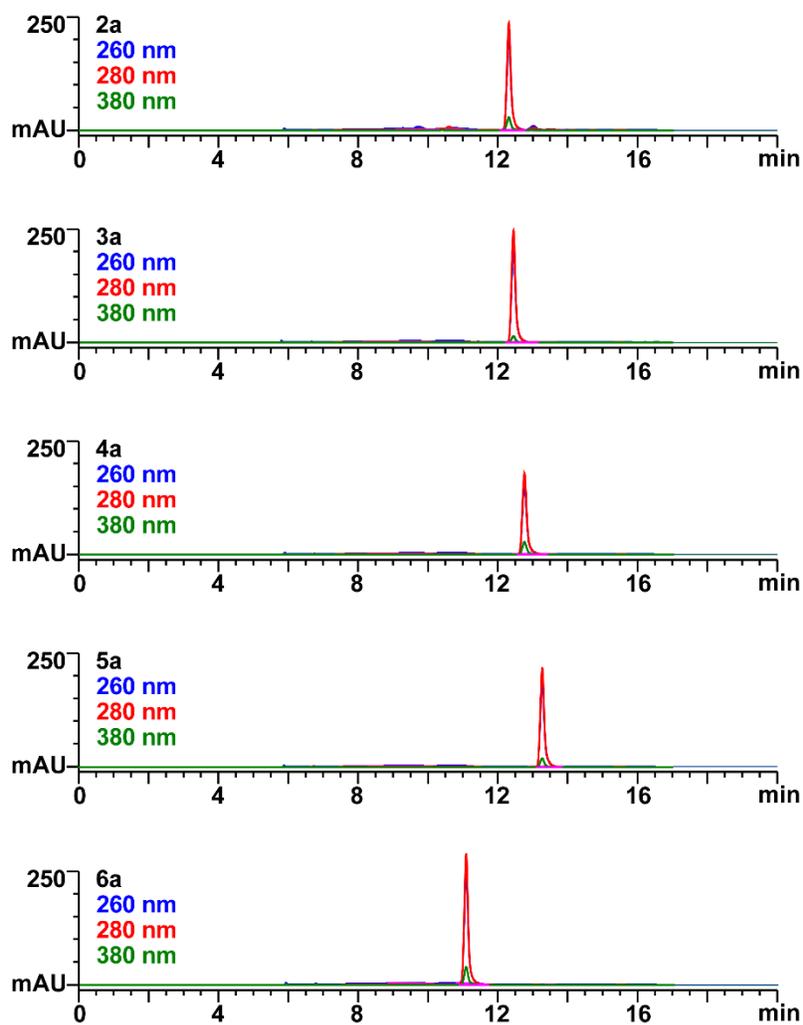
Figure 2.30. Raw HPLC traces of stock solutions for compounds **2–6a**.
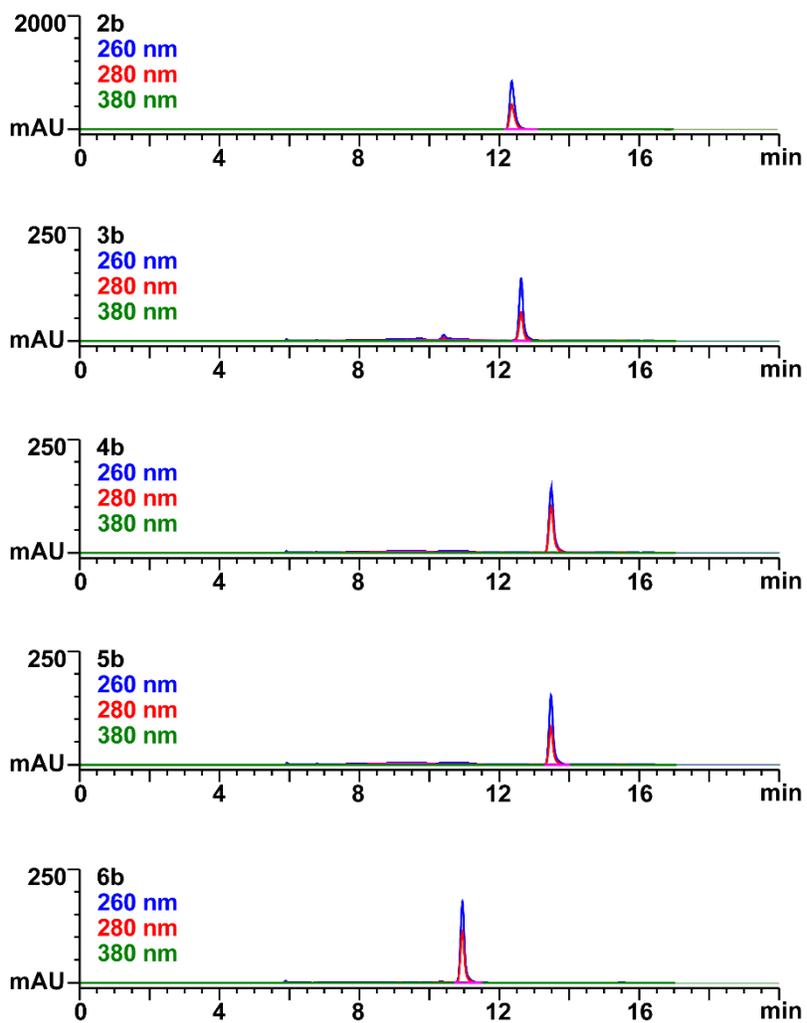
Figure 2.31. Raw HPLC traces for compounds **2–6b**.
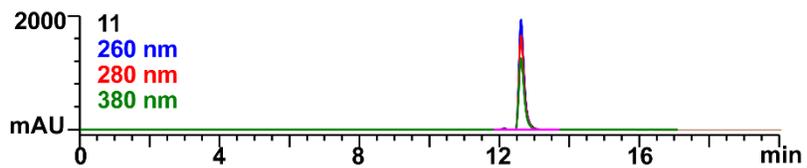
Figure 2.32. Raw HPLC traces for compound **11**.

2.9.6.2 NMR spectroscopy
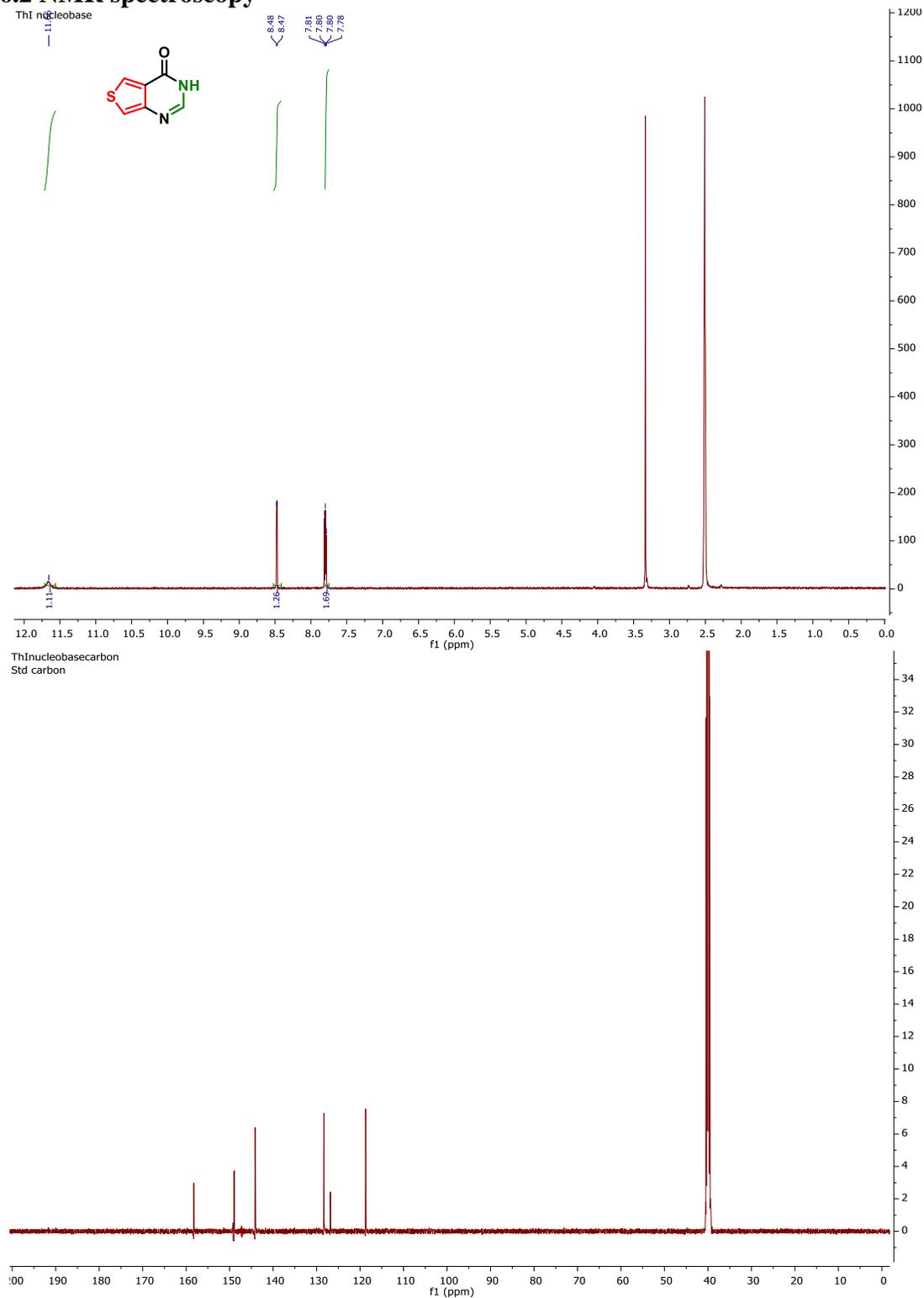
Figure 2.33. ^{1}H and ^{13}C NMR spectra of **6a**.
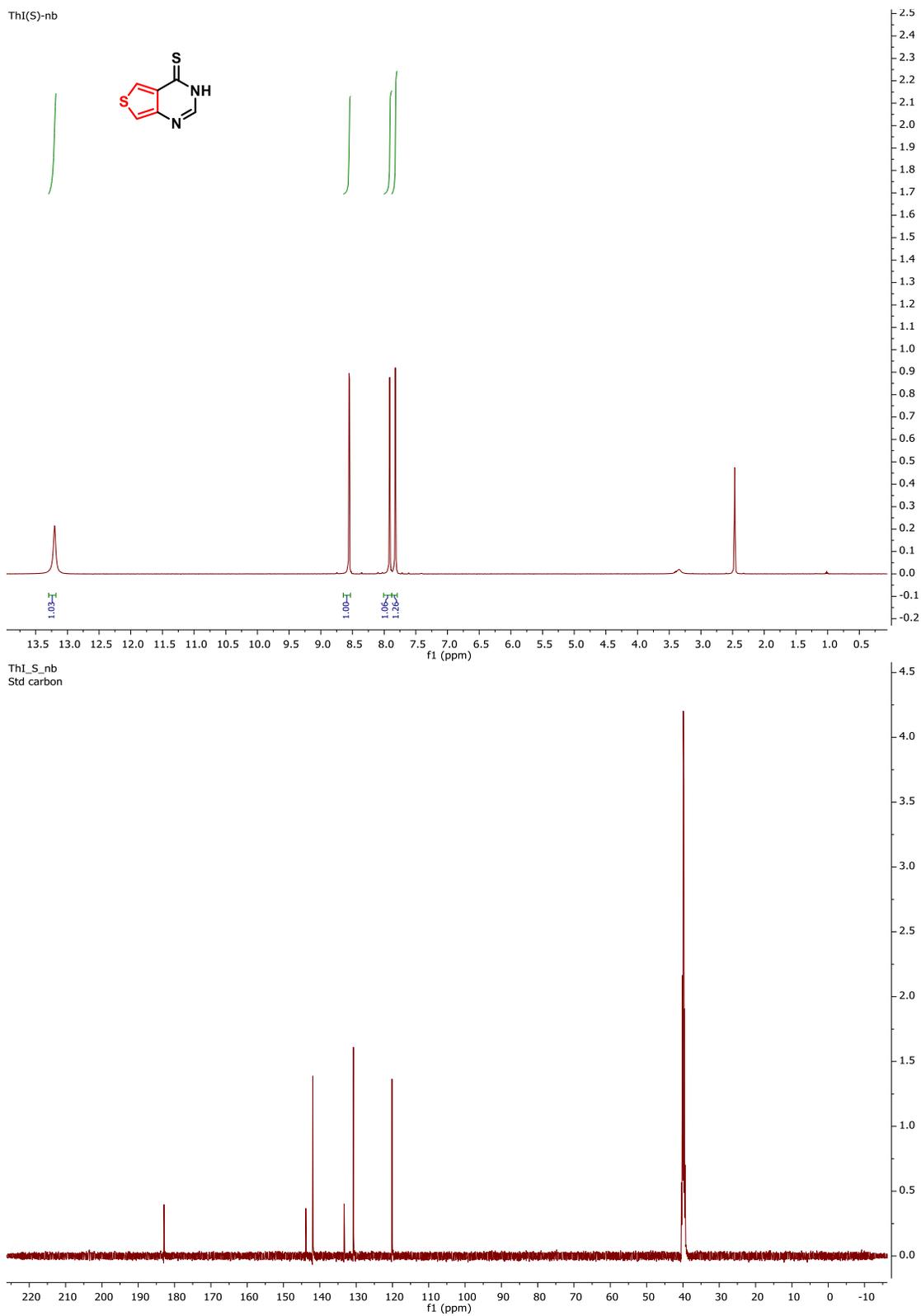
Figure 2.34. ^{1}H and ^{13}C NMR spectra of **7a**.
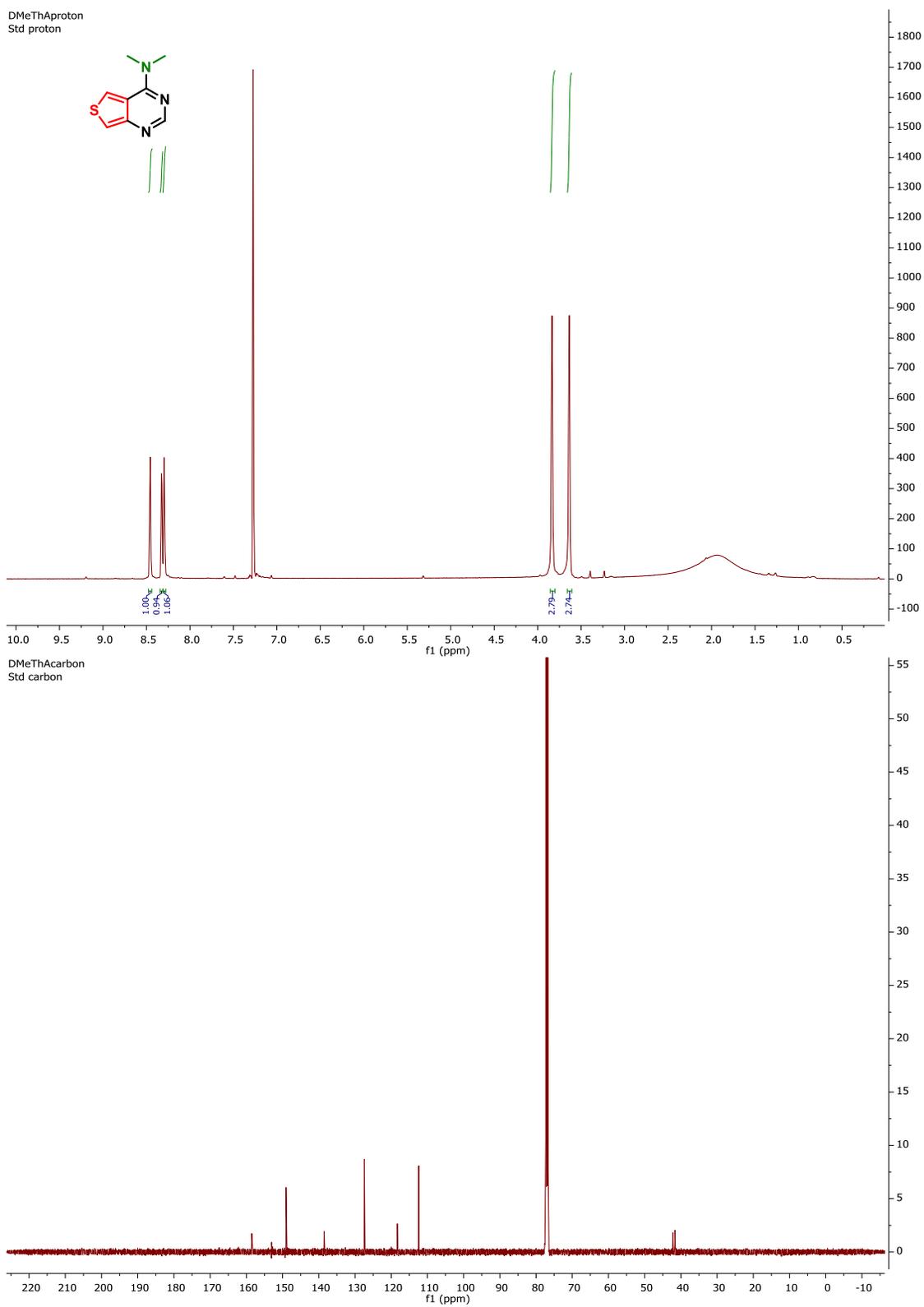
Figure 2.35. ^{1}H and ^{13}C NMR spectra of **2a**.
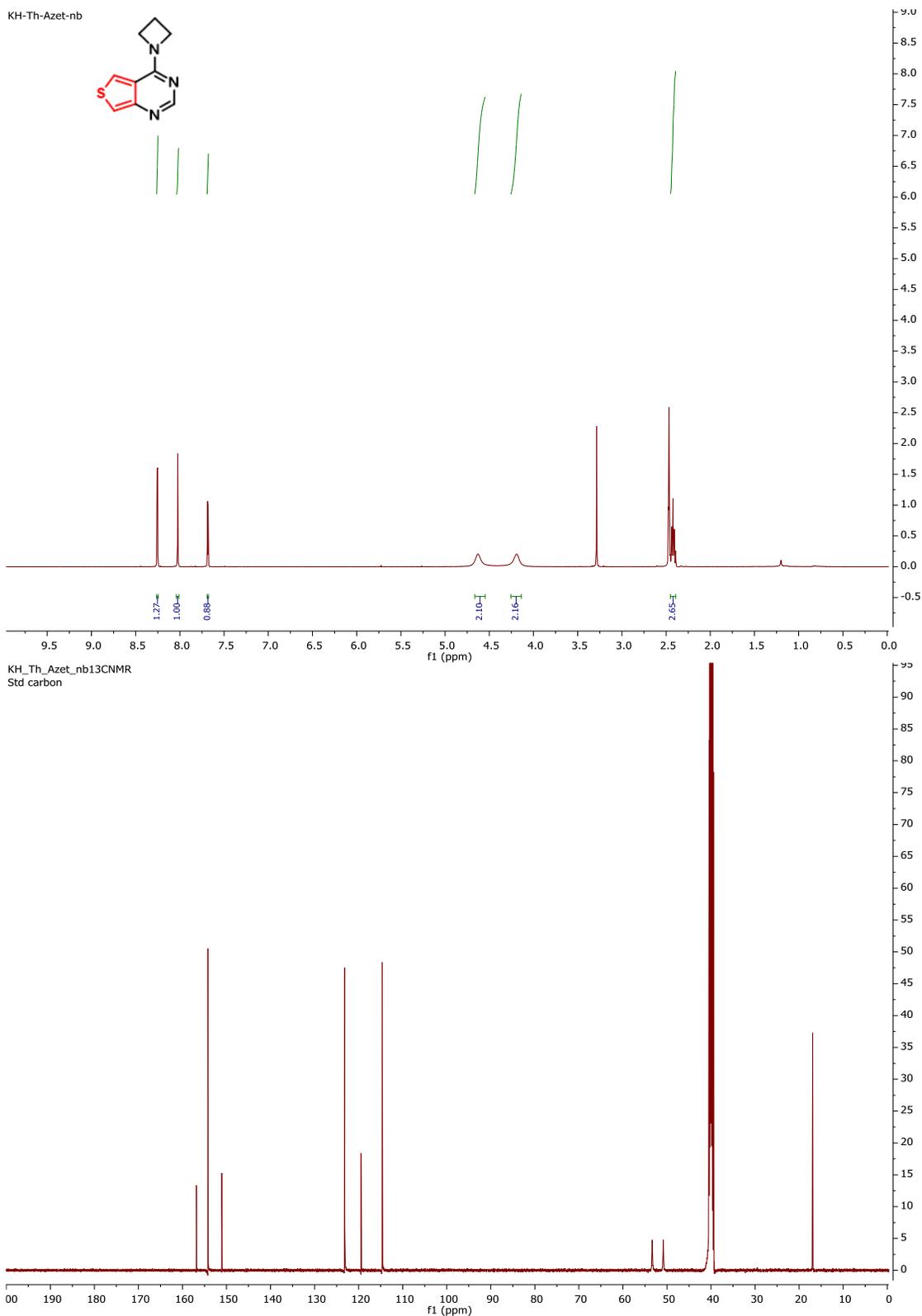
Figure 2.36. ^{1}H and ^{13}C NMR spectra of **3a**.
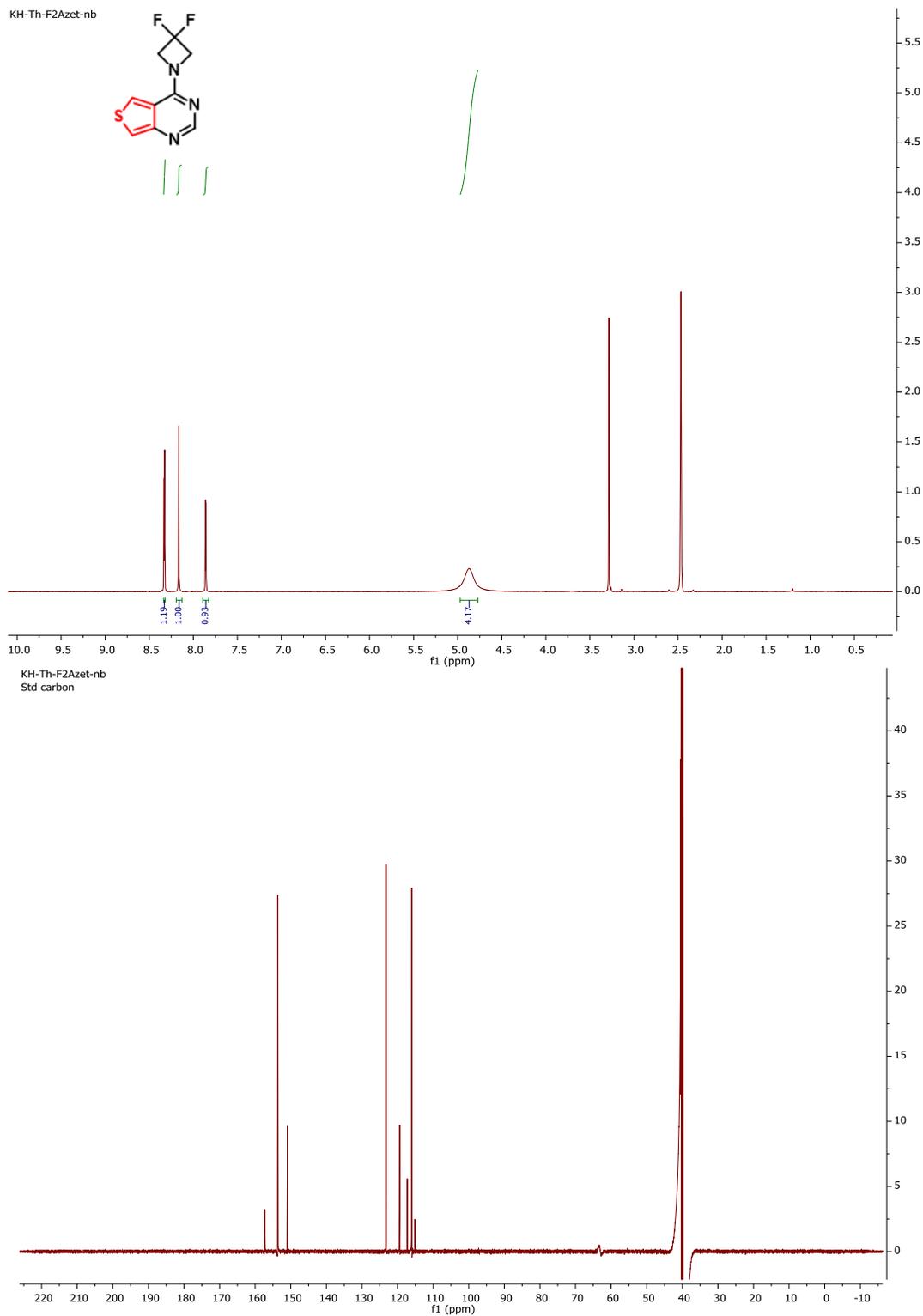
Figure 2.37. ^{1}H and ^{13}C NMR spectra of **4a**.

Figure 2.38. ^{1}H and ^{13}C NMR spectra of **5a**.
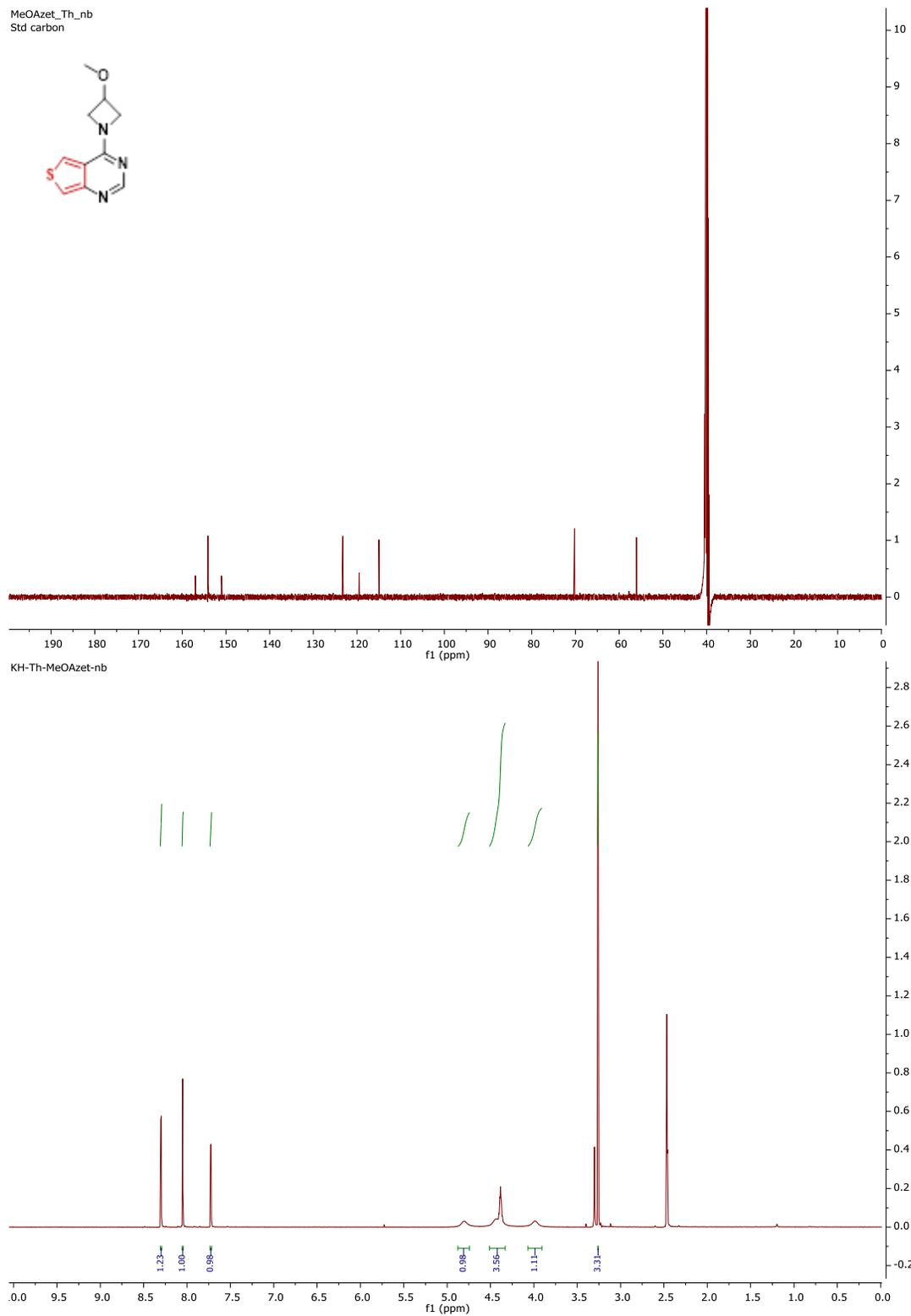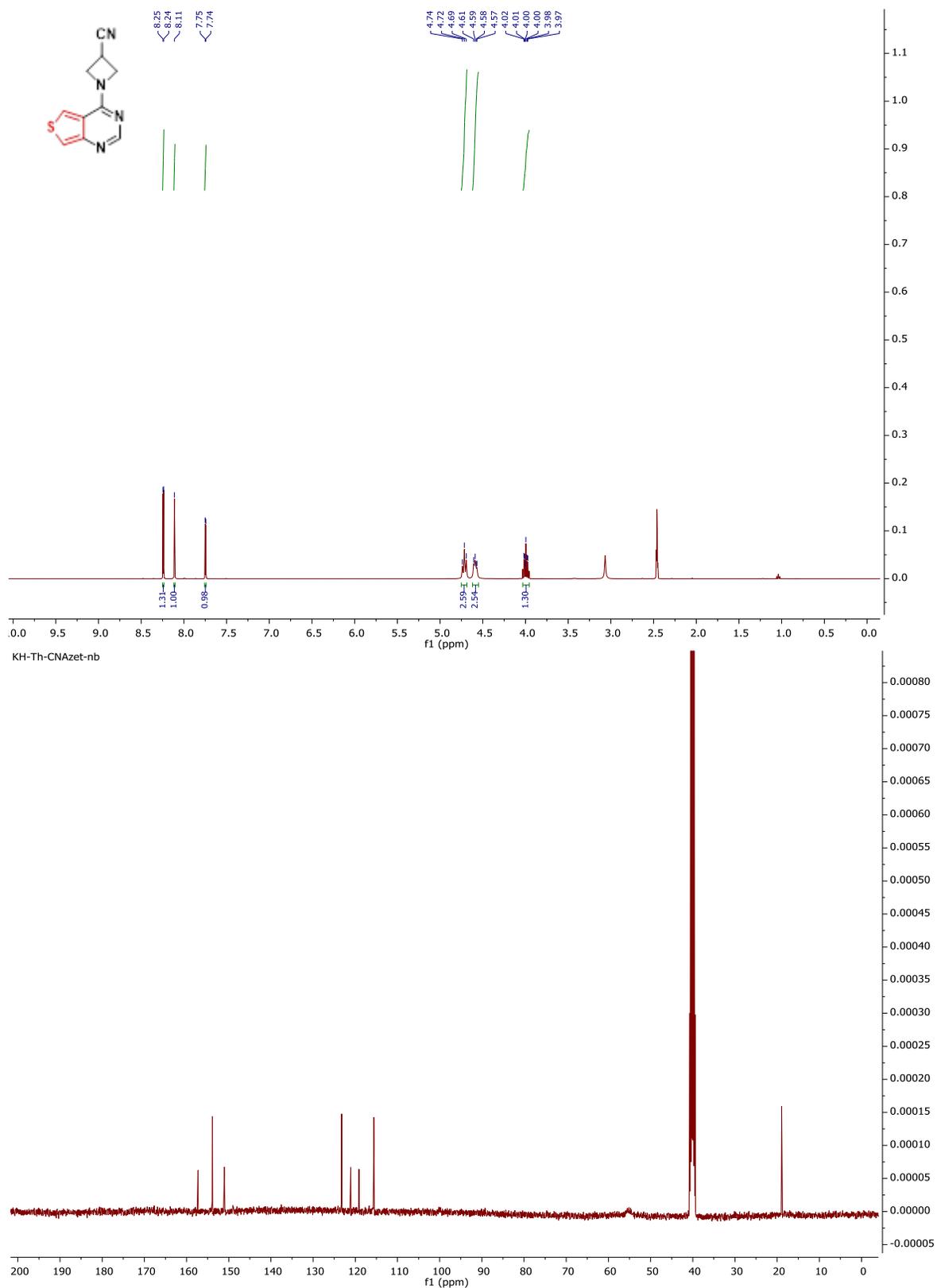
Figure 2.39. ^{1}H and ^{13}C NMR spectra of **6a**.
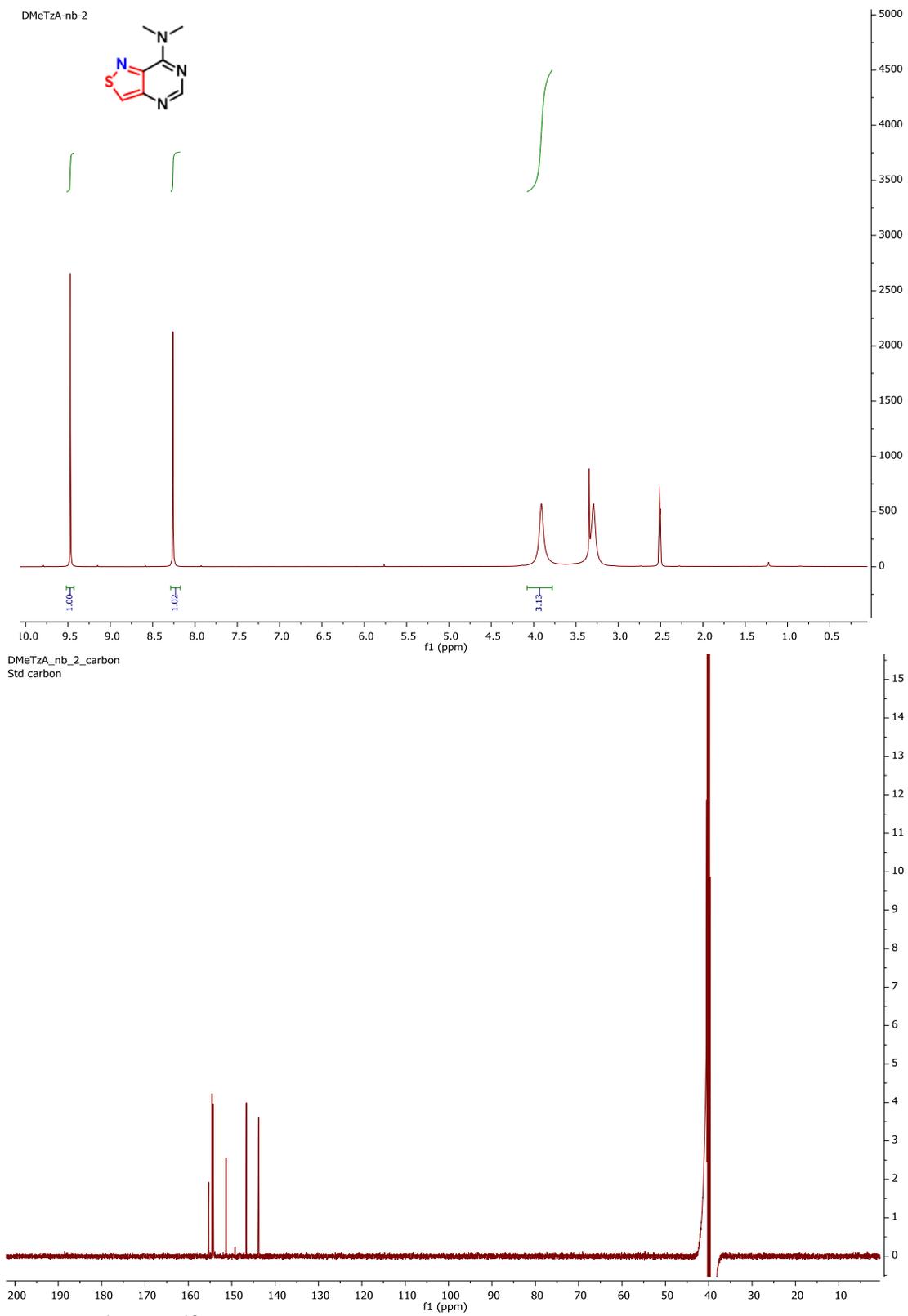
Figure 2.40. ^{1}H and ^{13}C NMR spectra of **2b**.
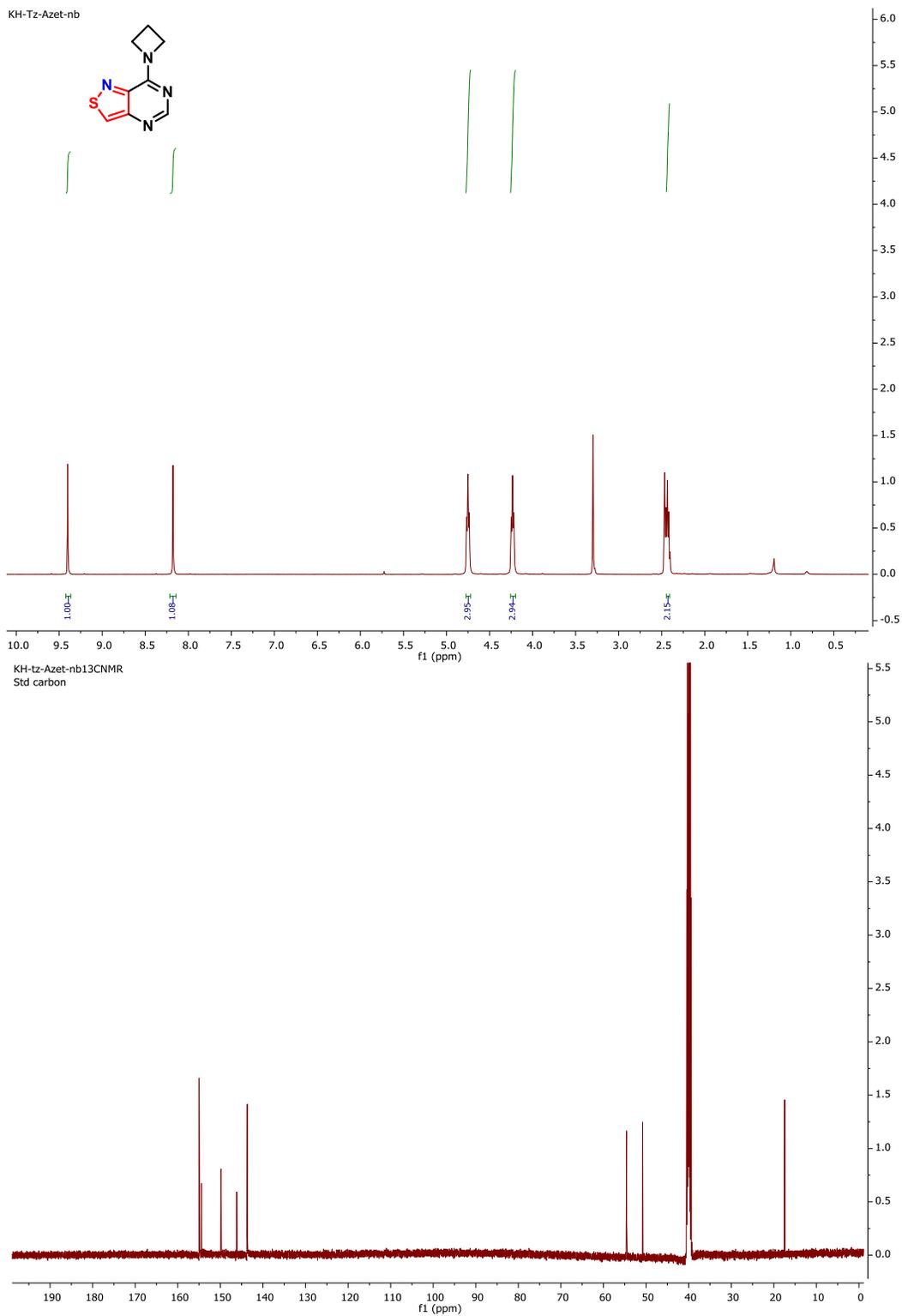
Figure 2.41. ^{1}H and ^{13}C NMR spectra of **3b**.
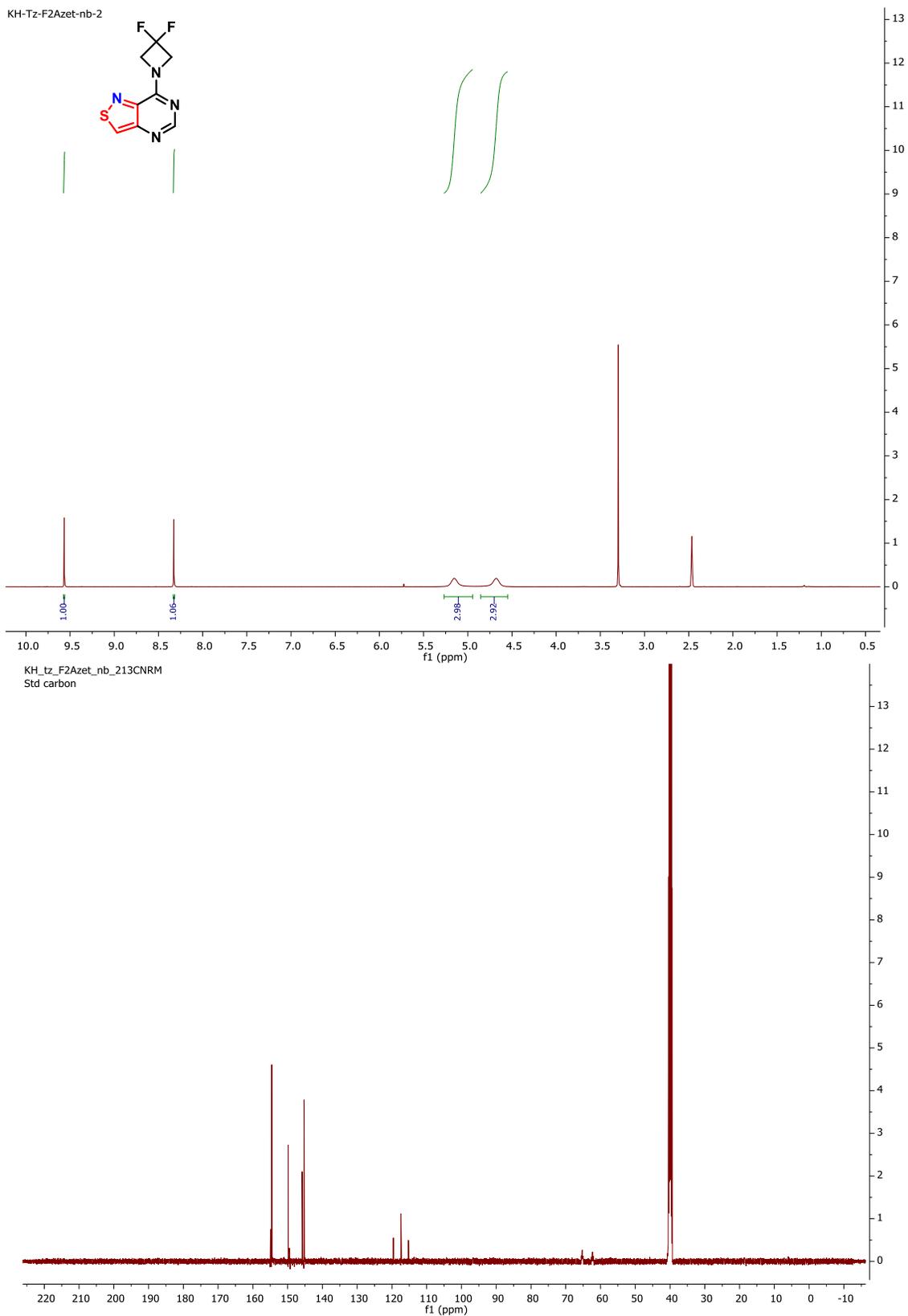
Figure 2.42. ^{1}H and ^{13}C NMR spectra of **4b**.
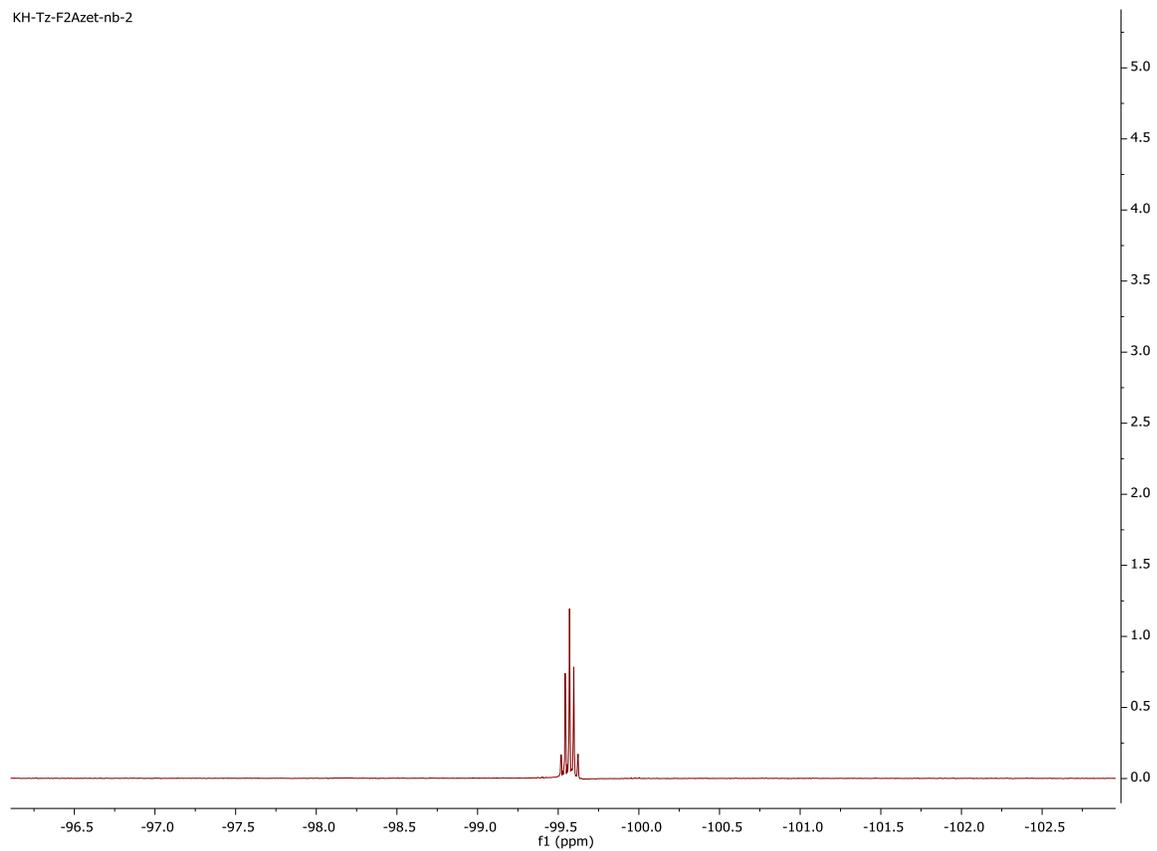
Figure 2.43. ^{19}F NMR spectra of compound **4b**.
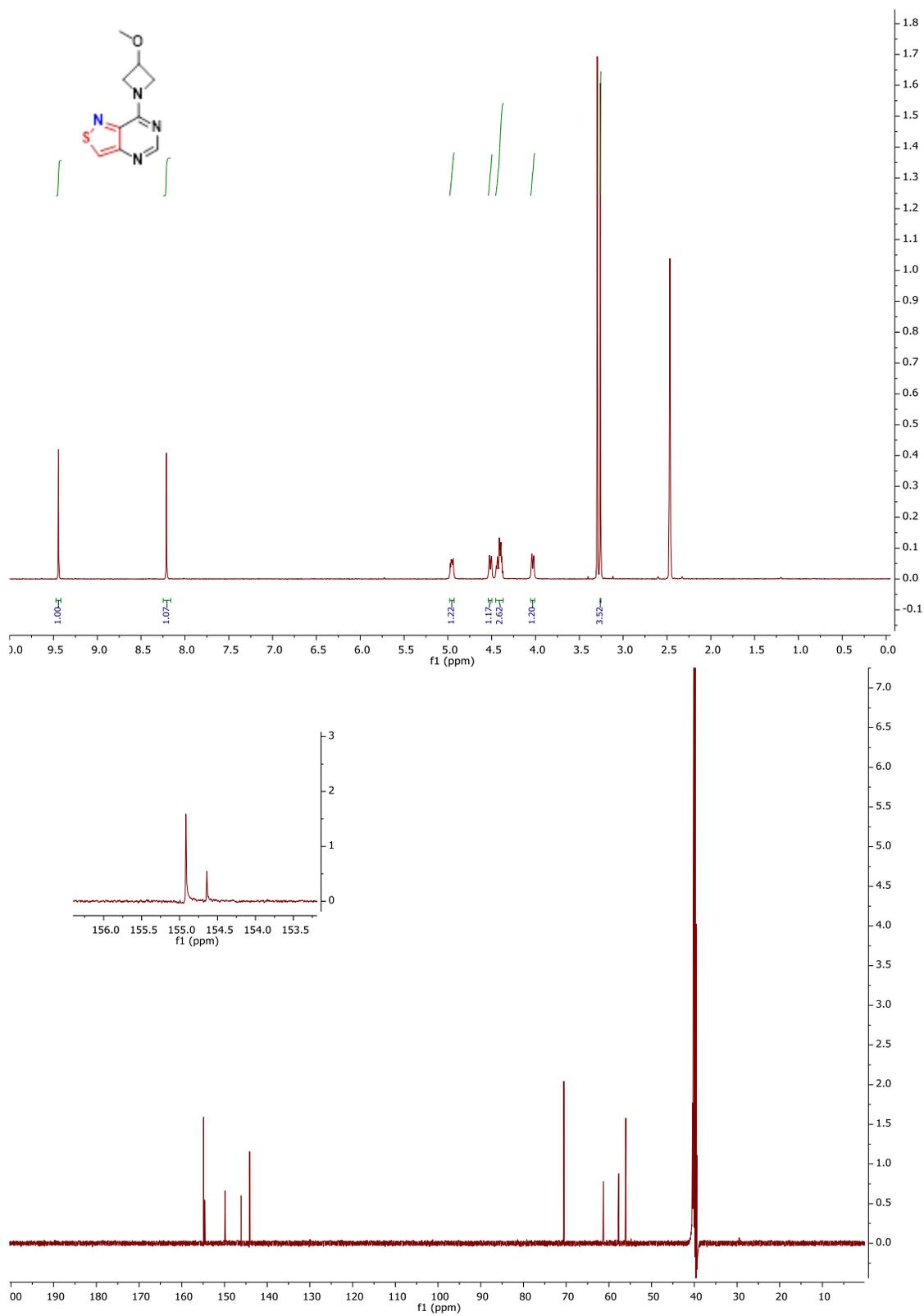
Figure 2.44. ^{1}H and ^{13}C NMR spectra of **5b**.
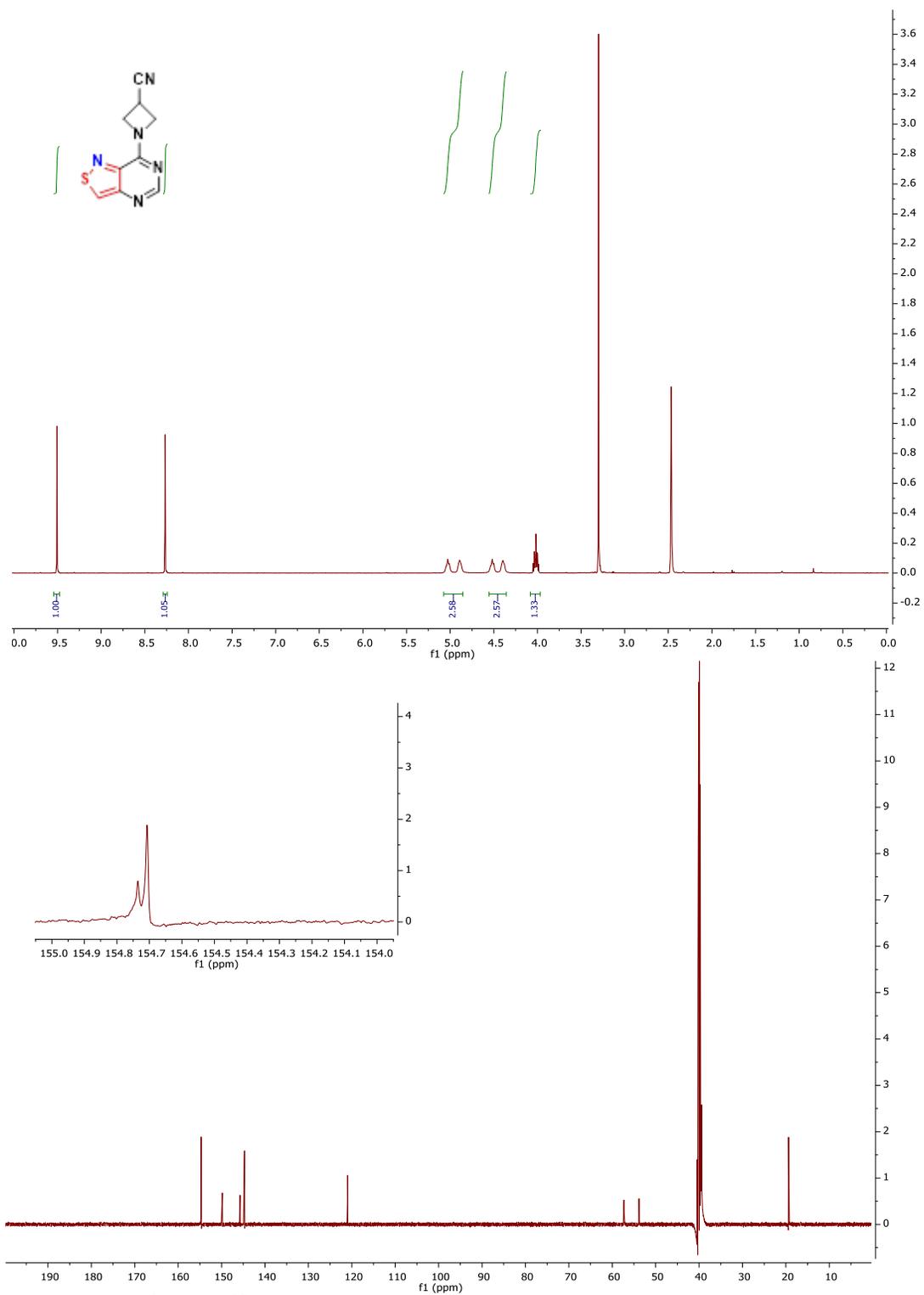
Figure 2.45. ^{1}H and ^{13}C NMR spectra of **6b**.
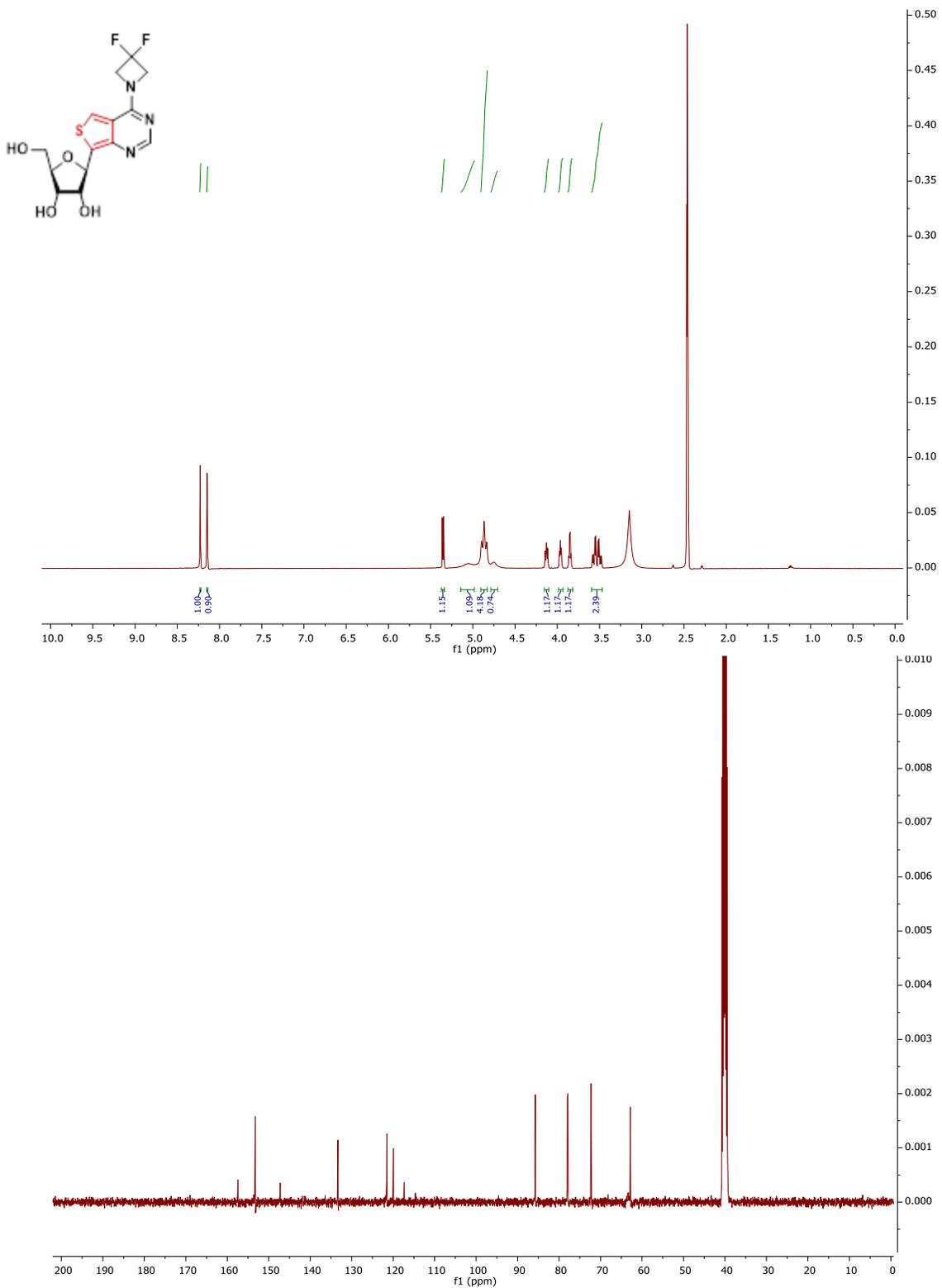
Figure 2.46. ^{1}H and ^{13}C NMR spectra of **11**.
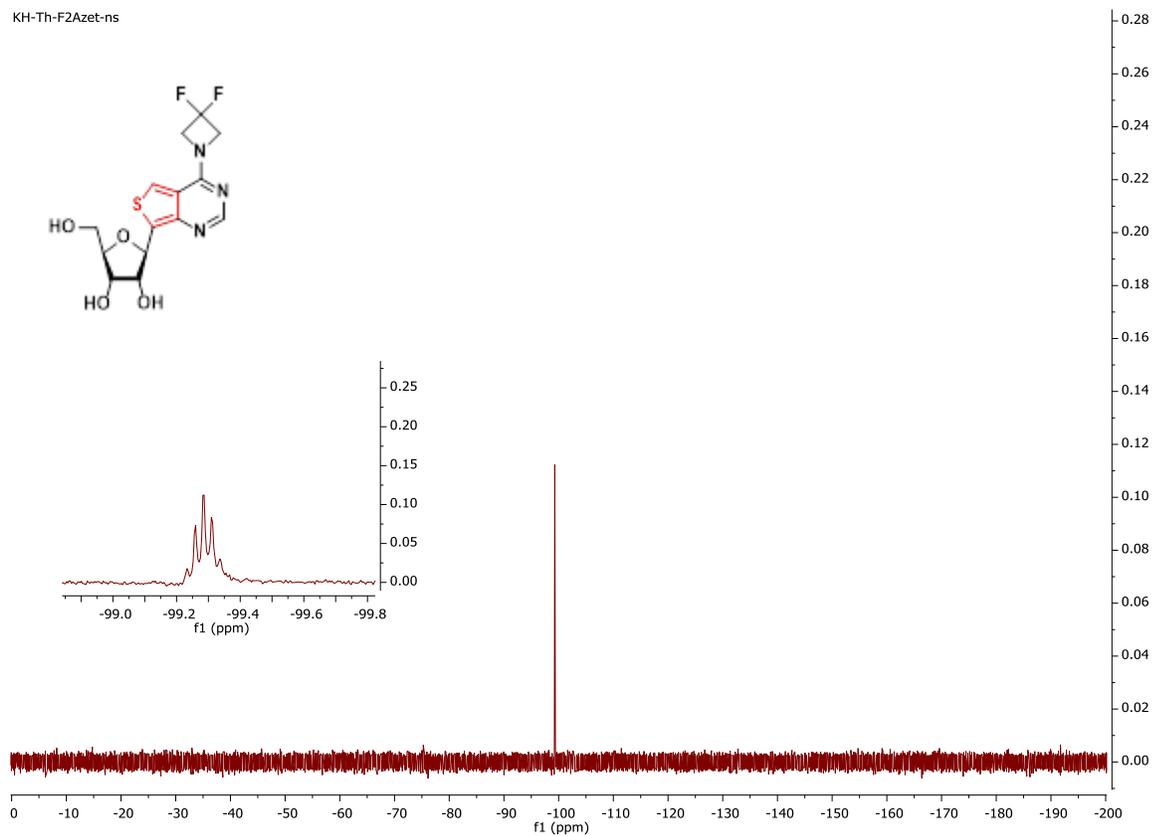
Figure 2.47. ^{19}F NMR spectra of **11**.

2.9.7 Supplementary references

(S1)   Rovira, A. R.; Fin, A.; Tor, Y. Chemical Mutagenesis of an Emissive RNA Alphabet. *J. Am. Chem. Soc.* **2015**, *137* (46), 14602–14605. https://doi.org/10.1021/jacs.5b10420.

(S2)   Shin, D.; Sinkeldam, R. W.; Tor, Y. Emissive RNA Alphabet. *J. Am. Chem. Soc.* **2011**, *133* (38), 14912–14915. https://doi.org/10.1021/ja206095a.

(S3)   Haidekker, M. A.; Brady, T. P.; Lichlyter, D.; Theodorakis, E. A. Effects of Solvent Polarity and Solvent Viscosity on the Fluorescent Properties of Molecular Rotors and Related Probes. *Bioorganic Chem.* **2005**, *33* (6), 415–425. https://doi.org/10.1016/j.bioorg.2005.07.005.

(S4)   Sinkeldam, R. W.; Marcus, P.; Uchenik, D.; Tor, Y. Multisensing Emissive Pyrimidine. *Chemphyschem Eur. J. Chem. Phys. Phys. Chem.* **2011**, *12* (12), 2260–2265. https://doi.org/10.1002/cphc.201100315.

(S5)   *Physical Properties of Glycerine and Its Solutions.*; Glycerine Producers' Association: New York, 1963.

(S6)   Vural, U. S.; Muradoglu, V.; Vural, S. Excess Molar Volumes, and Refractive Index of Binary Mixtures of Glycerol + Methanol and Glycerol + Water at 298.15 K and 303.15 K. *Bull. Chem. Soc. Ethiop.* **2011**, *25* (1). https://doi.org/10.4314/bcse.v25i1.63374.

(S7)   Afshar Ghotli, R.; Abdul Aziz, A. R.; Atadashi, I. M.; Hasan, D. B.; Kong, P. S.; Aroua, M. K. Selected Physical Properties of Binary Mixtures of Crude Glycerol and Methanol at Various Temperatures. *J. Ind. Eng. Chem.* **2015**, *21*, 1039–1043. https://doi.org/10.1016/j.jiec.2014.05.013.

(S8)   Kosower, E. M.; Dodiuk, H.; Kanety, H. Intramolecular Donor-Acceptor System. 4. Solvent Effects on Radiative and Nonradiative Processes for the Charge- Transfer States of N-Arylaminonaphthalenesulfonates. *J. Am. Chem. Soc.* **1978**, *100* (13), 4179–4188. https://doi.org/10.1021/ja00481a028.

CHAPTER 3: Synthesis of fluorescent puromycin analogues

3.1 Introduction
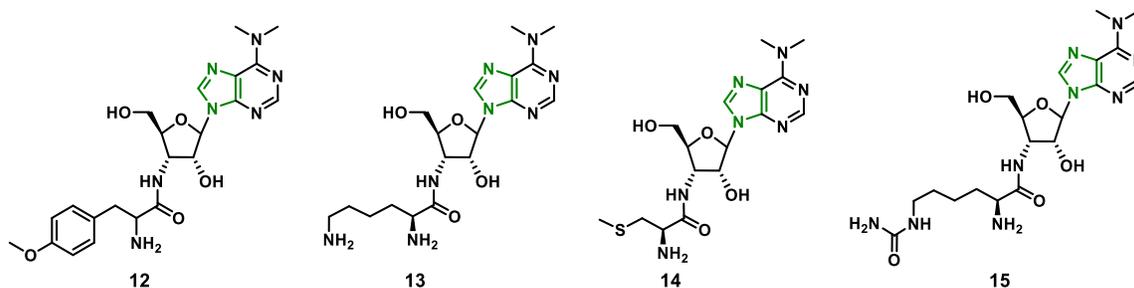
Figure 3.1. Chemical structure of puromycin (**12**) and similar antibiotics (lysylaminoadenosine **13**, Cystocin **14**, and homocitrullylaminoadenosine **15**) isolated from various strains of *streptomyces* bacteria (reiterated from figure 1.5)

Puromycin is an aminonucleoside antibiotic originally isolated from *Streptomyces Alboniger*. Its structure consists of adenosine dimethylated at the N^6 position of the purine and an O-methyl tyrosine residue bound at the 3' position of the ribonucleoside (figure 3.1) through an acetamido linkage. As one of the first nucleoside antibiotics discovered, puromycin possesses a rich history in the context of synthetic chemistry, not only in the total synthesis of the full antibiotic,[1] but in the 3'-amino-3'-deoxy nucleoside metabolite as well.[2,3] In this chapter, we will focus on synthetic strategies utilized to create puromycin and relevant analogues for the synthesis of fluorescent analogues employing both old and novel emissive nucleobases developed by Tor *et al.*[4,5]

3.2 Methodologies for Puromycin Synthesis

3.2.1 The Robins Method

A rich library of puromycin analogues containing modified amino acids, carbohydrate, and/or base moieties has been prepared. Notably, the lab of Morris J. Robins synthesized puromycin from adenosine in high overall yield through the installation of a 2',3'-epoxide[6] followed by a regiospecific epoxide ring opening of **17** with dimethylboron bromide to give the
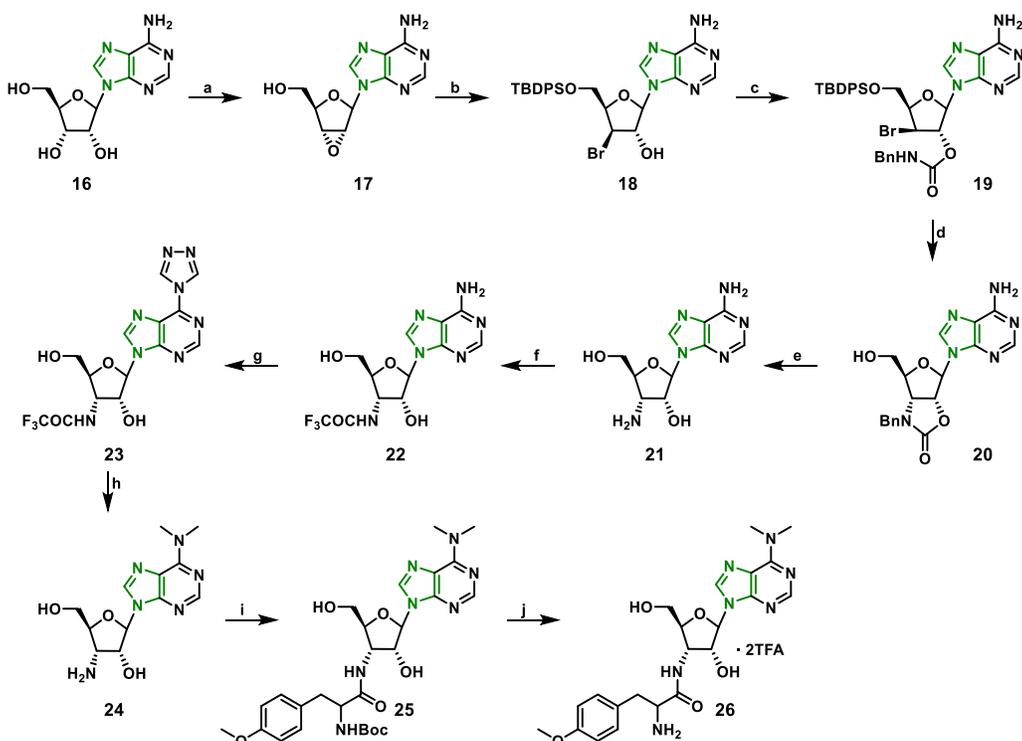
Scheme 3.1 The Robins synthesis of puromycin[1] (a). i. a-acetoxyisobutyryl bromide, ACN/H_2O; ii. DOWEX 1x2 (-OH), MeOH, 90%. (b) i. TBDPSCl, Pyridine; ii. Me_2BBr, DCM, 97%. (c) Benzyl isocyanate, Et_3N, THF/ACN, 89%. (d) i. NaH, THF, -20°C; ii. Bu_4NF, THF, 97% (e) i. NaOH, THF/H_2O; ii. Pd(OH)2, NH_4HCO_2, MeOH/H_2O, 92%. (f) i. Tf_2O, Pyridine; ii. DOWEX 1x2 (-OH), MeOH. (g) i. *N,N'*-bis[(dimethylamino)methylene]- hydrazine (BDMAMH) dihydrochloride, Me_3SiCl, Pyridine, 100 °C; ii. MeOH. (h) $HNME_2$, H_2O, 71% over 3 steps. (i) Boc (4-methoxy-L-phenylalanine), DCC, NHS, DMF, 85%. (j). i. TFA; ii. DOWEX 1x2 (-OH), 36%.

corresponding xylo isomer **18** (Scheme 3.1). The product bromohydrin was then converted into a carbamate with benzyl isocyanate, and subsequent decarbonylation produced 3-amino-3-deoxy adenosine **21**. Activation of the exocyclic amine of the adenosine heterocycle into a 1,2,4-triazole[7] followed by treatment with dimethylamine yielded the desired heterocycle, and simple peptide coupling to commercially available boc-protected 4-methoxy-L-phenylalanine yielded puromycin **12**. The same synthetic method yielded 7-deazapuromycin using 7-deazaadenosine as a precursor.[1]

3.2.2 The Strazewski Synthesis

Strazewski published a 9-step synthesis of puromycin and related analogues from adenosine in 2006, albeit missing the characteristic exocyclic dimethylamine (Scheme 3.2).[8,9] The synthetic methodology differs from that of Robins-in contrast, Strazewski's synthesis progresses through 3'-azido-3'-deoxy adenosine analogue **29**, as opposed to Robins 3'-amino-3'-deoxy intermediate **21**.[1] Strazewski synthesized this derivative through inversion of the 3'stereocenter into the xylo isomer **28**, followed by activation of the 3' alcohol with triflic chloride and subsequent displacement with lithium azide, producing the desired 3'-azido diastereomer **29**.[8] The cornerstone of the synthesis is a novel Stuadinger-Vilarassa ligation, combining the phosphine mediated azide reduction and subsequent peptide coupling into one pot to produce compounds **31a–c**. Removal of silyl and fmoc protecting groups yielded the desired puromycin derivatives **32a–c** in moderate yield.[8] Both Strazewski and Robins' synthesis provided significant inspiration in the development of our desired fluorescent analogues.
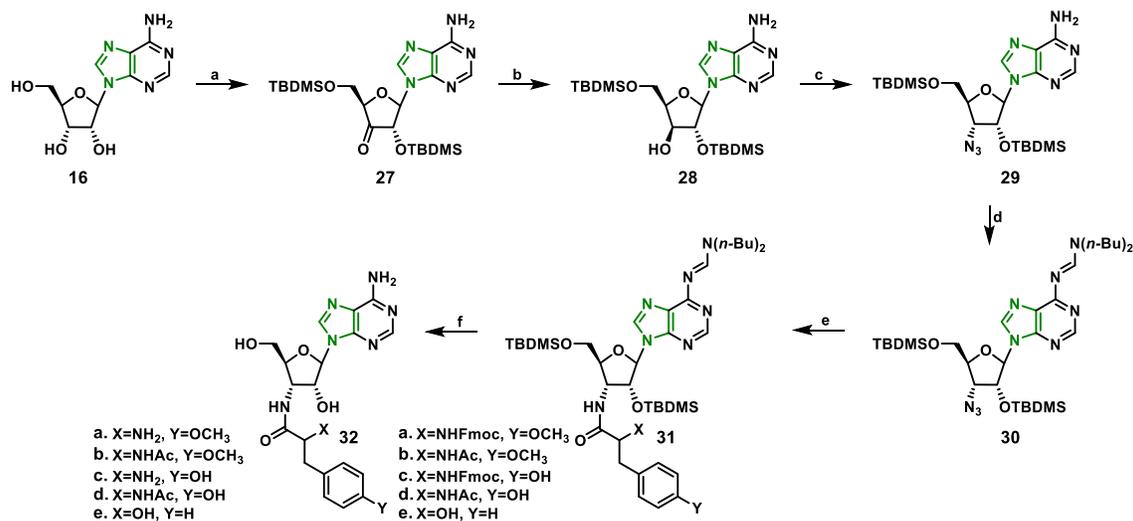
Scheme 3.2. The Strazewski synthesis of puromycin: (a) i. TBDPSCl, Pyridine, 48%; ii. CrO_3, Ac_2O, DCM, 81%. (b) $NaHB(OAc)_3$, ACN/AcOH, H_2O_2 workup, 90%. (c) i. TfCl, DMAP, DCM; ii. LiN_3, DMF, 85%. (d) N-butyl-N-(dimethoxymethyl)butan-1-amine, MeOH, 95%. (e) i. N-Fmoc/Ac-protected amino acid, DIC, HOBt, THF; ii. $(n\text{-}Bu)_3P$, 54-100%. (f) i. $NEt_3 \cdot 3HF$, THF; ii. CH_3NH_2, EtOH, HPLC.

3.3 Forethoughts for the Synthesis of Fluorescent Puromycin Analogues
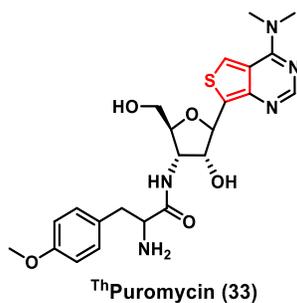
Figure 3.2. Chemical structure of target compound ThPuromycin **33**.

The endeavor to synthesize fluorescent nucleoside-amino acid conjugate Thpuromycin (**33**, figure 3.2) constitutes one of the most difficult synthetic endeavors undergone in the Tor lab. To perform such a feat, various factors needed to be considered and incorporated into the synthetic design to address functional group compatibility, ease of synthesis, number of steps, and structural variability. Here, we discuss the challenges associated with synthesizing this antibiotic and related derivatives.

3.3.1 Basicity of N^6, N^6-dimethylamino-adenosine
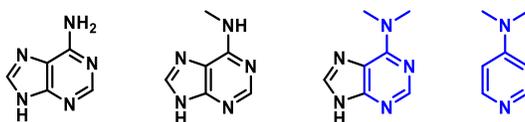
Figure 3.3. Adenine (left), N^6-methyladenine (second from left), N^6,N^6-dimethyladenine (**1**) (second from right) and DMAP (right). The similarity between the latter two are shown in blue.

Unlike N^6-monomethyl-adenosine, a commonly found modification in tRNAs,[10] the natural occurrence of dimethyl analogue **1** is exponentially rarer in nature (figure 3.3). A potential explanation for this observation is the removal of the ability to Watson-Crick base pair once the exocyclic nitrogen of adenosine is fully methylated. Base-pairing is vital to the function and recognition of the native nucleosides and their naturally occurring derivatives, so it stands to

reason that completely blocking the Watson-Crick face would not be favorable. If anything, the presence of such a modality in puromycin exemplifies the antibiotic's chemoselectivity in biological systems. However, for synthetic purposes, the exocyclic dimethylamine presents a unique problem due to the increased electron density introduced into the purine scaffold. N^6,N^6-dimethyladenine resembles 4-dimethylaminopyridine (DMAP) (figure 3.3, similarity highlighted in blue), a derivative of pyridine and a nucleophilic catalyst. DMAP is more basic than pyridine, and N^6,N^6-dimethyladenine also possesses significant basicity as well. In the context of synthetic chemistry, the presence and subsequent exposure of a basic functional group presents issues depending on the type of chemistry necessary.

3.3.2 Need for Structural Variability

The design of novel fluorescent probes ultimately relies on a gamble. Although the energies of certain electronic states within a fluorophore are predictable, other photophysical characteristics such as quantum yield and extinction coefficient remain difficult to predict, although recent advancements appear promising.[11] While this point is insignificant for compounds synthesized through simple means, many compounds need longer, more time-consuming reaction sequences that are highly inefficient when considering that a particular probe structure may ultimately lack desired photophysical properties. To circumvent this issue, flexible syntheses that facilitate structural diversity are highly desirable. If one or more designs prove lackluster, the ability to implement facile structural changes could provide rapid access to a library of probes and improve the chances of producing preferred photophysics amongst them.
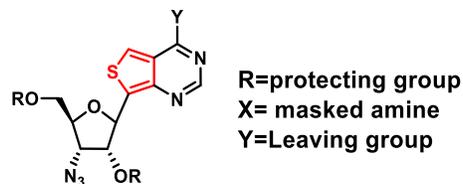
Figure 3.4. Chemical structure of a fluorescent antibiotic precursor.

In connection to synthesizing fluorescent puromycin analogues, it became imperative to incorporate a divergent synthesis allowing access to numerous antibiotic structures if one or more probes prove lackluster. The divergence stems from a common nucleoside intermediate, with reactive functionalities at the 3' position for peptide coupling reactions and a reactive leaving group at the 6-position of the purine (using native numbering) to alter the chromophore. This core structure can produce various fluorophore combinations through displacement of leaving group Y with different alkylamines, and peptide coupling of an amino acid to a 3' amine following azide reduction. The design introduces structural diversity solely at the two points predicted to affect photophysical properties – either the fluorophore itself (purine) and/or any side chains introduced by the amino acid, opening possibilities to more complex photophysical properties such as FRET and fluorescence quenching, among others.

3.3.3 C-Nucleoside Synthesis

Nucleosides containing sugars bound to a heterocyclic nitrogen or carbon atom are termed N- or C- linked nucleosides based on the identity of the atom in question. N- nucleosides constitute the majority of naturally occurring nucleosides, including the native ribo- and 2-deoxy nucleosides as well as several other non-cannonical nucleosides such as xanthine, isoguanine, and inosine. In terms of stability, the hemi-aminal of nitrogen-linked glyosidic bonds suffer susceptibility to hydrolysis by various enzymes and exposure to acidic conditions. Replacing the nitrogen atom of an N-nucleoside with a carbon atom significantly enhances their stability and

are often used in medicinal chemistry approaches as substitutes for N-nucleosides, as well as antivirals and various other drug targets. Several C-nucleosides are also naturally occurring, such as pseudouridine,[12] pirazomycin, and forodesine (figure 3.5),[13] and indeed most carbon-linked glycosides found in nature exhibit antibiotic, anticancer, and/or antiviral activity. Pseudouridine has recently found monumental importance as a substitute for uridine in the mRNA sequence of the Pzifer COVID-19 vaccine, increasing bioavailability and overall half-life of an otherwise delicate oligonucleotide.

pseudouridine (34)   pirazomycin (35)   forodesine (36)
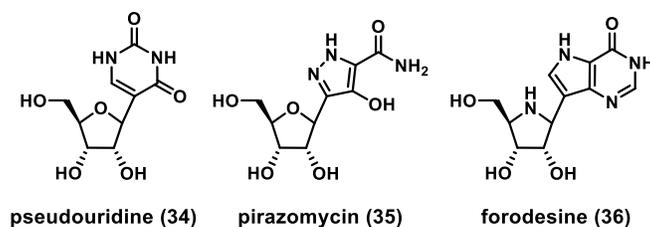
Figure 3.5. C-nucleosides pseudouridine (**34**), pirazomycin (**35**), and forodesine (**36**) found in nature.[12,13]

As C-nucleosides continue to demonstrate efficacy across various sections of medicinal chemistry and chemical biology, many synthetic strategies have been developed to construct novel structural frameworks. Given that the target antibiotics are funtionalized C-nucleoside analogues, the synthetic approaches commonly applied for their construction are also important in this context. The methods utilized for novel C-nucleoside synthesis can be divided into several approaches. These include functionalization of already existing C-nucleosides, direct coupling of a nucleobase or nucleobase precursor to sugar molecules, and lastly synthesizing a nucleobase from a modified sugar skeleton.[14] Here we will discuss the first two methods in the context of this work.

3.3.3.1 Glycosylation

Formation of a carbon-carbon bond between a nucleobase and the anomeric position of a sugar can be performed through various types of chemistries. The most frequently used methods involve Lewis acid mediated Friedal Crafts reactions, organometallic catalysis, and Grignard-type nucleophilic addition/substitution.[14] While the selection of a particular reaction is dependent on the reactivities of the substrates in question, it important to note that all the methods listed require the presence of a moderately electron rich heterocycle to facilitate sigma bond formation. Most aromatic nucleobases fulfill this requirement. However, there have been issues employing glycosylation methods in C-nucleosides containing more electron-poor aromatics.

Lewis acid-mediated glycosylation possess a distinct advantage in the synthesis of novel ribonucleosides through the directing effect of the neighboring 2' oxygen (figure 3.6). The substrate for reactions of this type, aside from the nucleobase, are tetra-O-esterified ribose derivatives with an 1' O-acetylated hemiacetal and acetate or benzoate protecting groups on the 2, 3, and 5' alcohols. Lewis acid coordination of the acetate at the anomeric position soon yields the corresponding oxocarbenium intermediate, which is then available for attack by the desired nucleobase. Significantly, stabilization of the oxocarbenium ion by the 2' esters directs carbon-carbon bond formation to occur from the β face, which is a highly desired stereochemical orientation considering most if not all naturally occurring nucleosides contain glycosidic linkages in the β configuration.[14]
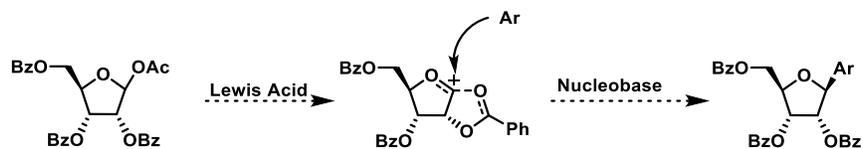
Figure 3.6. Mechanism of lewis-acid mediated glycosylation. β-selectivity is driven through neighboring-group participation of the 2' benzoate.

Palladium catalyzed synthesis of C-glycosides is particularly useful in the formation of deoxynucleoside analogues. A standard Heck coupling is usually employed between a halogen-modified nucleobase and ribose analogue containing an alkene between the 1' and 2' positions, forming a 2'-3' enol ether (figure 3.7). Subsequent reduction of the enol ether yields the desired deoxynucleoside. The stereochemistry observed is highly dependent on the substrates as well as the specific catalyst/ligands used.[14]
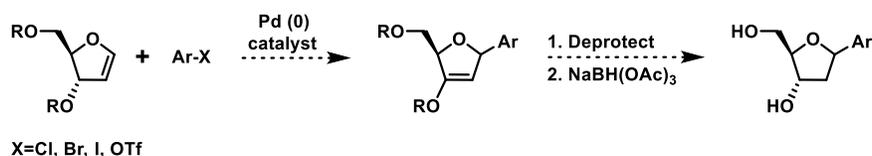
Figure 3.7. C-glycosylation of aromatic systems through palladium catalysis.

Halogenated aromatics can provide unique access to certain nucleoside scaffolds through the formation of an organometallic nucleobase. Grignard reagent formation and lithium halogen exchange are commonly utilized to form the nucleophilic intermediate, which subsequently reacts with an electrophilic sugar moiety at the anomeric carbon to form a new carbon-carbon bond. The formation of the organometallic aromatic can be tuned through the addition of certain additives such as HMPA and TMEDA which affect the butyllithium substructure. Organolithium and organomagnesium compounds tend to be extremely reactive as C-Li bond polarity results in the carbon attracting most of the electron density and resembling a carbanion. The "harder"

reactivity of these intermediates can often be reduced through transmetallation to "softer" metals such as zinc and copper allowing for milder reactions to soft electrophiles (figure 3.8).
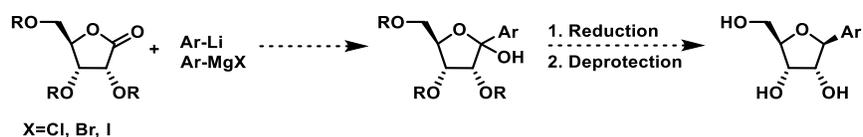
Figure 3.8. C-glycosylation of halogenated aromatics through lithium-halogen exchange

3.4 3'-Azido-3'-deoxy Nucleoside Synthesis

In our initial plan for the synthesis of fluorescent antibiotics, we decided to create a 3'-deoxy-3'-azido protected ribose analogue as a precursor to the full antibiotic skeleton through lewis acid-mediated glycosylation and peptide coupling reactions based on previously published procedures (figure 3.9, figure 3.10).[15] Whilst the 3'α-azide provides easy access to an amine through classical azide reduction methods, the 2' and 5' oxygens bear benzoate protecting groups in order to direct incoming heterocycles to the β face at the anomeric position following oxocarbenium formation.

We further prepared two nucleobases, the formyl-protected 4-aminothiopene **37** and nucleobase **2a** as substrates for glycosylation with sugar **44** (scheme 3.3).[4] Additionally, the synthesis of fluorescent nucleosides through glycosylation of a preformed nucleobase is often favorable as it allows one to preliminarily assess the photophysical properties of the chromophore without fully investing in the formation of the full nucleoside, which was performed in chapter 2.

While the synthesis may ultimately begin from glucose, we found it more efficient to start from the commercially available 1,2:5,6-Di-O-isopropylidene-α-D-glucofuranose **38** as reported previously in the literature (figure 3.10).[15] Activation of the unprotected 3' alcohol with triflic anhydride followed by displacement with sodium azide yielded the product alkylazide **39**

in moderate yield. It should be noted that only 3' triflates are sufficiently reactive enough to undergo an Sn2 reaction due to the conformation of the sugar pucker and steric hinderance from
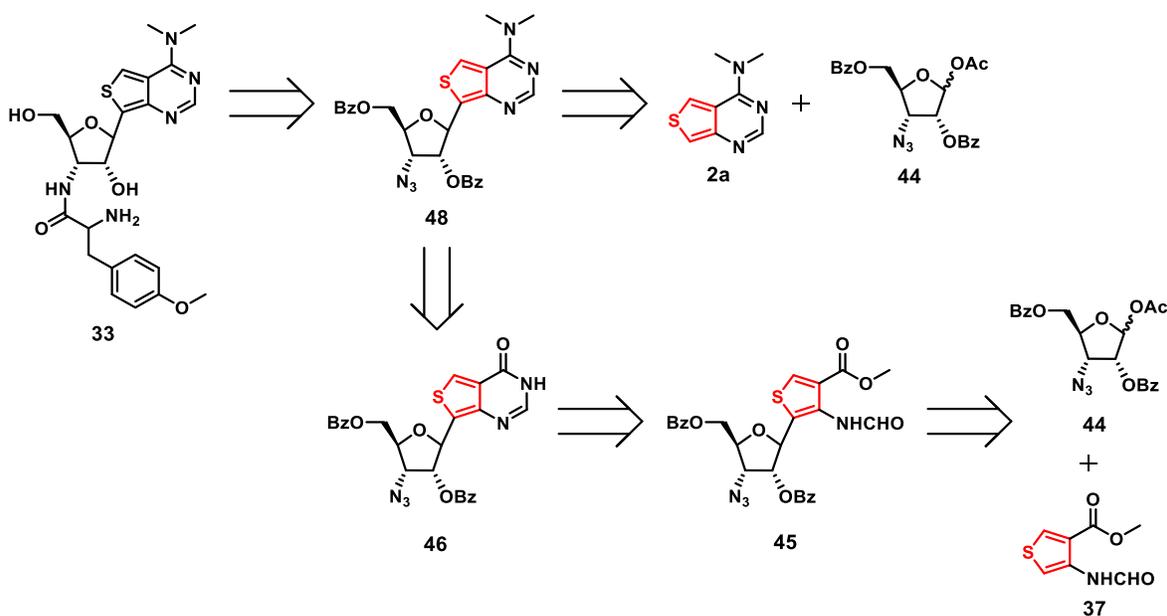
Figure 3.9. Retrosynthetic pathway from ThPuromycin **33** to potential starting precursors **37** and **44**.
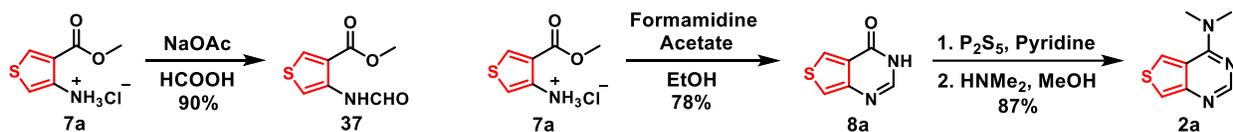
Scheme 3.3. Synthesis of formyl-protected thiophene **37** and nucleobase **2a**.

the neighboring 1'-2' isopropylidene.[16] Additionally, the other major product observed was a dehydrated product, with a double bond between the 3' and 4' positions of the ribose ring. Deprotection of the 5',6' isoproylidene group gave vicinal diol **40** and successive cleavage with sodium periodate and borohydride reduction gave the protected ribo derivative **42**. Isopropylidene removal in acidic methanol and protection of the newly free 2' alcohol gave acetal **43**, which was acetylated with acetic anhydride and acetic acid to give protected sugar **44**.

Stannic chloride glycosylation of sugar **44** with thiophene **37** successfully yielded

riboside **45** in a modest 27% yield (figure 3.11). Acid-mediated cleavage of the formyl protecting

group followed by cyclization with formamidine acetate yielded inosine analogue **46**. However,
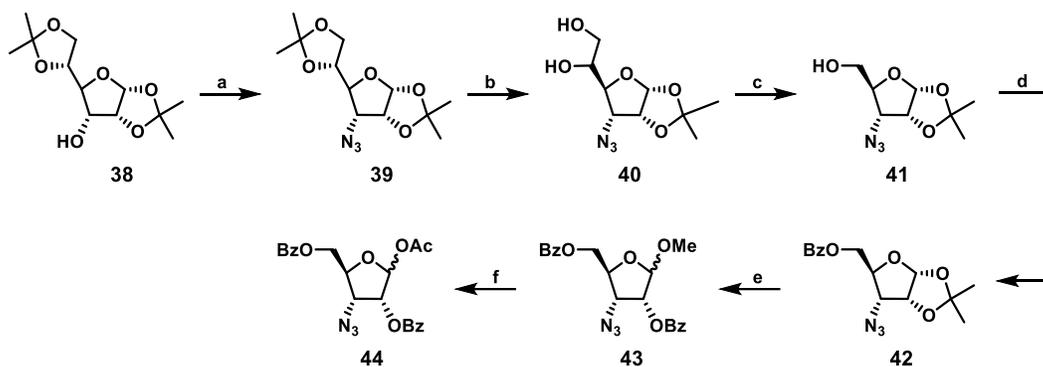
Figure 3.10. Synthesis of 1'-O-acetyl-3'-azido-3'-deoxy sugar **44** for lewis-acid mediated c-glycosylation.[15] a). i. Tf$_2$O, Pyridine, 1 hr 0 °C; ii. NaN$_3$, DMF, ON, 35%. (b) 75% AcOH, 2 hr, 80%. (c) i. NaIO$_4$, EtOH, 1.5 hr, 0 °C; ii. NaBH$_4$, 2 hr, 0 °C, 65%. (d) BzCl, Pyridine, 0 °C to RT, 96% (e) i. AcCl, MeOH, 3 h; ii. BzCl, Pyridine, 1 h, 70%. (f) Ac$_2$O, AcOH, 55 °C, 30 min, 90%

standard protocols for the activation of the heterocyclic amide with phosphorous pentasulfide

proved unsuccessful due to the temperature requirements necessary for the sulfurylation to

progress, as the 3' azide proved incompatible. Due to the technical difficulty of running

glycosylation reactions as well as the poor yields attained for compound **46**, various reaction

conditions were tested on nucleobase **8a** synthesized as a precursor for analog **2a**. Unfortunately,

exposing **8a** to various reactions including Vilsmeier–Haack and Mitsunobu did not yield the
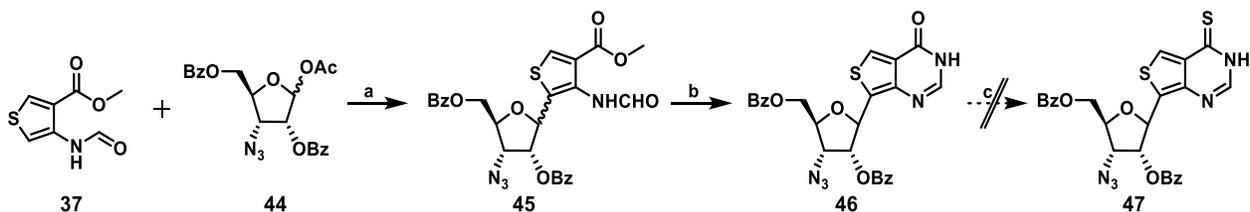
Figure 3.11. Attempted synthesis of 3'-azido-3'-deoxy c-nucleoside **47**. The final thionation step proved unsuccessful. (a). SnCl$_4$, MeNO$_2$, 60 °C 27%. (b) Formamidine acetate, EtOH, reflux, 20%. (c) P$_2$S$_5$, Pyridine, 105 °C, 3 hr.

target products (figure 3.12), prompting investigations into glycosylation using the fully-formed nucleobase **2a** instead (figure 3.13).
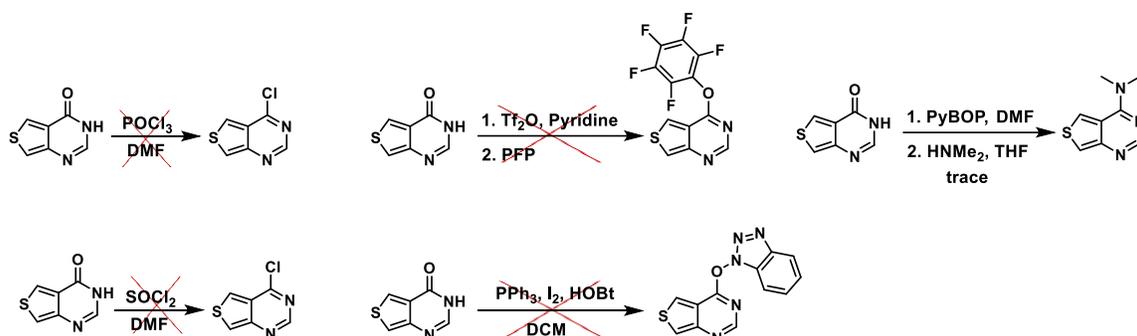
Figure 3.12. Different chemical reactions attempted to activate the imide in the inosine nucleobase analogue. All were unsuccessful.

Lewis acid-mediated glycosylation attempts between fully formed nucleobase **2a** and sugar **44** initially appeared successful, however problems arose during subsequent benzoate deprotection attempts of expected product **49** (figure 3.13). Despite a large swath of conditions tested, including amine-, hydroxide-, alkoxide-, acidic-, and lewis-acid mediated benzoate ester cleavage, the entire nucleoside underwent decomposition. This prompted us to analyze the compound further, and we were surprised to discover that N-glycoside **49** had formed in place of the desired C-glycoside **48**. Crystallographic analysis confirmed this theory indicating the positively charged N-nucleoside had eluted with a hexachlorotin ion, corroborating the strange reactivity observed with the different conditions utilized for benzoate cleavage (figure 3.14). Given that this compound contained incorrect bond connectivity, different chemistries were explored to attain the desired C glycoside.
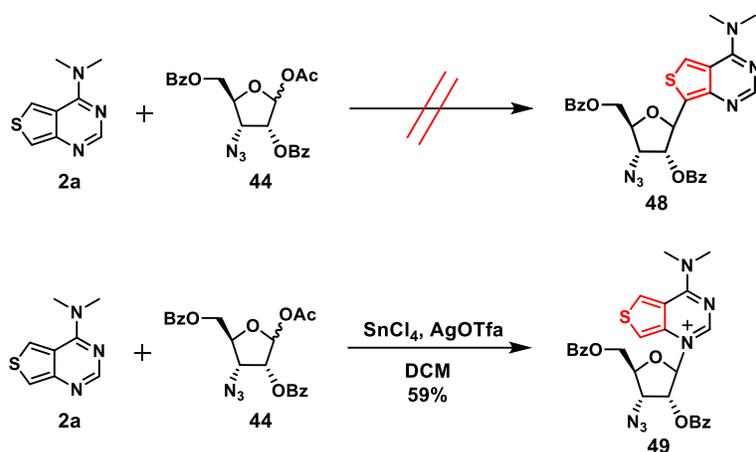
Figure 3.13. Reaction conditions leading to formation of n-nucleoside 49 instead of c-nucleoside **48**.
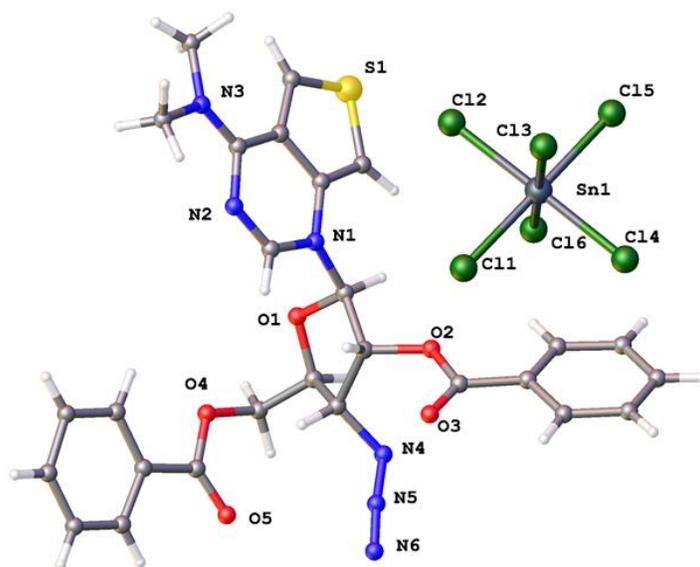
Figure 3.14. X-ray crystal structure of n-nucleoside **49**.

3.5 Lithiation and organolithium based reactions

Although the lewis-acid mediated glycosylation attempts failed, significant information was obtained about the reactivity of the species in question: nucleobase **2a** possessed sufficient electron density required to react with a designated sugar analogue, even though it was reacting through an undesired electron flow. From this point, we hypothesized that electron-rich nucleobase **2a** could still perform the desired carbon-carbon bond formation through the

installation of a reactive handle on C9 position of the purine (according to native numbering). For this reactive handle, we opted to introduce a brominate the C9 carbon with the intention of forming an organometallic nucleobase through lithium halogen exchange. Through halogenation of the nucleobase, a lithium halogen exchange reation with an alkyllithium could potentially produce an aromatic carbanion at a specific position, allowing subsequent bond formation to occur in the desired regioselective fashion. Inspired by this idea, we synthesized brominated nucleobase **50** by subjecting the dimethylated nucleobase **2a** to standard bromination conditions (scheme 3.4).
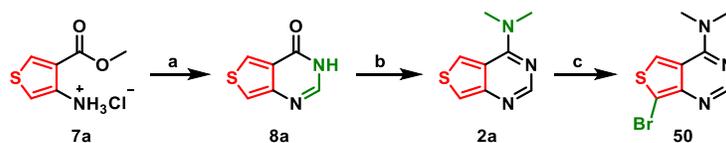
Scheme 3.4. Full Synthesis of brominated nucleobase **50**: (a) Formamidine acetate, EtOH, reflux, ON, 78%. (b) i. P_2S_5, Pyridine, 110 °C, 2 h; ii. 2M $HNMe_2$ in MeOH, 60 °C, ON, 87%. (c) Br_2, AcOH, 100 °C, 2.5 h, 58%.

We also synthesized azidolactone **55** as a corresponding electrophilic moiety to react with the organolithium intermediate and produce a C-glycoside (scheme 3.5). The compound was synthesized from D-xylose **51** mimicking a synthetic pathway used for the synthesis of sugar **44**. Initially, the xylofuranose was protected at the 1' and 2' positions via an isoproylidene and at the 5' position with a benzyl group. Activation of the 3' alcohol with triflic anhydride and subsequent displacement with an azide ion yielded the desired 3' α-azide **53** in 53% yield. Isoproylidene deprotection with acidic methanol and protection of the newly freed 2' alcohol produced intermediate **64**, with a methoxy group at the anomeric position, two benzyl groups at the 2' and 5', and an azide at the 3' of the ribose. Hydrolysis of the anomeric acetal with hydrochloric acid yielded the hemi-acetal, which was then oxidized to the desired lactone **55** with pyridinium chlorochromate. This intermediate was deemed sufficient as an "acceptor" for

lithium halogen exchange, with robust benzyl groups protecting the alcohols and a lactone as a

mild electrophile.
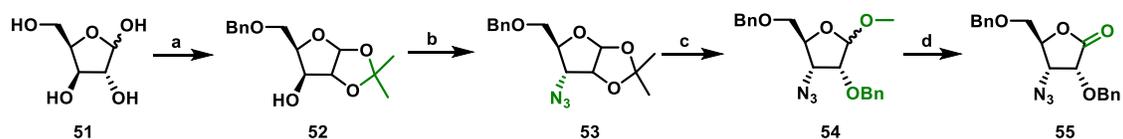
Scheme 3.5. Synthesis of azidolactone **55**: (a) i. H₂SO₄, Acetone, RT, 2.5 hr; ii. TBAI, DIPEA, BnBr, 90 °C, 22h, 65%. (b) i. Tf₂O, Pyridine, DCM, 0 °C, 2h; ii. NaN₃, DMF, RT, 48 hr, 53%. (c) i. AcCl, MeOH, RT, ON; ii. NaH, BnBr, TBAI, DMF, 0 °C to RT, ON, 83%. (d) i. HCl, H₂O, Dioxane, 60 °C, ON; ii. PCC, DCM, RT, ON, 78%.
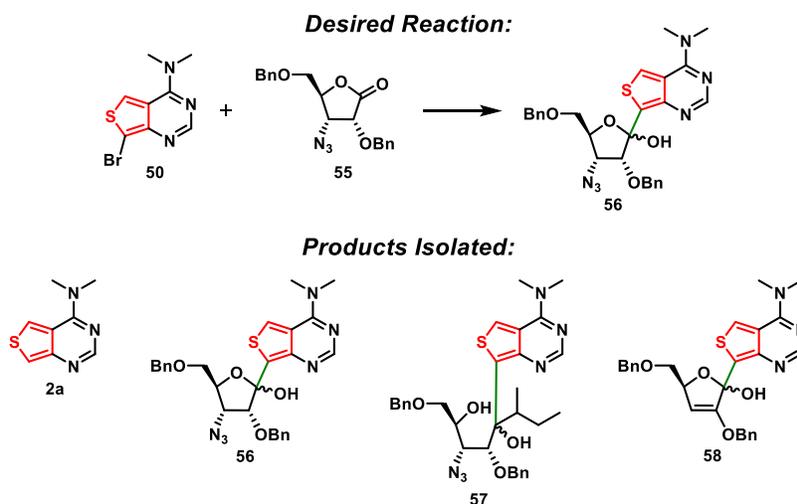
Figure 3.15. Lithium halogen exchange reaction conditions tested between nucleobase 50 and azidolactone **55**.

Table 3.1. Lithium halogen exchange reactions conditions tested for the formation of **56**.

Time	Solvent	Alkyllithium (3 eq)	Temp	Lactone eq	Lactone addition	Product Yield
1 hr	THF	s-Buli	-78 °C	2	dropwise	15-20%
1 hr	Toluene	s-Buli	-78 °C	2	dropwise	trace
30 min	THF	s-Buli	-78 °C	2	dropwise	19%
1 hr	THF	s-Buli (2 eq)	-78 °C	2	dropwise	19%
1 hr	THF	s-Buli	-78 °C	4	dropwise	19%
1 hr	THF	s-Buli	-78 °C	2	dropwise	32%
1 hr	THF	t-Buli	-78 °C	2	dropwise	33%
1 hr	THF	t-Buli	-78 °C	2	dropwise	40%

Lithium halogen exchange with nucleobase **50** and sugar **55** was examined under varying reaction conditions (figure 3.15, table 3.1). The reactions were all performed at -78 °C, changing solvent, equivalents, alkyllithiums, and mode of addition. The first reaction utilized sec-butyllithium and 2 equivalents of lactone with a dropwise addition in tetrahydrofuran and produced the desired product in 15-20% yield, although several other side products were identified, including the dehalogenated nucleobase, alkylated product **57**, and retro-michael product **58**. Switching the solvent to toluene yielded trace product and using 4 equivalents of lactone **55** instead of 2 showed no change in product yield. However, changing the addition method of the lactone to cannulation increased the yield from approximately 20 to 32%. Lastly, in order to prevent the formation of alkylated product **57**, tert-butyllithium was employed due to its increased steric bulk and inability to undergo nucleophilic addition (despite its increased reactivity). This hypothesis proved correct as lithium halogen exchange with t-BuLi yielded the desired product upwards of 40%, with no alkylated product detected. A potential mechanism for reaction progressin is provided in figure 3.16.
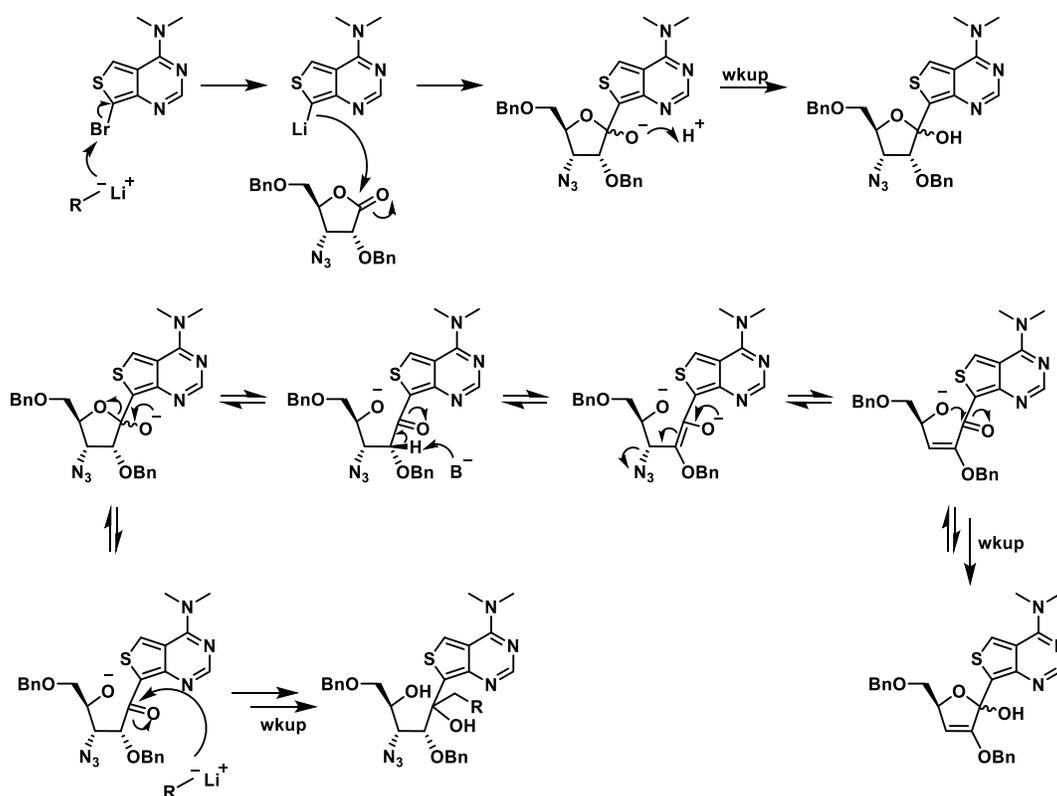
Figure 3.16. Proposed mechanism of action for carbon-carbon bond formation between nucleobase **50** and azidolactone **55** including formation of various side products through lithium-halogen exchange.

3.6 First generation synthesis of ThPuromycin

Despite low yields, the success of the lithium halogen exchange produced intermediate **56** with both a dimethylamine on the purine and a 3' azide for the first time. The manipulations remaining were hemiketal reduction, azide reduction/peptide coupling, and deprotection in a particular order (scheme 3.6). To set the desired β stereochemistry at the anomeric position, reduction of the hemiketal with BF$_3$·Et$_2$O in dry acetonitrile after stringent drying yielded the desired riboside intermediate **59**. With the planned intermediate in hand, we moved to peptide coupling to attach the O-methyl tyrosine at the 3' position. Mimicking the Strazewski synthesis, Staudinger-Vilarassa coupling of the 3' azide **59** to commercially available Fmoc protected O-methyl-L-tyrosine yielded the benzyl and fmoc protected ThPuromycin **60**. Lastly, with only two

deprotections remaining, we postulated that removal of the benzyl groups should proceed first. Although benzyl groups are easily removed through hydrogenolysis, undesired reduction of the dimethylated nucleobase was evident during the silane hemiketal reduction, so we opted to avoid this method.
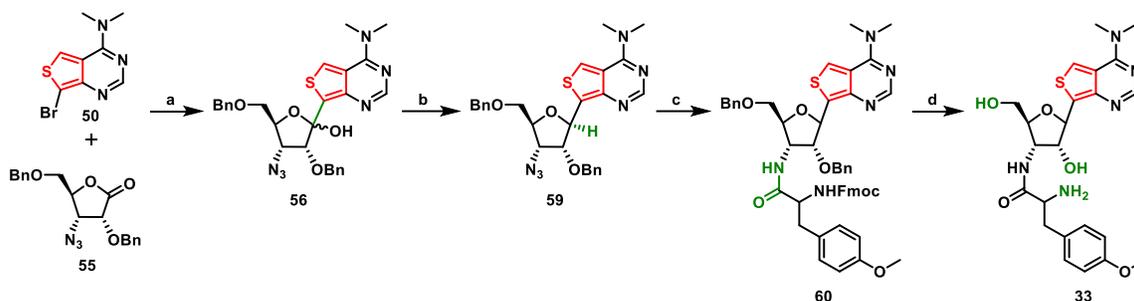
Scheme 3.6. Chemical synthesis of Thpuromycin **33**. (a) t-BuLi, THF, -78°C, 40%. (b) BF$_3$·OEt$_2$, Et$_3$SiH, ACN;DCM 3:1, -40°C to RT, 36%. (c) Fmoc-O-methyl-L-tyrosine, DCC, HOBt, (n-Bu)$_3$P, THF, 58%. (d) i. Piperidine, DMF, RT; ii. BCl$_3$, DCM, -78°C to 0°C, 60%.

Without hydrogenolysis, benzyl ethers are difficult to cleave, and the potential necessity for harsher chemistries suggested that the Fmoc protecting group should remain present to avoid any unnecessary side reactions with the free amine of the tyrosine. BF$_3$·Et$_2$O and ethanedithiol mediated benzyl removal obliterated the precursor. However, lewis acid mediated removal seemed promising, and indeed boron trichloride in DCM cleanly removed the benzyl groups. Once the sugar alcohols were freed, subjecting the crude to direct deprotection of the fmoc group using piperidine in DMF yielded the free antibiotic in good yield over two steps following HPLC purification.

3.7 2nd generation synthesis of ThPuromycin and related analogues

Upon completing the synthesis of antibiotic **33**, it became apparent that the synthesis needed to be modified to incorporate different amines at the 6 position of the purine once analysis of the azetidine modified nucleobases exhibited stronger photophysical properties. This seemed virtually impossible on paper, as we needed to install a moderately electrophilic

functional group at the 6 position of the thienopurine that would survive a similar reaction sequence as the dimethyl substituent, and lithium halogen exchange with highly reactive nucleophilic organolithium in the presence of such an electrophile seemed counterintuitive. Fortunately, this was not the case, as LHE reactions with lactone **55** and nucleobase **62** containing a thiomethyl group at the 6 position of the thienopurine and a bromine atom at the 9 position (native purine numbering) were successful (scheme 3.8). The synthesis of nucleobase **62** was performed in a similar fashion to the pathway utilized for compound **50** (scheme 3.7). Instead of displacing the sulfur of nucleobase **9a**, methylation with methyl iodide and potassium carbonate selectively produced thiomethyl product **61**, which was brominated with DBDMH to product the halogenated substrate **62** utilized for lithium halogen exchange.
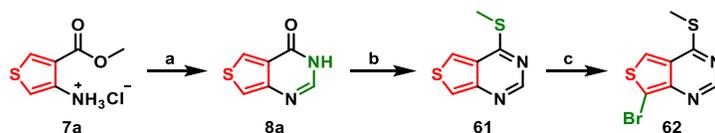
Scheme 3.7. Synthesis of brominated thiomethyl purine analog **62**. (a) Formamidine acetate, EtOH, reflux, ON, 78%. (b) i. P_2S_5, Pyridine, 110 °C, 2 h; ii. MeI, K_2CO_3, DMF, RT, 1 hr, 65%. (c) DBDMH, DMF, 0 °C, 2.5 h, 45%.
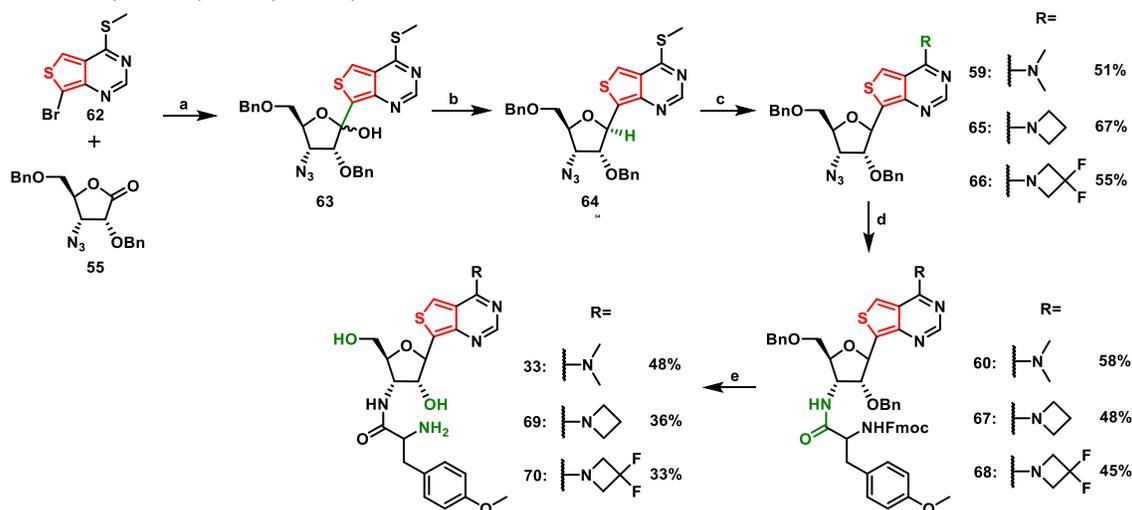
Scheme 3.8. Chemical synthesis of azetine- and difluoroazetidine-modified puromycin analogues **69** and **70**. (a) n-BuLi, THF, -78°C, 27%. (b) $BF_3·OEt_2$, Et_3SiH, DCM, -78°C to RT, 55%. (c) RH_2Cl, DBU, DMSO, RT. (d) Fmoc-O-methyl-L-tyrosine, EDC, HOBt, $(Me)_3P$, THF. (e) i. Piperidine, DMF, RT; ii. The same synthetic pathway was found to produce Thpuromycin **33** as well.
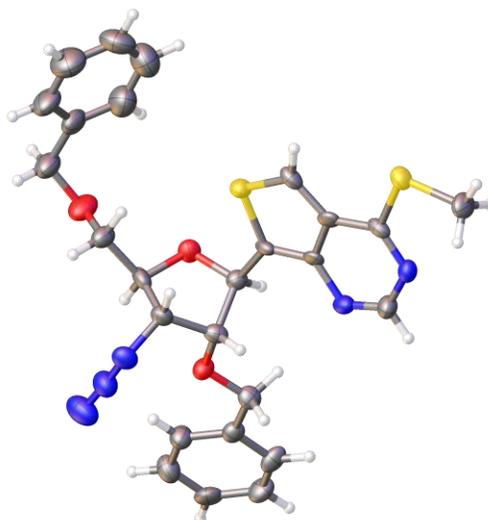
Figure 3.17. X-ray crystal structure of synthetic precursor **64**.

A similar BF$_3$·Et$_2$O mediated triethylsilane reduction of hemiketal **63** yielded nucleoside **64**, a key precursor to various antibiotic skeletons (scheme 3.8). The stereochemistry of the anomeric position was confirmed with X-ray crystallography (figure 3.17). S$_n$Ar reactions displacing the thiomethyl group of **64** with differing secondary amines in DMSO allowed for production of fluorescent adenosine analogues **59**, **65** and **66,** the latter two containing azetidines. Subsequent reduction of the 3' azide with trimethylphosine in THF followed by peptide coupling with Fmoc protected O-methyl-L-tyrosine yielded the fully protected azetidino- and difluoroazetidino [Th]Puromycin derivatives **67** and **68**. Classical Fmoc deprotection with piperidine in DMF followed by benzyl deprotection with boron trichloride produced the fluorescent antibiotics **69** and **70** in good yield following HPLC purification.

3.8 Concluding remarks

The fluorescent antibiotics were produced using two different methods employing convergent synthesis with 17- and 18 steps, respectively. Lewis acid mediated methods proved

unsuccessful, producing the undesired N-glycoside, whereas lithium halogen exchange and conversion to organolithium intermediates followed the desired reaction sequence. The extensive synthetic studies performed showcases the difficulty of C-glycoside and thiophene chemistry and the significant electronic and reactive differences displayed by the dialkylated heterocycles. This is further reinforced by the photophysics studies performed in the following chapter.

3.9 Acknowledgements

Chapters 3 is currently being prepared for submission for publication of the material. Hadidi, K; Steinbuch, K; Tor, Y. The dissertation author was the primary investigator and author of the material.

3.10 References

1. Robins, M. J., Miles, R. W., Samano, M. C. & Kaspar, R. L. Syntheses of Puromycin from Adenosine and 7-Deazapuromycin from Tubercidin, and Biological Comparisons of the 7-Aza/Deaza Pair1. *J. Org. Chem.* **66**, 8204–8210 (2001).

2. García-Alles, L. F., Magdalena, J. & Gotor, V. Synthesis of Purine and Pyrimidine 3'-Amino-3'-deoxy- and 3'-Amino-2',3'-dideoxyxylonucleosides. *J. Org. Chem.* **61**, 6980–6986 (1996).

3. Samano, M. C. & Robins, M. J. High-yield synthesis of 3′-amino-3′-deoxynucleosides. Conversion of adenosine to 3′-amino-3′-deoxyadenosine. *Tetrahedron Lett.* **30**, 2329–2332 (1989).

4. Shin, D., Sinkeldam, R. W. & Tor, Y. Emissive RNA Alphabet. *J. Am. Chem. Soc.* **133**, 14912–14915 (2011).

5. Rovira, A. R., Fin, A. & Tor, Y. Chemical Mutagenesis of an Emissive RNA Alphabet. *J. Am. Chem. Soc.* **137**, 14602–14605 (2015).

6. Hansske, F. & J. Robins, M. Regiospecific and stereoselective conversion of ribonucleosides to 3′-deoxynucleosides. A high yield three-stage synthesis of cordycepin from adenosine. *Tetrahedron Lett.* **26**, 4295–4298 (1985).

7. Miles, R. W., Samano, V. & Robins, M. J. Nucleic Acid Related Compounds. 86. Nucleophilic Functionalization of Adenine, Adenosine, Tubercidin, and Formycin Derivatives via Elaboration of the Heterocyclic Amino Group into a Readily Displaced 1,2,4-Triazol-4-yl Substituent. *J. Am. Chem. Soc.* **117**, 5951–5957 (1995).

8. Chapuis, H. & Strazewski, P. Shorter puromycin analog synthesis by means of an efficient Staudinger–Vilarrasa coupling. *Tetrahedron* **62**, 12108–12115 (2006).

9. Krishnakumar, K. S., Goudedranche, S., Bouchu, D. & Strazewski, P. The Shortest Synthetic Route to Puromycin Analogues Using a Modified Robins Approach. *J. Org. Chem.* **76**, 2253–2256 (2011).

10. Jackman, J. E. & Alfonzo, J. D. Transfer RNA modifications: Nature's combinatorial chemistry playground. *Wiley Interdiscip. Rev. RNA* **4**, 35–48 (2013).

11. Ju, C.-W., Bai, H., Li, B. & Liu, R. Machine Learning Enables Highly Accurate Predictions of Photophysical Properties of Organic Fluorescent Materials: Emission Wavelengths and Quantum Yields. *J. Chem. Inf. Model.* **61**, 1053–1065 (2021).

12. Rasmuson, T. & Björk, G. R. Pseudouridine: a modified nucleoside as biological marker in malignant lymphomas. *Cancer Detect. Prev.* **6**, 293–296 (1983).

13. Nicotra, F., Airoldi, C. & Cardona, F. 1.16 - Synthesis of C- and S-Glycosides. in *Comprehensive Glycoscience* (ed. Kamerling, H.) 647–683 (Elsevier, 2007). doi:10.1016/B978-044451967-2/00017-9.

14. Jan Štambaský, Hocek, M. & Pavel Kočovský. C-Nucleosides: Synthetic Strategies and Biological Applications. *Chem. Rev.* **109**, 6729–6764 (2009).

15. Liang, C. W., Kim, M. J., Jeong, L. S. & Chun, M. W. Synthesis of 2-(3′-Azido- and 3′-Amino-3′-deoxy-β-D-ribofuranosyl)thiazole-4-carboxamide. *Nucleosides Nucleotides Nucleic Acids* **22**, 2039–2048 (2003).

16. Swamy, K. C. K., Kumar, N. N. B., Balaraman, E. & Kumar, K. V. P. P. Mitsunobu and Related Reactions: Advances and Applications. *Chem. Rev.* **109**, 2551–2651 (2009).

3.11 Methods

3.11.1 General Methods

Reagents were purchased from Sigma-Aldrich, TCI, Spectrum, Acros, Fisher Scientific, and VWR and used without further purification. Methyl 4-amino thiophene 3-carboxylate hydrochloride was purchased from Matrix Scientific. Azetidine hydrochloride and difluoroazetidine hydrochloride were purchased from Synthonix. Solvents were purchased from Sigma-Aldrich and Fisher Scientific and dried by standard techniques. NMR solvents were purchased from Cambridge Isotope Laboratories (Andover, MA). All reactions were monitored with analytical TLC (Merck Kieselgel 60 F254). Column Chromatography was carried out with a Teledyne ISCO Combiflash RF with silica gel particle size 40-63 µm. ^{1}H NMR and GDQCOSY spectra were obtained on either a Bruker Avance 300 MHz or Jeol ECA 500 MHz. ^{13}C NMR, GHSQC and GHMBC spectra and were taken on Varian VS 500 MHz. Mass spectra were obtained on an Agilent 6230 HR-ESI-TOF MS at the Molecular Mass Spectrometry Facility at the UCSD Chemistry and Biochemistry Department.

3.11.2 Synthesis of N-glycoside 49 and halogenated nucleobase 50
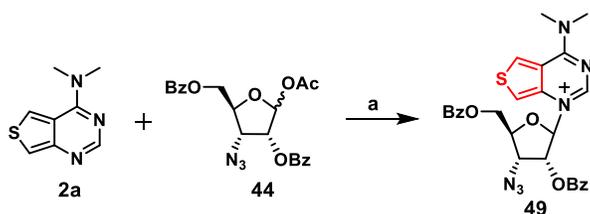
Scheme 3.9. Synthesis of N-glycoside **49**: (a) SnCl$_4$, AgOTfa, DCM, relflux, 2.5 h, 59%

1-((3R,4R,5S)-4-azido-3-(benzoyloxy)-5-((benzoyloxy)methyl)tetrahydrofuran-2-yl)-4-(dimethylamino)thieno[3,4-*d*]pyrimidin-1-ium (49)

To a flame-dried 25 ml round-bottom flask purged with argon was added AgOTfa (325 mg, 1.5 mmol, 1.5 eq) and nucleobase **50** (375 mg, 2.1 mmol, 1.5 eq). Compound **44** (575 mg,

1.35 mmol, 1 eq) was added and the components were allowed to dry on high vacuum in the presence of phosphorous pentoxide overnight. The following day anhydrous DCM was added (7 ml) and SnCl$_4$ (1 M in DCM, 650 mg, 2.5 mmol, 2.5 ml, 2.2 eq) was added dropwise. The mixture was refluxed for 3.5 h, then quenched with cold methanol and evaporated to dryness. The leftover residue was subjected to column chromatography with a gradient of 0-10% MeOH in DCM to yield a brownish solid (420 mg, 57%). ^{1}H NMR (300 MHz, DMSO-d_6) δ 9.17 (d, J = 2.9 Hz, 1H), 8.81 (s, 1H), 8.19 – 8.02 (m, 6H), 7.76 – 7.56 (m, 6H), 6.59 (d, J = 2.3 Hz, 1H), 6.03 (dd, J = 5.6, 2.2 Hz, 1H), 4.99 (dd, J = 9.0, 5.7 Hz, 1H), 4.91 – 4.70 (m, 2H), 4.63 (dd, J = 9.0, 3.5 Hz, 1H), 3.79 (s, 3H), 3.55 (s, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 165.87, 165.20, 157.32, 148.07, 147.87, 135.52, 134.66, 134.37, 134.20, 130.35, 130.21, 129.98, 129.90, 129.46, 129.20, 128.89, 117.98, 110.25, 110.16, 91.98, 80.28, 80.17, 75.05, 74.86, 62.75, 59.16, 58.91, 43.26, 43.09. ESI-HRMS calculated for C$_{27}$H$_{25}$N$_6$O$_5$S$^+$ [M+H]$^+$ 545.1602, found 545.1604.
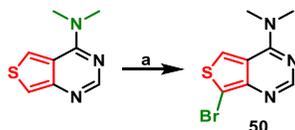
Scheme 3.10. Synthesis of nucleobase **50** from **2a**: (a) Br$_2$, AcOH, 100 °C, 2.5 h, 58%.

7-bromo-N,N-dimethylthieno[3,4-*d*]pyrimidin-4-amine (50)

To a flame-dried 500 ml round-bottom flask purged with argon was added **2a** (3.4 g, 18.7 mmol) and AcOH (282 ml). Once the starting material was solubilized, bromine (3.1 g, 19.5 mmol, 1 ml, 1 eq) was added dropwise at room temperature and the solution was stirred at 100 °C for 2.5 hours. Upon reaction completion the solution was cooled to room temperature and quenched with saturated Na$_2$S$_2$O$_3$. The volatiles were evaporated off and the crude residue was partitioned between DCM and water. The aqueous phase was washed 3 times with DCM and the

organic phases were combined, dried with sodium sulfate, filtered, and evaporated to dryness. The leftover residue was subjected to column chromatography with a gradient of 0-5% MeOH in DCM to yield a yellowish solid (2.8 g, 58%). 1H NMR (500 MHz, DMSO-d_6) δ 8.72 (s, 1H), 8.15 (s, 1H), 3.45 (s, 6H). 13C NMR (126 MHz, DMSO- d_6) δ 157.72, 154.86, 149.66, 126.56, 120.29, 101.29. ESI-HRMS calculated for $C_8H_9N_3S$ [M+H]$^+$ 257.9695, found 257.9696.

3.11.3 Synthetic procedures for 3'-azido-3'deoxy lactone sugar preparation
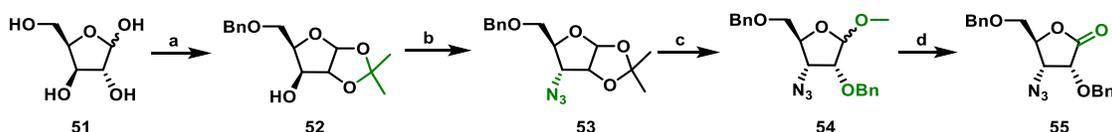
Scheme 3.11. Reiteration of scheme 3.5: Synthesis of azidolactone **55**: (a) i. H_2SO_4, Acetone, RT, 2.5 hr; ii. TBAI, DIPEA, BnBr, 90 °C, 22h, 65%. (b) i. Tf_2O, Pyridine, DCM, 0 °C, 2h; ii. NaN_3, DMF, RT, 48 hr, 53%. (c) i. AcCl, MeOH, RT, ON; ii. NaH, BnBr, TBAI, DMF, 0 °C to RT, ON, 83%. (d) i. HCl, H_2O, Dioxane, 60 °C, ON; ii. PCC, DCM, RT, ON, 78%.

(3aR,5R,6S,6aR)-5-((benzyloxy)methyl)-2,2-dimethyltetrahydrofuro[2,3-*d*] [1,3]dioxol-6-ol (52)

Compound **52** was synthesized according to previously published procedures. Briefly, D-xylose (50 g, 333 mmol) was dissolved in a solution of sulfuric acid (50 ml) in acetone (1.3 L). The solution was allowed to stir for 30 minutes at room temperature using a water bath to keep the temperature constant. A solution of sodium carbonate (65 g, 613 mmol, 1.8 eq) in water (560 ml) was added slowly at room temperature and allowed to stir for 2.5 hours. Solid sodium carbonate was then added (35 g, 330 mmol, 1 eq) and the solution was stirred for 20 minutes. The resulting solid was filtered, and the filtrate was evaporated to dryness leaving a crude oil. To the crude was added DIPEA (85.23 g, 660 mmol, 115 ml, 2 eq), benzyl bromide (90 g, 526 mmol, 62.5 ml, 1.6 eq) and tetrabutylammonium iodide (36.75 g, 100 mmol, 0.3 eq). The

reaction mixture was stirred at 90 °C for 5 hours. Both DIPEA (85.23 g, 660 mmol, 115 ml, 2 eq) and benzyl bromide (90 g, 526 mmol, 625 ml, 1.6 eq) were added again, and the reaction was stirred at 90 °C for 17 hours. The reaction mixture was then cooled to room temperature and partitioned between DCM and water. The aqueous phase was washed with DCM 6 times and the organic phases were combined, dried over sodium sulfate, filtered, and evaporated to dryness. The resulting residue was subjected to column chromatography with a gradient of 0 to 45% EtOAc in hexanes to yield a clear oil (60.7 g, 65% over 2 steps) that was recrystallized from toluene/pentane. ^{1}H NMR (500 MHz, chloroform-d) δ 7.36 – 7.28 (m, 5H), 5.95 (d, J = 3.7 Hz, 1H), 4.58 (ABq, J = 11.9 Hz, 2H), 4.48 (d, J = 3.7 Hz, 1H), 4.26 – 4.20 (m, 2H), 3.90 (dd, J = 10.9, 4.1 Hz, 1H), 3.86 (dd, J = 10.9, 3.6 Hz, 1H), 3.68 (d, J = 3.0 Hz, 1H), 1.47 (s, 3H), 1.29 (s, 3H).^{13}C NMR (126 MHz, Chloroform-d) δ 137.15, 128.60, 128.09, 127.93, 111.59, 104.84, 85.30, 78.18, 76.29, 74.05, 68.16, 26.77, 26.19. ESI-HRMS calculated for $C_{15}H_{20}O_5Na$ [M+Na]$^+$ 303.1203, found 303.1202.

(3aR,5S,6R,6aR)-6-azido-5-((benzyloxy)methyl)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxole (53)

Compound **52** (60.7 g, 217 mmol) was azeotroped with toluene 3 times, then dissolved in anhydrous DCM. Anhydrous pyridine (43 ml) was added, and the solution was cooled to 0 °C with an ice bath and trifluoromethanesulfonic anhydride (91.9 g, 326 mmol, 60.9 ml, 1.5 eq) was added dropwise via a flame-dried addition funnel. Following addition, the reaction was allowed to stir at 0 °C for 2 hours. The reaction was then partitioned between DCM and water. The organic phase was dried over sodium sulfate, filtered and evaporated to dryness yielding an oil

that was immediately reacted further. The crude triflate was azeotroped with toluene 3 times and purged with argon, then dissolved in anhydrous DMF (400 ml). Sodium azide (60.6 g, 933 mmol, 4.3 eq) was added and the reaction was stirred at room temperature for 2 days. The reaction was concentrated and partitioned between EtOAc and water. The organic phase was washed 5 times with water, dried over sodium sulfate, filtered, and evaporated to dryness. The crude oil was subjected to column chromatography with a gradient of 0-30% EtOAc in hexanes to yield a yellowish oil (35.5g, 53% over 2 steps). ^{1}H NMR (500 MHz, chloroform-d) δ 7.38 – 7.27 (m, 5H), 5.82 (d, J = 3.6 Hz, 1H), 4.71 (dd, J = 4.2, 3.6 Hz, 1H), 4.60 (ABq, J = 12.1 Hz, 2H), 4.21 – 4.17 (m, 1H), 3.80 (dd, J = 11.3, 2.5 Hz, 1H), 3.64 (dd, J = 11.4, 3.6 Hz, 1H). 3.62 (dd, J = 9.6, 4.7 Hz, 1H). 1.57 (s, 3H), 1.36 (s, 3H). ^{13}C NMR (126 MHz, chloroform-d) δ 137.59 (s), 128.48 (s), 127.86 (s), 127.79 (s), 113.06 (s), 104.13 (s), 79.77 (d, J = 20.4 Hz), 77.26 (s), 73.70 (s), 67.53 (s), 60.43 (s), 26.41 (s), 26.40 (s). ESI-HRMS calculated for $C_{15}H_{19}N_3O_4Na$ [M+Na]$^+$ 328.1268, found 328.1271.

(2S,3R,4R)-3-azido-4-(benzyloxy)-2-((benzyloxy)methyl)-5-methoxytetrahydrofuran (54)

Compound **53** (35.5 g, 232 mmol) was dissolved in MeOH (305 ml) and acetyl chloride (4.51 g, 57.5 mmol, 4.1 ml) was added dropwise. The solution was allowed to stir at room temperature overnight, then evaporated to dryness and partitioned between DCM and saturated NaHCO$_3$. The aqueous phase was extracted with EtOAc twice and the organic phases were combined, dried over sodium sulfate, and evaporated to dryness leaving an oily crude that was dried on high vacuum overnight. After purging with argon, the riboside was dissolved in anhydrous DMF (594 ml) and tetrabutylammonium iodide (429 mg, 1.16 mmol, 0.01 eq). The solution was brought to 0 °C with an ice bath and sodium hydride (5.58 g, 232 mmol, 2 eq) was

added slowly under an argon stream. The solution was allowed to stir for 15 minutes at 0 °C. Benzyl bromide (39.8 g, 232 mmol, 2 eq) was added dropwise via an addition funnel. The reaction was left to warm to room temperature overnight. The reaction was quenched with 45 ml saturated NH₄Cl solution and concentrated. The remaining crude was partitioned between EtOAc and water. The aqueous phase was extracted with EtOAc 3 times, and the organic phases were combined, dried with sodium sulfate, and concentrated. Column chromatography of the resulting residue with a gradient of 0 to 30% EtOAc in hexanes yielded an oil with two separable isomers that were combined for the next reaction (35.5 g, 83% over 2 steps). ^{1}H NMR (500 MHz, chloroform-*d*) δ 7.46 – 7.30 (m, 10H), 4.92 (s, 1H), 4.73 (s, 2H), 4.63 (d, J = 1.4 Hz, 2H), 4.35 (dt, J = 7.5, 4.9 Hz, 1H), 4.00 (d, J = 4.8 Hz, 1H), 3.84 (dd, J = 7.5, 4.9 Hz, 1H), 3.68 – 3.58 (m, 2H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 137.91, 137.25, 128.52, 128.41, 127.98, 127.71, 127.63, 106.09, 82.50, 79.74, 77.49, 73.41, 72.79, 70.95, 61.45, 55.14. ESI-HRMS calculated for $C_{20}H_{23}N_3O_4Na$ [M+Na]$^+$ 392.1581, found 328.1578. 1H NMR (300 MHz, CDCl3) δ 7.45 – 7.29 (m, 10H), 4.88 (d, J = 4.2 Hz, 1H), 4.71 (ABq, J = 12.2 Hz, 2H), 4.57 (ABq, J = 12.1 Hz, 2H), 4.14 (dd, J = 7.3 Hz, 3.5 Hz, 1Hz), 3.94 (ABq, J = 7.5, 3.9 Hz, 2H), 3.60 (d, J = 3.6 Hz, 2H), 3.46 (s, 3H). ^{13}C NMR (126 MHz, chloroform-*d*) δ 137.64, 137.17, 128.52, 128.48, 128.12, 128.09, 127.85, 127.65, 101.81, 80.73, 78.33, 77.49, 73.61, 72.94, 69.83, 59.62, 55.31. ESI-HRMS calculated for $C_{20}H_{23}N_3O_4Na$ [M+Na]$^+$ 392.1581, found 392.1581.

(3R,4R,5S)-4-azido-3-(benzyloxy)-5-((benzyloxy)methyl)dihydrofuran-2(3H)-one (55)

Compound **54** (13.6 g, 37 mmol) was dissolved in a mixture of dioxane (170 ml) and water (108 ml). 37% HCl (54 ml) was added to the solution slowly over 20 minutes and the solution was stirred at 60 °C overnight. The reaction was then cooled in an ice bath and

quenched with saturated NaHCO$_3$. The mixture was poured into a separatory funnel and extracted 3 times with EtOAc. The organic phases were combined, dried with sodium sulfate, filtered, and concentrated to yield an oil. The crude hemi-acetal was then coevaporated once with DCM and 3 times with toluene, then purged with argon and dissolved in anhydrous DCM (297 ml). Pyridinium chlorochromate (12 g, 55.5 mmol, 1.5 eq) was added and the reaction was stirred at room temperature overnight. Upon reaction completion the solution was filtered over a celite cake and evaporated to dryness. The crude residue was subjected to column chromatography with a gradient of 0-40% EtOAc in hexanes to yield the lactone as an off-yellow oil (9.7 g, 74% over 2 steps). ^{1}H NMR (500 MHz, chloroform-*d*) δ 7.40 – 7.13 (m, 10H), 5.02 (d, J = 12.0 Hz, 1H), 4.78 (d, J = 12.0 Hz, 1H), 4.62 (d, J = 6.2 Hz, 1H), 4.47 (dd, J = 50.1, 11.8 Hz, 2H), 4.35 (dd, J = 3.9, 2.2 Hz, 1H), 4.14 (dd, J = 6.2, 1.5 Hz, 1H), 3.68 (dd, J = 11.1, 2.6 Hz, 1H), 3.61 (dd, J = 11.1, 2.3 Hz, 1H). ^{13}C NMR (126 MHz, chloroform-*d*) δ 172.95 (s), 136.77 (s), 136.43 (s), 128.65 (s), 128.61 (s), 128.34 (s), 128.19 (s), 128.18 (s), 127.70 (s), 80.65 (d, J = 33.8 Hz), 74.16 (s), 73.78 (s), 73.29 (s), 68.90 (s), 60.21 (s). ESI-HRMS calculated for C$_{19}$H$_{19}$N$_3$O$_4$Na [M+Na]$^+$ 376.1268, found 376.1265.

3.11.4 Synthesis of fluorescent antibiotics
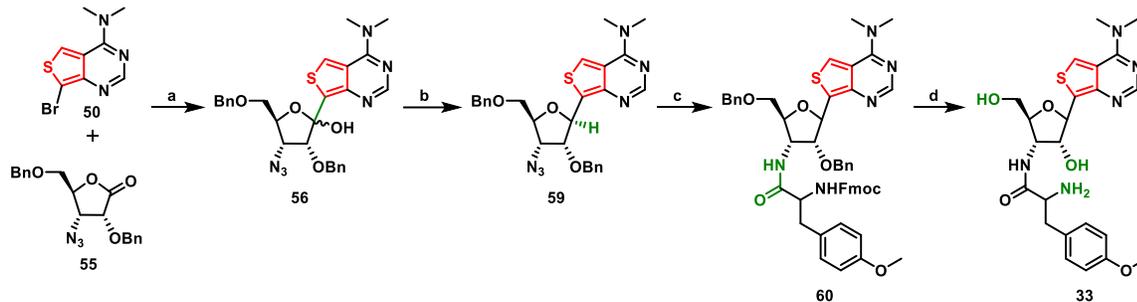
Scheme 3.12. Reiteration of scheme 3.6: Chemical synthesis of [Th]puromycin **33**. (a) t-BuLi, THF, -78°C, 40%. (b) BF$_3$·OEt$_2$, Et$_3$SiH, ACN;DCM 3:1, -40°C to RT, 36%. (c) Fmoc-O-methyl-L-tyrosine, DCC, HOBt, (n-Bu)$_3$P, THF, 58%. (d) i. Piperidine, DMF, RT; ii. BCl$_3$, DCM, -78°C to 0°C, 60%.

7-((2R,3R,4R,5S)-4-azido-3-(benzyloxy)-5-((benzyloxy)methyl)tetrahydrofuran-2-yl)-N,N-dimethylthieno[3,4-*d*]pyrimidin-4-amine (59)

To a 25 ml round-bottom flask that was flame-dried and purged with argon was added nucleobase **50** (48 mg, 0.19 mmol) and anhydrous THF (3.5 ml). The solution was cooled to -78 °C and t-BuLi (1.5 M in pentane, 375 µl, 3 eq) was added dropwise. The solution was allowed to stir at -78 °C for 1 hr before adding a cooled solution of azidolactone **55** (286 mg, 0.81 mmol, 4.3 eq) in anhydrous THF (0.5 ml). The reaction was allowed to stir at -78 °C for 1 hr, then quenched with water and EtOAc. The solution was evaporated to dryness. The crude was subjected to column chromatography with a gradient of 0-75% EtOAc in hexanes to yield a slightly crude yellowish solid (40 mg) immediately subjected to the next reaction. The solid was placed in a 10 ml round bottom flask and coevaporated with acetonitrile 3 times. Anhydrous ACN:DCM 3:1 (1 ml) was added along with triethylsilane (53 mg, 0.45 mmol, 6 eq). The solution was brought to -40 °C in a dry ice/acetonitrile bath and BF$_3$·OEt$_2$ (32 mg, 0.23 mmol, 3 eq) was added very slowly dropwise over the course of 20-30 minutes. It is imperative to add the lewis acid slowly as reduction of the nucleobase can occur if the addition is too quick. After

stirring for 1 hour at -40 °C following lewis acid addition, the solution was allowed to stir at room temperature for 40 minutes. The solution was then partitioned between EtOAc and 1M NaOH, then extracted. The aqueous phase was washed 3 times with EtOAc and the organic phases were combined, dried with sodium sulfate, filtered, and evaporated to dryness. The crude riboside was subjected to column chromatography using 0-70% EtOAc in hexanes to yield a brownish solid (13 mg, 14% over 2 steps). ^{1}H NMR (500 MHz, chloroform-*d*) δ 8.29 (s, 1H), 7.89 (s, 1H), 7.38 – 7.26 (m, 10H), 5.93 (d, J = 3.8 Hz, 1H), 4.82 (Abq, J = 12.2 Hz, 2H), 4.65 (Abq, J = 12.2 Hz, 2H), 4.39 (dd, J = 5.4, 3.9 Hz, 1H), 4.34 – 4.30 (m, J = 7.3, 4.0, 3.2 Hz, 1H), 3.92 (dd, J = 7.4, 5.4 Hz, 1H), 3.82 (dd, J = 10.9, 3.1 Hz, 1H), 3.73 (dd, J = 10.9, 4.2 Hz, 1H), 3.48 (s, 6H).^{13}C NMR (126 MHz, chloroform-*d*) δ 157.94 (s), 153.20 (s), 148.42 (s), 138.08 (s), 137.60 (s), 130.46 (s), 128.40 (s), 128.26 (s), 128.06 (s), 127.78 – 127.54 (m), 121.38 (s), 120.76 (s), 83.87 (s), 80.04 (s), 78.74 (s), 73.55 (s), 72.11 (s), 69.67 (s), 60.88 (s), 40.48 (s). ESI-HRMS calculated for $C_{27}H_{29}N_6O_3S$ [M+H]$^+$ 517.2016, found 517.2012.

(9H-fluoren-9-yl)methyl((S)-1-(((2S,3R,4R,5R)-4-(benzyloxy)-2-((benzyloxy)methyl)-5-(4-(dimethylamino)thieno[3,4-*d*]pyrimidin-7-yl)tetrahydrofuran-3-yl)amino)-3-(4-methoxyphenyl)-1-oxopropan-2-yl)carbamate (60)

To a 10 ml round bottom flask that was flame-dried and purged with argon was added protected nucleoside **59** (54 mg, 0.10 mmol) and THF (1.1 ml) and water (100 µl). A solution of trimethylphosine in THF (1 M, 312 µl, 3 eq) was added dropwise and the solution was stirred at room temperature until disappearance of starting material was observed 30 minutes later. The solution was then evaporated to dryness and the crude aminonucleoside was coevaporated 3 times with acetonitrile and left to dry on high vacuum for 3 hours. Subsequently, in a separate

flame-dried and argon-purged 10 ml round bottom flask was added Fmoc-O-methyl-L-tyrosine (57 mg, 0.14 mmol, 1.3 eq) and anhydrous THF (4 ml). The solution was brought to 0 °C in an ice bath and HOBt (26 mg, 0.19 mmol, 1.7 eq) was added. Following 15 minutes of stirring on ice, EDC·HCl (37 mg, 0.19 mmol, 1.7 eq) was added. The newly formed suspension was stirred at 0 °C for 30 minutes before adding a cooled solution of the crude aminonucleoside in THF (1 m) dropwise. The reaction was allowed to warm to room temperature overnight. The solution was then rotorvapped to dryness, then dissolved in DCM and extracted with water. The aqueous phase was washed with DCM 3 times and the organic phases were combined, dried with sodium sulfate, filtered, and evaporated to dryness. The crude protected antibiotic was subjected to column chromatography with a gradient of 0-5% MeOH in DCM to yield the desired product as a slightly impure cream solid (54 mg, 58%). The product was crudely purified and was immediately subjected to deprotection.

(S)-2-amino-N-((2S,3S,4R,5R)-5-(4-(dimethylamino)thieno[3,4-*d*]pyrimidin-7-yl)-4-hydroxy-2-(hydroxymethyl)tetrahydrofuran-3-yl)-3-(4-methoxyphenyl) propanamide (33)

To a flame-dried 25 ml round-bottom flask under argon was added protected Thpuromycin **60** (38 mg, 0.04 mmol) and a solution of 20% piperidine in DMF (1 ml). The reaction was allowed to stir at room temperature for 15 minutes, then evaporated to dryness. The crude was coevaporated 3 times with heptane and left to dry on high vaccuum overnight, then subsequently flushed with argon and dissolved in anhydrous DCM (2.5 ml) The solution was brought to -78 °C with a dry-ice acetone bath and BCl$_3$ (1M in DCM, 0.4 ml, 0.4 mmol, 10 eq) was added dropwise. The solution was allowed to stir at -78 °C for 1 hour, then at room temperature for another hour. Upon proof of reaction completion with ESI-MS, the reaction was

quenched with methanol and evaporated to dryness. The crude antibiotic was purified using

reverse-phase HPLC (0-40% ACN with 0.1% TFA in water with 0.1% TFA over 22 minutes) to

yield the desired product as a white solid (10 mg, 48% over 2 steps). ^{1}H NMR (500 MHz,) δ

8.61 (s, 1H), 8.08 (s, 1H), 7.13 – 7.07 (m, 2H), 6.92 – 6.83 (m, 2H), 5.03 (d, J = 5.5 Hz, 1H),

4.26 (dd, J = 6.5, 5.5 Hz, 1H), 4.13 – 4.09 (m, 2H), 3.67 (d, J = 1.4 Hz, 6H), 3.63 (ddd, J = 6.8,

4.1, 2.6 Hz, 1H), 3.53 (dd, J = 12.4, 2.6 Hz, 1H), 3.42 (s, 3H), 3.37 (dd, J = 12.5, 4.2 Hz, 1H),

3.17 – 2.87 (m, 2H). ^{13}C NMR (126 MHz, D$_2$O) δ 169.13 (s), 163.03 (s), 162.75 (s), 158.29 (s),

151.09 (s), 147.25 (s), 130.63 (s), 130.47 (s), 126.27 (s), 118.28 (s), 117.29 (s), 114.97 (s),

114.28 (s), 82.14 (s), 79.11 (s), 74.96 (s), 60.76 (s), 55.13 (s), 54.36 (s), 51.98 (s), 42.46 (s),

41.44 (s), 35.98 (s). ESI-HRMS calculated for C$_{23}$H$_{30}$N$_5$O$_5$S [M+H]$^+$ 488.1962, found 488.1955.
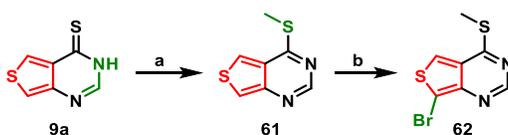
Scheme 3.13. Chemical synthesis of brominated thiomethyl purine analog **62**. (a) MeI, K$_2$CO$_3$, DMF, RT, 1 hr, 78%. (b) DBDMH, DMF, 0 °C, 2.5 h, 45%.

4-(methylthio)thieno[3,4-*d*]pyrimidine (61)

To a 2 L flask flame-dried and purged with argon was added thionated nucleobase **9a**

(14.5 g, 168 mmol, 1 eq), anhydrous DMF (375 ml) and K$_2$CO$_3$ (50 g, 362 mmol, 4.5 eq). MeI

(34 g, 240 mmol, 15 ml, 3 eq) was added dropwise and the solution was allowed to stir for 1

hour at room temperature before adding another equivalent of MeI (11.3 g, 80 mmol, 5 ml, 1 eq)

dropwise. The reaction was stirred for another 30 minutes at room temperature before deemed

complete by TLC. The solution was evaporated to dryness and the crude was partitioned between

DCM and water. The organic phase was washed 3 times with water, dried with sodium sulfate,

filtered, and evaporated to dryness. The remaining crude was subjected to column

143

chromatography with a gradient of 0-3% MeOH in DCM to yield a light-brownish solid (12 g, 78%). ^{1}H NMR (500 MHz, chloroform-*d*) δ 8.64 (s, 1H), 8.01 (d, *J* = 3.3 Hz, 1H), 7.83 (d, *J* = 3.3 Hz, 1H), 2.70 (s, 3H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 168.23, 151.87, 148.20, 127.48, 120.15, 117.74, 12.08. ESI-HRMS calculated for $C_7H_7N_2S_2$ [M+H]$^+$ 183.0045, found 183.0047.

7-bromo-4-(methylthio)thieno[3,4-*d*]pyrimidine (62)

To a 1 L flame-dried round-bottom flask purged with argon was added nucleobase **61** (12 g, 66 mmol) and anhydrous DMF (92 ml). The solution was cooled to 0 °C in an ice bath and 1,3-Dibromo-5,5-Dimethylhydantoin (2.35 g, 8.2 mmol, 0.12 eq) was added slowly. The solution was stirred at 0 °C for one hour before adding another eighth of an equivalent of 1,3-Dibromo-5,5-Dimethylhydantoin (2.35 g, 8.2 mmol, 0.12 eq). After stirring for 15 minutes at 0 °C more 1,3-Dibromo-5,5-Dimethylhydantoin (2.35 g, 8.2 mmol, 0.12 eq) was added, then following another 15 minutes at 0 °C a final eighth of an equivalent (total 0.5 eq) was added. The reaction was stirred on ice for another hour before deemed complete by TLC. The reaction was quenched with saturated sodium thiosulfate and partitioned between EtOAc and water. The aqueous phase was washed 4 times with EtOAc and the organic phases were combined, dried with sodium sulfate, filtered, and evaporated to dryness. The crude halogenated nucleobase was subjected to column chromatography using a gradient of 0-20% EtOAc in hexanes to yield a yellowish solid (7.8 g, 45%). ^{1}H NMR (500 MHz, chloroform-*d*) δ 8.72 (s, 1H), 8.06 (d, J = 3.7 Hz, 1H), 2.72 (s, 3H). ^{13}C NMR (126 MHz, chloroform-*d*) δ 169.21 (s), 152.50 (s), 145.38 (s), 127.40 (s), 120.41 (s), 106.56 (s), 12.22 (d, J = 3.3 Hz). ESI-HRMS calculated for $C_7H_6N_2BrS_2$ [M+H]$^+$ 260.9150, found 260.9149.
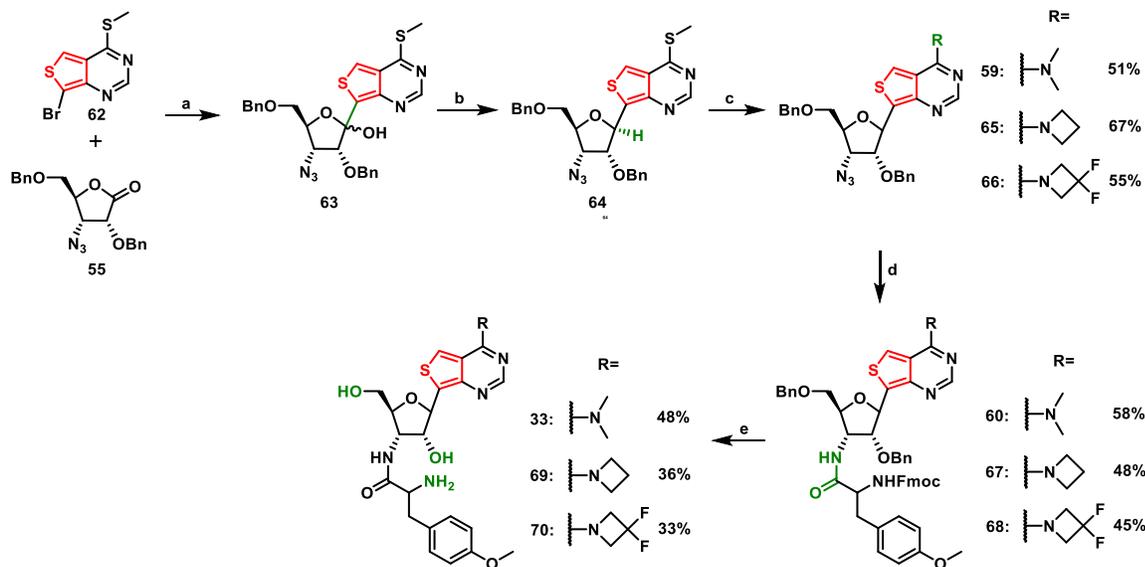
Scheme 3.14. Reiterated from scheme 3.8: Chemical synthesis of azetine- and difluoroazetidine-modified puromycin analogues **69** and **70**. (a) n-BuLi, THF, -78°C, 27%. (b) BF$_3$·OEt$_2$, Et$_3$SiH, DCM, -78°C to RT, 55%. (c) RH$_2$Cl, DBU, DMSO, RT. (d) Fmoc-O-methyl-L-tyrosine, EDC, HOBt, (Me)$_3$P, THF. (e) i. Piperidine, DMF, RT; ii. The same synthetic pathway was found to produce Thpuromycin **33** as well.

(3R,4R,5S)-4-azido-3-(benzyloxy)-5-((benzyloxy)methyl)-2-(4-(methylthio)thieno[3,4-*d*]pyrimidin-7-yl)tetrahydrofuran-2-ol (63)

To a 100 ml flame-dried round bottom flask purged with argon was added nucleobase **62** (1 g, 3.8 mmol), anhydrous THF (10 ml), azidolactone **55** (1.6 g, 4.6 mmol, 1.2 eq), and activated 3 Å molecular sieves. The solution was allowed to dry overnight and was then brought to -78° C in a dry-ice acetone bath. *n*-BuLi (2.4 M in hexanes, 1.9 ml, 4.6 mmol, 1.2 eq) was added slowly dropwise. The reaction was allowed to stir for 2 hours at -78° C, then quenched with 1 ml saturated ammonium chloride. The solution was partitioned between EtOAc and water. The organic phase was dried with sodium sulfate, filtered, and evaporated to dryness. The crude hemi-ketal was subjected to column chromatography using a gradient of 0-30% EtOAc in

hexanes to yield a yellowish solid (550 mg, 27%). The product was crudely purified and used immediately in the next reaction.

7-((2R,3R,4R,5S)-4-azido-3-(benzyloxy)-5-((benzyloxy)methyl)tetrahydrofuran-2-yl)-4-(methylthio)thieno[3,4-*d*]pyrimidine (64)

To a 25 ml round bottom flask that was flame-dried and purged with argon was added hemi-ketal **63** (530 mg, 1.0 mmol) and anhydrous DCM (12 ml). Triethylsilane (690 mg, 5.9 mmol, 950 μl, 6 eq) was added and the solution was allowed to cool to -78 °C in a dry ice acetone bath. $BF_3 \cdot Et_2O$ (420 mg, 3.0 mmol, 370 μl, 3 eq) was added dropwise. The reaction was initially allowed to stir at -78 °C for 1 hour before stirring at room temperature for 2 hours, then quenched with triethylamine (506 mg, 5 mmol, 700 μl, 5 eq). The solution was partitioned between EtOAc and water, then extracted. The aqueous phase was washed with EtOAc and the organic phases were combined, dried with sodium sulfate, filtered, and evaporated to dryness. The crude glycoside was subjected to column chromatography using a gradient of 0-15% EtOAc in hexanes to yield the desired product as a yellowish solid (284 mg, 55%). ^{1}H NMR (500 MHz, chloroform-*d*) δ 8.62 (s, 1H), 7.87 (s, 1H), 7.38 – 7.20 (m, 10H), 5.91 (d, J = 4.3 Hz, 1H), 4.76 (ABq, J= 12.2 Hz, 2H), 4.64 (ABq, J = 12.0 Hz, 2H), 4.44 (t, J = 4.9 Hz, 1H), 4.32 (dt, J = 6.9, 3.5 Hz, 1H), 3.97 (dd, J = 6.8, 5.3 Hz, 1H), 3.81 (dd, J = 10.9, 3.1 Hz, 1H), δ 2.70 (s, 1H). ^{13}C NMR (126 MHz, Chloroform-d) δ 168.14, 151.15, 144.35, 137.94, 137.32, 134.76, 128.43, 128.27, 128.21, 128.01, 127.82, 127.73, 127.65, 118.58, 84.15, 80.46, 78.20, 73.59, 72.40, 69.64, 61.01, 11.97. ESI-HRMS calculated for $C_{26}H_{26}N_5O_3S_2$ [M+H]$^+$ 520.1472, found 520.1474.

7-((2R,3R,4R,5S)-4-azido-3-(benzyloxy)-5-((benzyloxy)methyl)tetrahydrofuran-2-yl)-N,N-dimethylthieno[3,4-*d*]pyrimidin-4-amine (59)

To a 10 ml round-bottom flask that was flame-dried and purged with argon was added protected nucleoside **64** (45 mg, 0.1 mmol) and anhydrous DMSO (1 ml). Dimethylammonium hydorchloride (184 mg, 2.3 mmol, 26 eq) was added foolowed by DBU (490 mg, 3.6 mmol, 450 µl, 37 eq). The reaction was allowed to stir at room temperature overnight. Upon proof of completion the solution was partitioned betweeen EtOAc and water and extracted. The organic phase was washed with water 6 times, dried with sodium sulfate, filtered, and evaporated to dryness. The crude was subjected to column chromatography using a gradient of 0-2% MeOH in DCM to yield the desired product as a yellowish film (23 mg, 51%).

4-(azetidin-1-yl)-7-((2R,3R,4R,5S)-4-azido-3-(benzyloxy)-5-((benzyloxy)methyl) tetrahydrofuran-2-yl)thieno[3,4-*d*]pyrimidine (65)

To a flame-dried 10 ml round bottom flask under argon was added protected nucleoside **64** (89 mg, 0.17 mmol) and anhydrous DMSO (3.4 ml). Azetidine hydrochloride (458 mmol, 5 mmol, 29 eq) was added along with DBU (1.08 g, 7.13 mmol, 1 ml, 42 eq) and the reaction was allowed to stir overnight. The reaction was partitioned between EtOAc and water and extracted. The organic phase was washed with water 6 times, dried over sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude nucleoside using a gradient of 0-2% MeOH in DCM yielded the desired product as a brownish solid (63 mg, 70%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.26 (s, 1H), 7.66 (s, 1H), 5.87 (d, *J* = 4.1 Hz, 1H), 4.85 – 4.74 (m, 2H), 4.69 – 4.58 (m, 2H), 4.40 (dd, *J* = 5.4, 4.1 Hz, 1H), 4.30 (ddd, *J* = 7.2, 4.1, 3.1 Hz, 1H), 3.94 (dd, *J* = 7.2, 5.4 Hz, 1H), 3.85 – 3.68 (m, 2H), 2.56 (tt, *J* = 8.4, 6.9 Hz, 2H). ^{13}C NMR (126 MHz,

Chloroform-*d*) δ 156.20, 153.21, 146.48, 138.05, 137.52, 131.16, 128.44, 128.30, 128.08, 127.75, 127.69, 127.65, 119.68, 83.89, 80.23, 78.47, 73.57, 73.56, 72.25, 72.23, 69.67, 60.93, 53.09, 50.97, 29.75, 17.16. ESI-HRMS calculated for $C_{28}H_{29}N_6O_3S$ [M+H]$^+$ 529.2016, found 529.2016.

7-((2R,3R,4R,5S)-4-azido-3-(benzyloxy)-5-((benzyloxy)methyl)tetrahydrofuran-2-yl)-4-(3,3-difluoroazetidin-1-yl)thieno[3,4-*d*]pyrimidine (66)

To a flame-dried 25 ml round bottom flask under argon was added protected nucleoside **64** (92 mg, 0.18 mmol) and anhydrous DMSO (1.7 ml). Difluoroazetidine hydorchloride (600 mg, 4.6 mmol, 26 eq) and DBU (1 g, 6.6 mmol, 0.9 ml, 37 eq) were added sequentially. The reaction was allowed to stir overnight at room temperature. The reaction solution was then partitioned between EtOAc and water. The organic phase was washed with water 6 times, dried over sodium sulfate, filtered, and evaporated to dryness. The crude nucleoside was subjected to column chromatography using a gradient of 0-40% EtOAc in hexanes to yield the desired product as a yellowish solid (57 mg, 57%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.34 (s, 1H), 7.59 (s, 1H), 7.36 – 7.22 (m, 13H), 5.90 (d, *J* = 3.8 Hz, 1H), 4.80 (dd, *J* = 13.5, 8.4 Hz, 7H), 4.70 – 4.60 (m, 2H), 4.42 (t, *J* = 4.7 Hz, 1H), 4.32 (dt, *J* = 7.2, 3.4 Hz, 1H), 3.94 (t, *J* = 6.4 Hz, 1H), 3.77 (ddd, *J* = 52.9, 10.9, 3.4 Hz, 2H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 156.64 (d, *J* = 4.1 Hz), 153.07, 146.95, 138.00, 137.45, 133.06, 128.42, 128.29, 128.05, 127.76, 127.69, 127.62, 120.24, 118.39, 117.98, 115.80, 113.62, 83.98, 80.24, 78.52, 73.56, 72.27, 69.56, 63.70 – 62.61 (m), 60.84. ESI-HRMS calculated for $C_{28}H_{27}F_2O_3N_6S$ [M+H]$^+$ 565.1828, found 565.1822.

(9H-fluoren-9-yl)methyl((S)-1-(((2S,3R,4R,5R)-5-(4-(azetidin-1-yl)thieno[3,4-*d*]pyrimidin-7-yl)-4-(benzyloxy)-2-((benzyloxy)methyl)tetrahydrofuran-3-yl)amino)-3-(4-methoxyphenyl)-1-oxopropan-2-yl)carbamate (67)

To a flame-dried 10 ml round-bottom flask was added azetidine-modified nucleoside **65** (55 mg, 0.10 mmol) and THF (1.1 ml). Trimethylphosphine (1 M in THF, 0.34 ml, 3.3 eq) was added along with water (0.1 ml). The reaction was stirred for 30 minutes at room temperature until all starting material dissappeared by TLC. The solution was then evaporated to dryness, coevaporated 3 times with acetonitrile, and reconstituted in anhydrous THF (1 ml). Simultaneously, in a separate flame-dried round bottom flask under argon was added Fmoc-O-methyl-L-tyrosine (62 mg, 0.15 mmol, 1.4 eq) and anhydrous THF (4 ml), The solution was cooled to in an ice bath to 0 °C and HOBt (26 mg, 0.19 mmol, 1.9 eq) was added. After stirring under ice for 15 minutes, EDC·HCl (37 mg, 0.19 mmol, 1.9 eq) was added. The reaction was allowed to stir at 0 °C for 30 minutes. A chilled solution of the 3'-amino-3'-deoxy nucleoside in THF prepared previously was added dropwise to the newly formed activated ester. The reaction was allowed to warm to room temperature overnight. Upon proof of completion the reaction was evaporated to dryness and the crude was partitioned between DCM and water. The aqueous phase was washed 3 times with DCM and the organic phases were combined, dried over sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude protected antibiotic using a gradient of 0-5% MeOH in DCM yielded the desired product as cream solid (45 mg, 48%). The product was crudely purified and subjected to deprotection immediately.

(9H-fluoren-9-yl)methyl((S)-1-(((2S,3R,4R,5R)-4-(benzyloxy)-2-((benzyloxy)methyl)-5-(4-(3,3-difluoroazetidin-1-yl)thieno[3,4-*d*]pyrimidin-7-yl)tetrahydrofuran-3-yl)amino)-3-(4-methoxyphenyl)-1-oxopropan-2-yl)carbamate (68)

To a flame-dried 10 ml round-bottom flask was added difluoroazetidine-modified nucleoside **66** (57 mg, 0.10 mmol) and THF (1.1 ml). Trimethylphosphine (1 M in THF, 0.34 ml, 3.3 eq) was added along with water (0.1 ml). The reaction was stirred for 30 minutes at room temperature until all starting material dissappeared by TLC. The solution was then evaporated to dryness slowly while keeping the solution cool, coevaporated 3 times with cold acetonitrile, and reconstituted in chilled anhydrous THF (1 ml). Simultaneously, in a separate flame-dried round bottom flask under argon was added Fmoc-O-methyl-L-tyrosine (62 mg, 0.15 mmol, 1.4 eq) and anhydrous THF (4 ml), The solution was cooled to in an ice bath to 0 °C and HOBt (26 mg, 0.19 mmol, 1.9 eq) was added. After stirring under ice for 15 minutes, EDC·HCl (37 mg, 0.19 mmol, 1.9 eq) was added. The reaction was allowed to stir at 0 °C for 30 minutes. A chilled solution of the 3'-amino-3'-deoxy nucleoside in THF prepared previously was added dropwise to the newly formed activated ester. The reaction was allowed to warm to room temperature overnight. Upon proof of completion the reaction was evaporated to dryness and the crude was partitioned between DCM and water. The aqueous phase was washed 3 times with DCM and the organic phases were combined, dried over sodium sulfate, filtered, and evaporated to dryness. Column chromatography of the crude protected antibiotic using a gradient of 0-5% MeOH in DCM yielded the desired product as cream solid (43 mg, 45%). The product was crudely purified and subjected to deprotection immediately.

(S)-2-amino-N-((2S,3S,4R,5R)-5-(4-(azetidin-1-yl)thieno[3,4-*d*]pyrimidin-7-yl)-4-hydroxy-2-(hydroxymethyl)tetrahydrofuran-3-yl)-3-(4-ethoxyphenyl)propanamide (69)

To a flame-dried 25 ml round-bottom flask under argon was added protected antibiotic **67** (35 mg, 0.04 mmol) and a solution of 20% piperidine in DMF (1 ml). The reaction was allowed to stir at room temperature for 15 minutes, then evaporated to dryness. The crude was coevaporated 3 times with heptane and left to dry on high vaccuum overnight, then subsequently flushed with argon and dissolved in anhydrous DCM (2.5 ml) The solution was brought to -78 °C with a dry-ice acetone bath and BCl_3 (1M in DCM, 0.4 ml, 0.4 mmol, 10 eq) was added dropwise. The solution was allowed to stir at -78 °C for 1 hour, then at room temperature for another hour. Upon proof of reaction completion with ESI-MS, the reaction was quenched with methanol and evaporated to dryness. The crude antibiotic was purified using reverse-phase HPLC (0-40% ACN with 0.1% TFA in water with 0.1% TFA over 22 minutes) to yield the desired product as a white solid (7 mg, 36% over 2 steps). ^{1}H NMR (500 MHz, D$_2$O) δ 8.28 (s, 1H), 8.01 (s, 1H), 7.12 – 7.05 (m, 2H), 6.89 – 6.80 (m, 2H), 5.03 (d, J = 5.4 Hz, 1H), 4.82 – 4.74 (m, 2H), 4.46 – 4.37 (m, 2H), 4.24 (dd, J = 6.4, 5.4 Hz, 1H), 4.15 – 4.04 (m, 2H), 3.65 (s, 3H), 3.62 (ddd, J = 6.8, 4.2, 2.5 Hz, 1H), 3.54 – 3.31 (m, 2H), 3.18 – 2.84 (m, 2H), 2.52 (p, J = 7.7 Hz, 2H). ^{13}C NMR (126 MHz, Deuterium Oxide) δ 169.11, 162.84 (q, J = 35.6 Hz), 158.28, 153.44, 147.47, 133.42, 130.45, 128.50, 126.25, 125.97, 119.61, 117.31, 114.98, 114.25, 112.65, 82.08, 79.18 (d, J = 2.5 Hz), 75.14, 60.75, 55.28 – 55.07 (m), 54.84, 54.36, 52.99, 51.94, 35.97, 16.14. ESI-HRMS calculated for $C_{24}H_{30}N_5O_5S$ [M+H]$^+$ 500.1962, found 500.1960.

(S)-2-amino-N-((2S,3S,4R,5R)-5-(4-(3,3-difluoroazetidin-1-yl)thieno[3,4-*d*] pyrimidin-7-yl)-4-hydroxy-2-(hydroxymethyl)tetrahydrofuran-3-yl)-3-(4-methoxyphenyl)propanamide (70)

To a flame-dried 25 ml round-bottom flask under argon was added protected antibiotic **68** (43 mg, 0.05 mmol) and a solution of 20% piperidine in DMF (1 ml). The reaction was allowed to stir at room temperature for 15 minutes, then evaporated to dryness. The crude was coevaporated 3 times with heptane and left to dry on high vaccuum overnight, then subsequently flushed with argon and dissolved in anhydrous DCM (2.5 ml) The solution was brought to -78 °C with a dry-ice acetone bath and BCl$_3$ (1M in DCM, 0.5 ml, 0.5 mmol, 10 eq) was added dropwise. The solution was allowed to stir at -78 °C for 1 hour, then at room temperature for another hour. Upon proof of reaction completion with ESI-MS, the reaction was quenched with methanol and evaporated to dryness. The crude antibiotic was purified using reverse-phase HPLC (0-40% ACN with 0.1% TFA in water with 0.1% TFA over 22 minutes) to yield the desired product as a white solid (8 mg, 33% over 2 steps). ^{1}H NMR (500 MHz, D$_2$O) δ 8.36 (s, 1H), 8.20 (s, 1H), 7.15 – 7.08 (m, 2H), 6.91 – 6.85 (m, 2H), 5.26 (t, *J* = 11.1 Hz, 2H), 5.07 (d, *J* = 5.4 Hz, 1H), 4.88 (t, *J* = 11.4 Hz, 2H), 4.31 – 4.22 (m, 1H), 4.11 (ddd, *J* = 8.2, 6.2, 1.4 Hz, 2H), 3.67 (s, 3H), 3.64 (ddd, *J* = 6.6, 4.1, 2.4 Hz, 1H), 3.54 (dd, *J* = 12.5, 2.6 Hz, 1H), 3.36 (dd, *J* = 12.5, 4.1 Hz, 1H), 3.19 – 2.85 (m, 2H). ^{13}C NMR (126 MHz, Deuterium Oxide) δ 176.12, 169.14, 162.91 (q, *J* = 35.6 Hz), 158.30, 155.83, 149.29, 132.88, 130.47, 128.27, 126.27, 125.86, 119.62, 117.65, 117.30, 116.78, 114.98, 114.62, 114.28, 112.66, 112.46, 104.99, 82.14, 79.23 (d, *J* = 2.2 Hz), 75.14, 64.49 (dt, *J* = 97.5, 30.0 Hz), 60.67, 55.69 – 55.07 (m), 54.36, 51.92, 38.83, 35.98, 35.38. ESI-HRMS calculated for C$_{24}$H$_{28}$F$_2$N$_2$O$_5$S [M+H]$^+$ 536.1774, found 536.1774.

3.11.5 Structural analysis

3.11.5.1 General experimental summary

The single crystal X-ray diffraction studies were carried out on a Bruker Apex II CCD diffractometer equipped with Cu K$_\square$ radiation (λ = 1.54178). A 0.200 x 0.180 x 0.150 mm colorless crystal was mounted on a Cryoloop with Paratone oil. Data were collected in a nitrogen gas stream at 100(2) K using ϕ and ϖ scans. Crystal-to-detector distance was 40 mm using exposure time 3,12 and 20s s with a scan width of 1.25°. Data collection was 97.1% complete to 66.567° in θ.

A total of 14870 reflections were collected covering the indices, -25<=h<=20, -9<=k<9, -23<=l<=24. 5308 reflections were found to be symmetry independent, with a R$_{int}$ of 0.069. Indexing and unit cell refinement indicated a primitive, **Monoclinic** lattice. The space group was found to be *C2*. The data were integrated using the Bruker SAINT software program and scaled using the SADABS software program. Solution by direct methods (SHELXT) produced a complete phasing model consistent with the proposed structure.

All nonhydrogen atoms were refined anisotropically by full-matrix least-squares (SHELXL-2014). All carbon bonded hydrogen atoms were placed using a riding model. Their positions were constrained relative to their parent atom using the appropriate HFIX command in SHELXL-2014. All other hydrogen atoms (H-bonding) were located in the difference map (N-H, O-H). There positions were refined using "riding" model.

3.11.5.2 X-ray crystal structures

Table 3.2. Crystal data and structure refinement for compound **49**.

Identification code	tor_kh_1035_sq
Empirical formula	C54 H50 Cl6 N12 O10 S2 Sn
Molecular formula	2(C27 H25 N6 O5 S), Cl6 Sn
Formula weight	1422.57
Temperature	100.0 K
Wavelength	1.54178 Å
Crystal system	Monoclinic
Space group	C 1 2 1
Unit cell dimensions	a = 22.563(3) Å $\qquad\qquad$ α= 90°.
	b = 8.0756(10) Å $\qquad\qquad$ β= 122.326(7)°.
	c = 20.720(2) Å $\qquad\qquad$ γ = 90°.
Volume	3190.3(7) Å^3
Z	2
Density (calculated)	1.481 Mg/m^3
Absorption coefficient	6.644 mm^{-1}
F(000)	1444
Crystal size	0.2 x 0.17 x 0.15 mm^3
Crystal color, habit	colorless irregular
Theta range for data collection	2.524 to 66.567°.
Index ranges	-25<=h<=20, -9<=k<=9, -23<=l<=24
Reflections collected	14870
Independent reflections	5308 [R(int) = 0.0690]
Completeness to theta = 66.567°	97.1 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.3219 and 0.1459
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	5308 / 115 / 418
Goodness-of-fit on F^2	1.055
Final R indices [I>2sigma(I)]	R1 = 0.0583, wR2 = 0.1297
R indices (all data)	R1 = 0.0929, wR2 = 0.1515
Absolute structure parameter	0.076(8)
Extinction coefficient	0.00084(11)
Largest diff. peak and hole	0.499 and -0.607 e.Å^{-3}
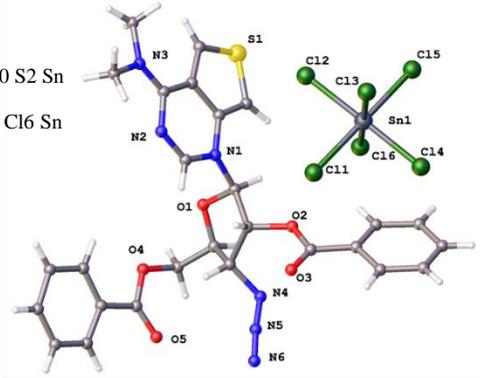
154

Table 3.3. Crystal data and structure refinement for compound **64**.

Identification code	tor153
Empirical formula	C26 H25 N5 O3 S2
Molecular formula	C26 H25 N5 O3 S2
Formula weight	519.63
Temperature	100.0 K
Wavelength	1.54178 Å
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
Unit cell dimensions	a = 4.7605(4) Å	$\alpha = 90°$.
	b = 12.8199(9) Å	$\beta = 90°$.
	c = 40.759(3) Å	$\gamma = 90°$.
Volume	2487.5(3) Å^3
Z	4
Density (calculated)	1.388 Mg/m^3
Absorption coefficient	2.263 mm^{-1}
F(000)	1088
Crystal size	0.135 x 0.045 x 0.04 mm^3
Crystal color, habit	light yellow plank
Theta range for data collection	2.168 to 63.776°.
Index ranges	-5<=h<=5, -14<=k<=13, -46<=l<=45
Reflections collected	14046
Independent reflections	4065 [R(int) = 0.0739]
Completeness to theta = 63.776°	99.4 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.5829 and 0.4428
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	4065 / 0 / 327
Goodness-of-fit on F^2	1.056
Final R indices [I>2sigma(I)]	R1 = 0.0590, wR2 = 0.1439
R indices (all data)	R1 = 0.0878, wR2 = 0.1597
Absolute structure parameter	0.04(2)
Extinction coefficient	0.0038(6)
Largest diff. peak and hole	0.318 and -0.368 e.Å^{-3}

155
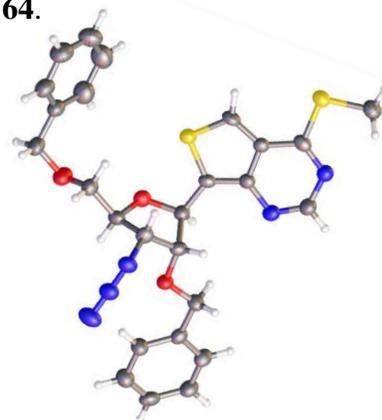
3.11.6 ^{1}H and ^{13}C NMR spectra
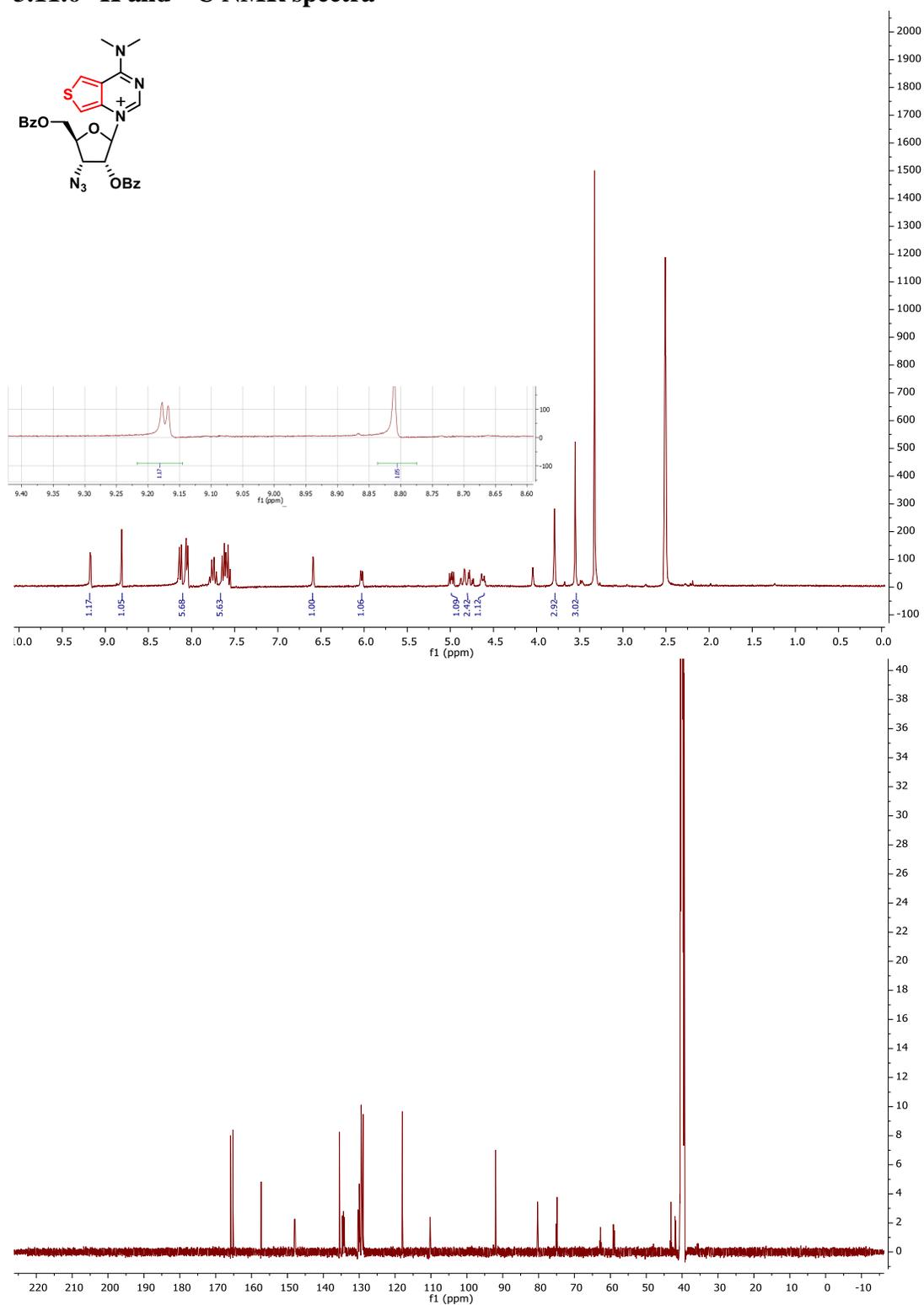
Figure 3.18. ^{1}H and ^{13}C NMR of compound **49.**
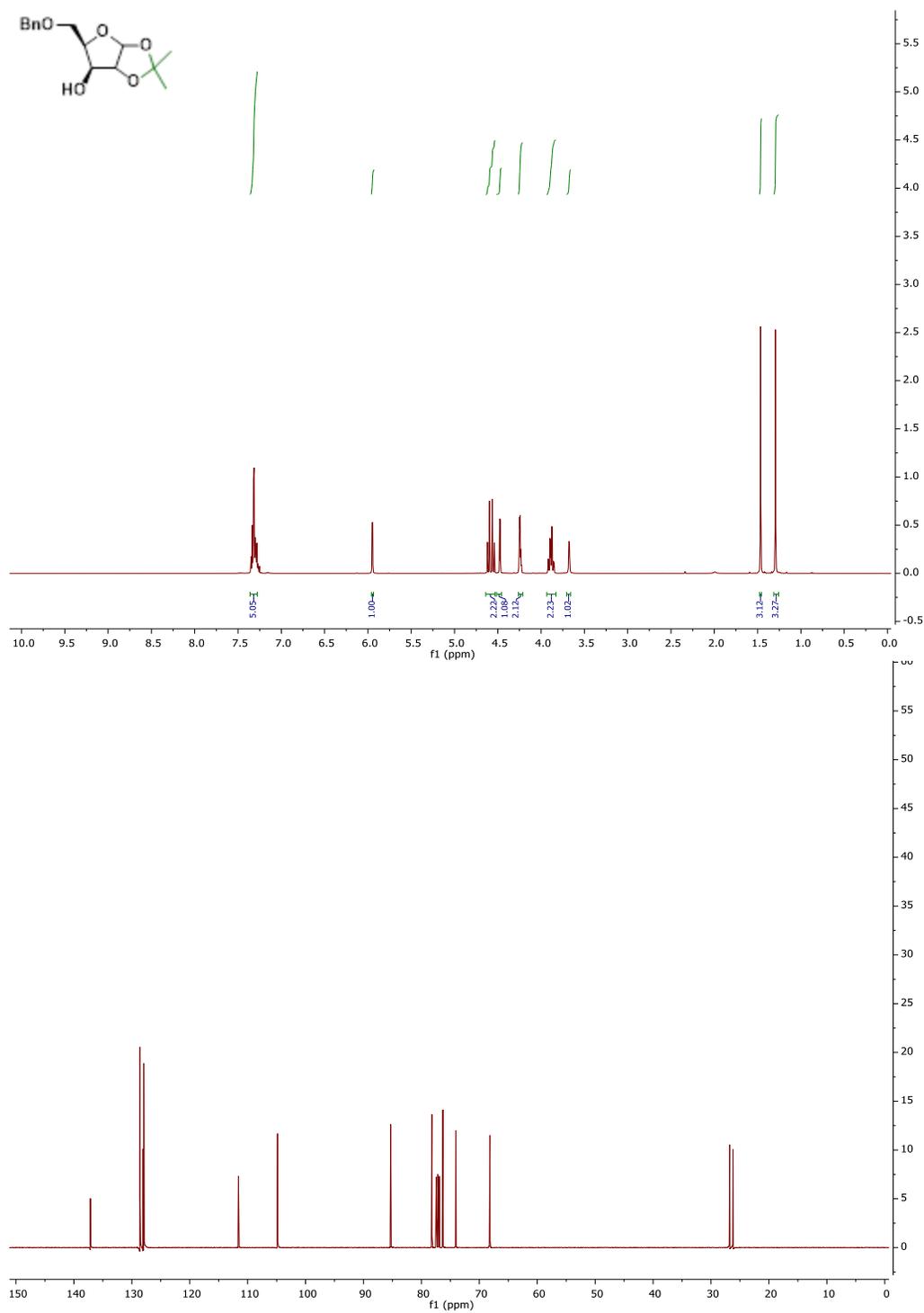
Figure 3.19. ^{1}H and ^{13}C NMR spectra of compound **52**.
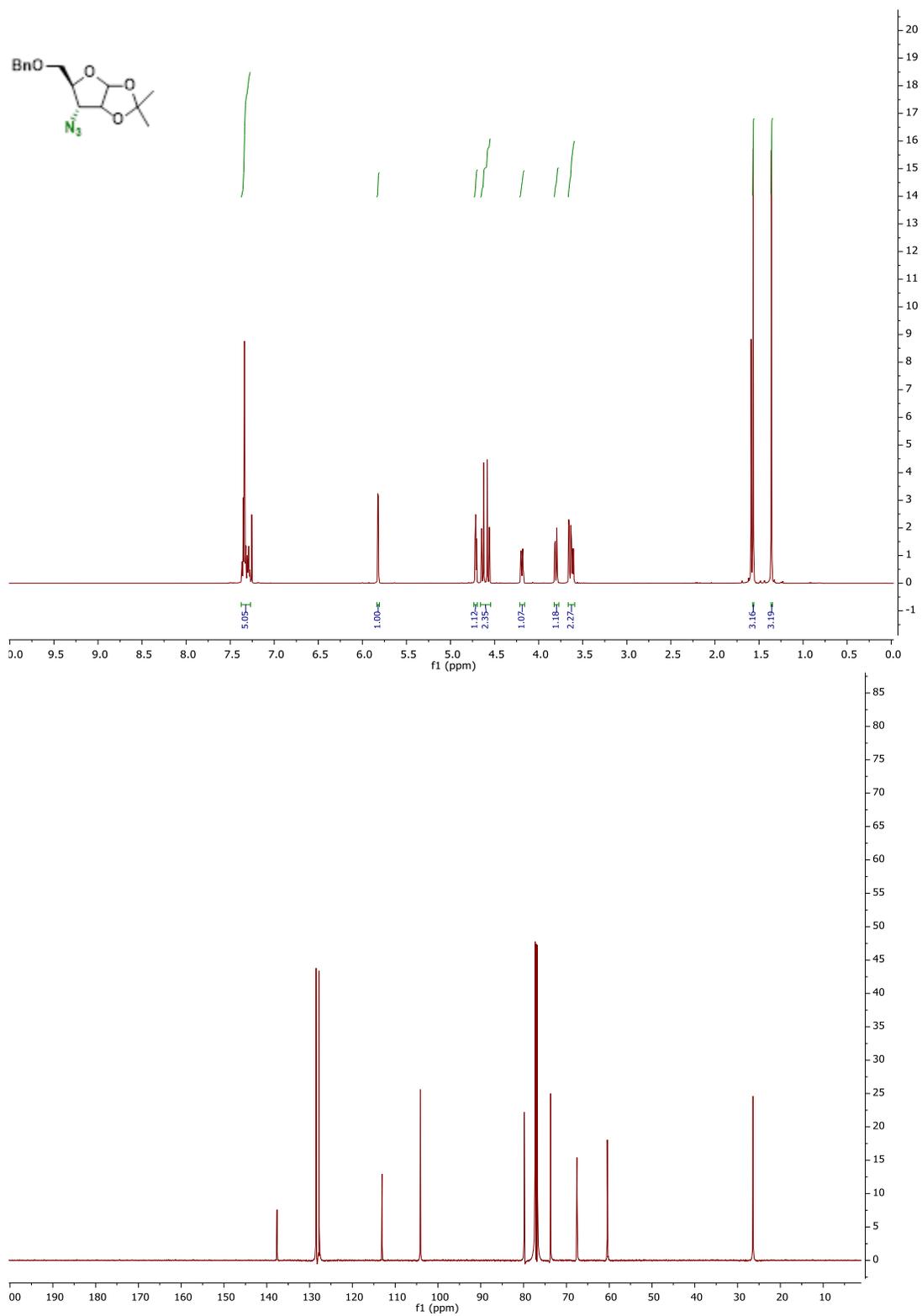
Figure 3.20. ^{1}H and ^{13}C NMR spectra of compound **53**.
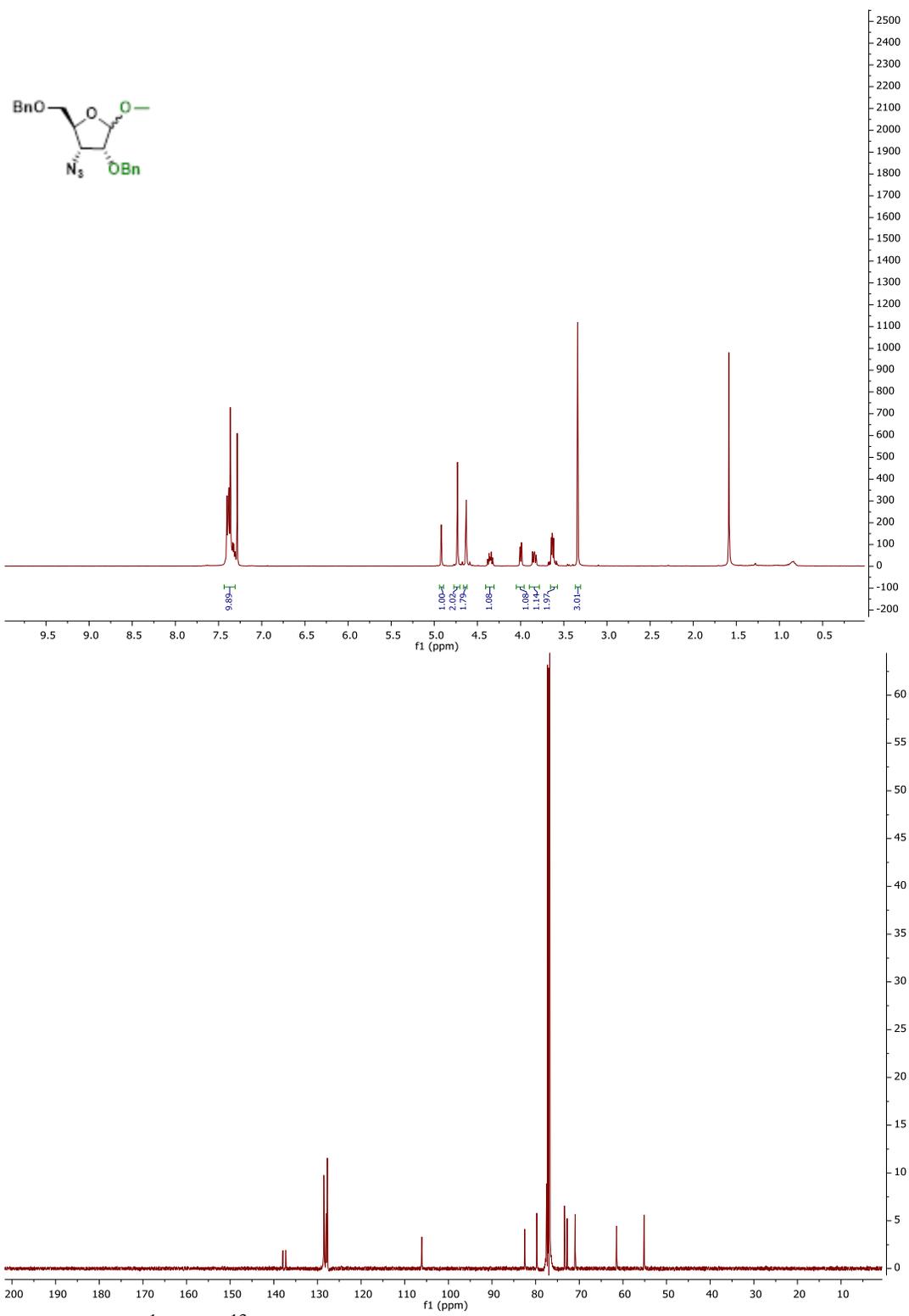
Figure 3.21. ^{1}H and ^{13}C NMR spectra of one anomer of compound **54**.
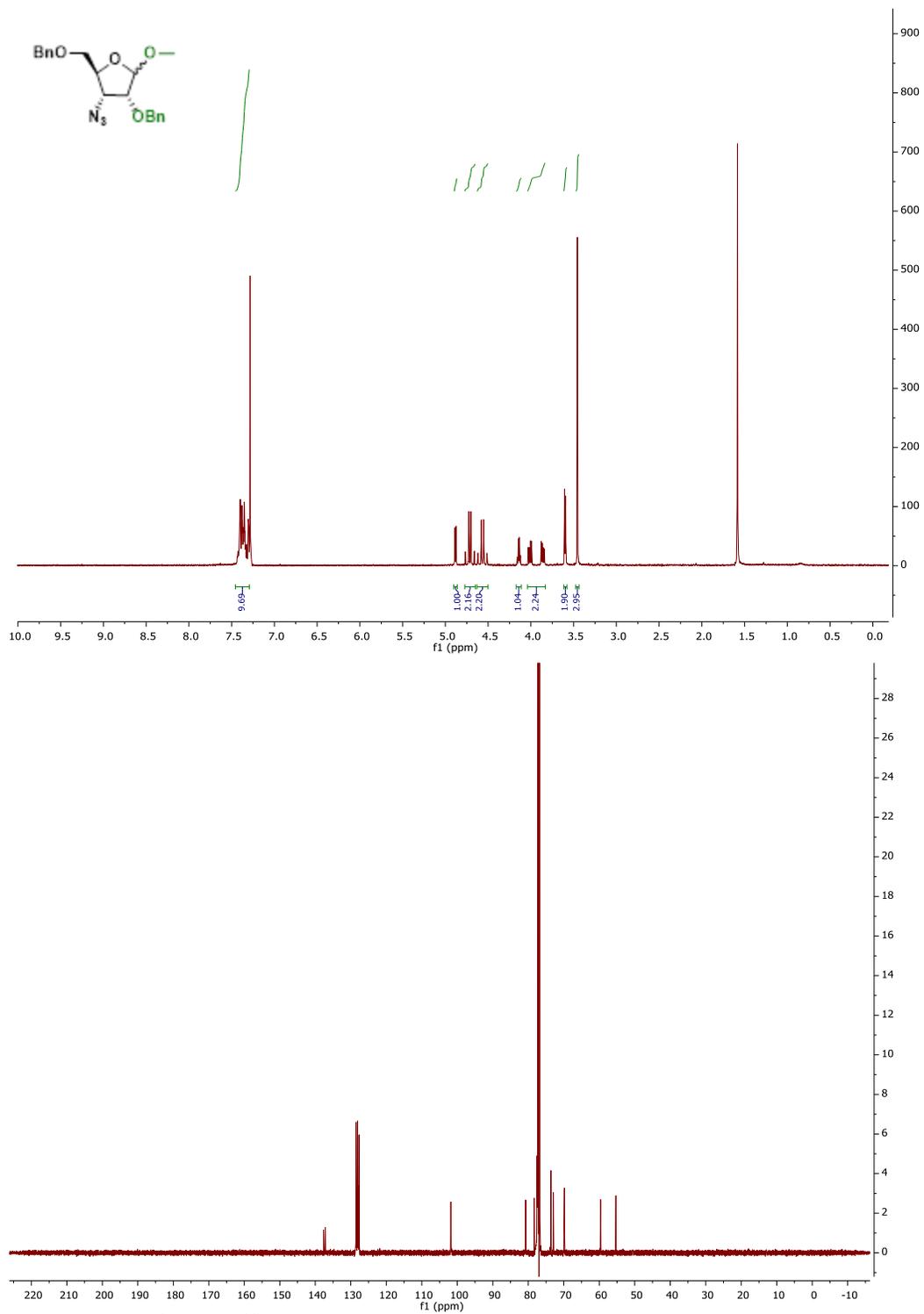
Figure 3.22. ^{1}H and ^{13}C NMR spectra of one anomer of compound **54**.
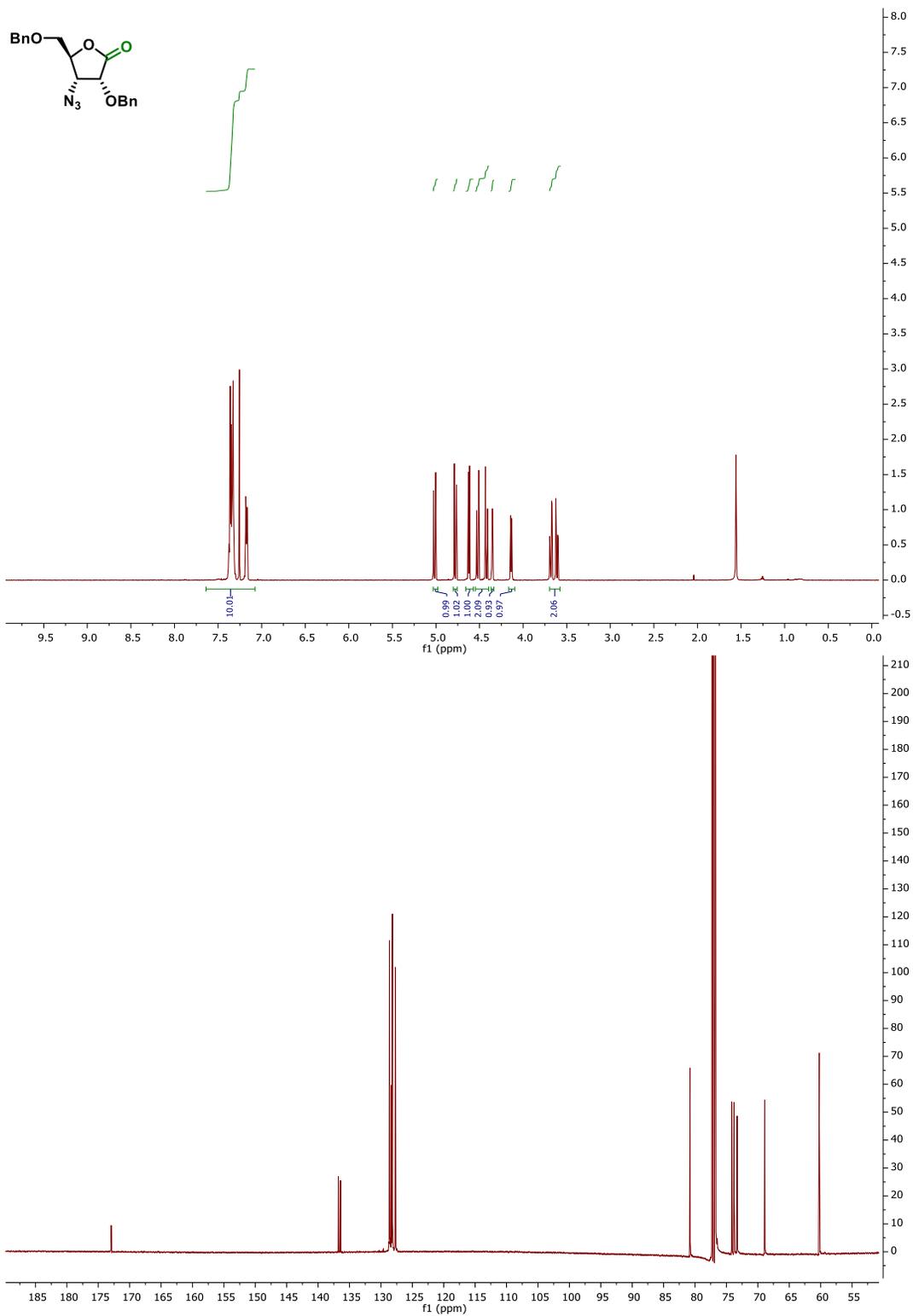
Figure 3.23. ^{1}H and ^{13}C NMR spectra of compound **55**.
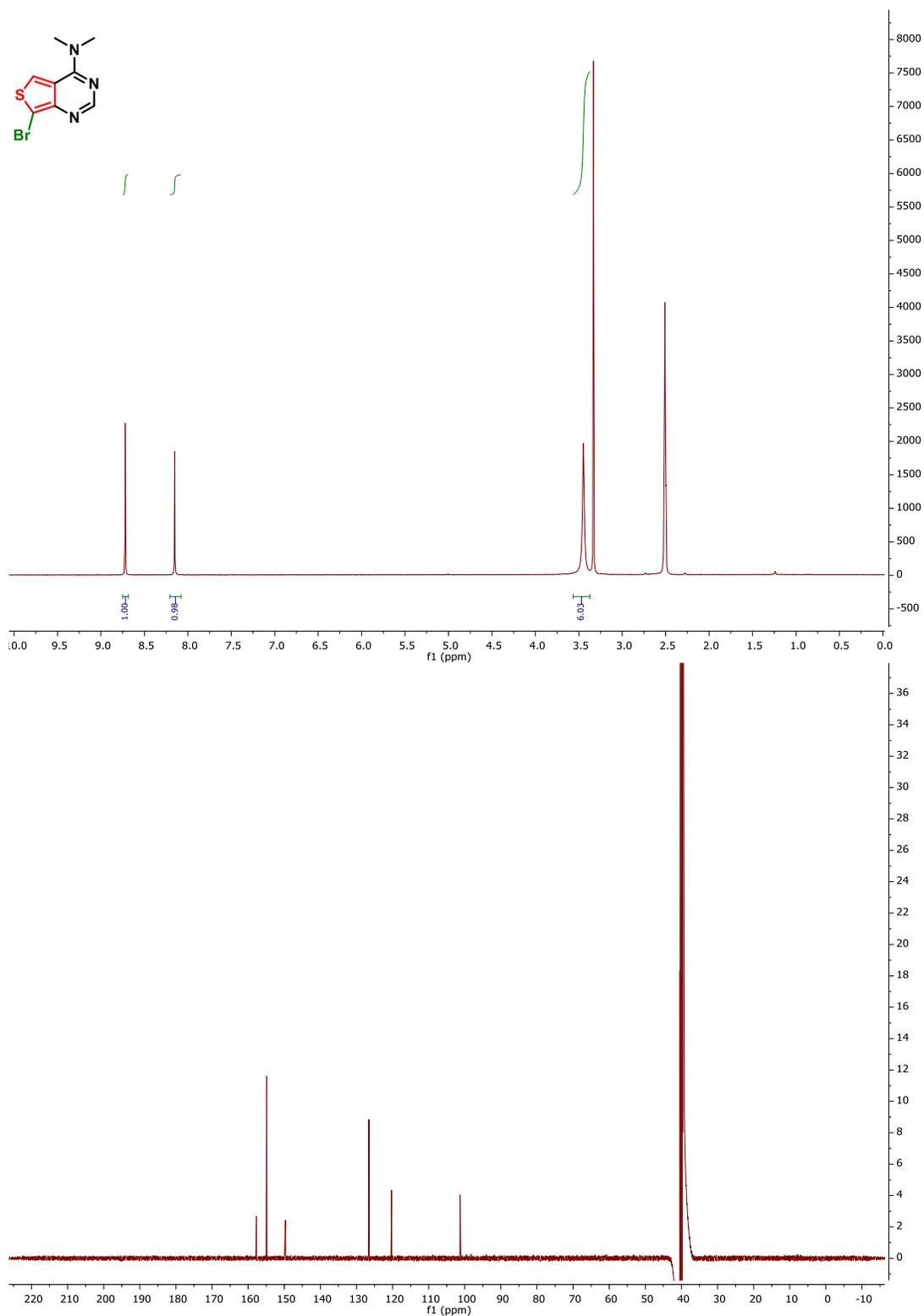
Figure 3.24. ^{1}H and ^{13}C NMR spectra of nucleobase **50**.
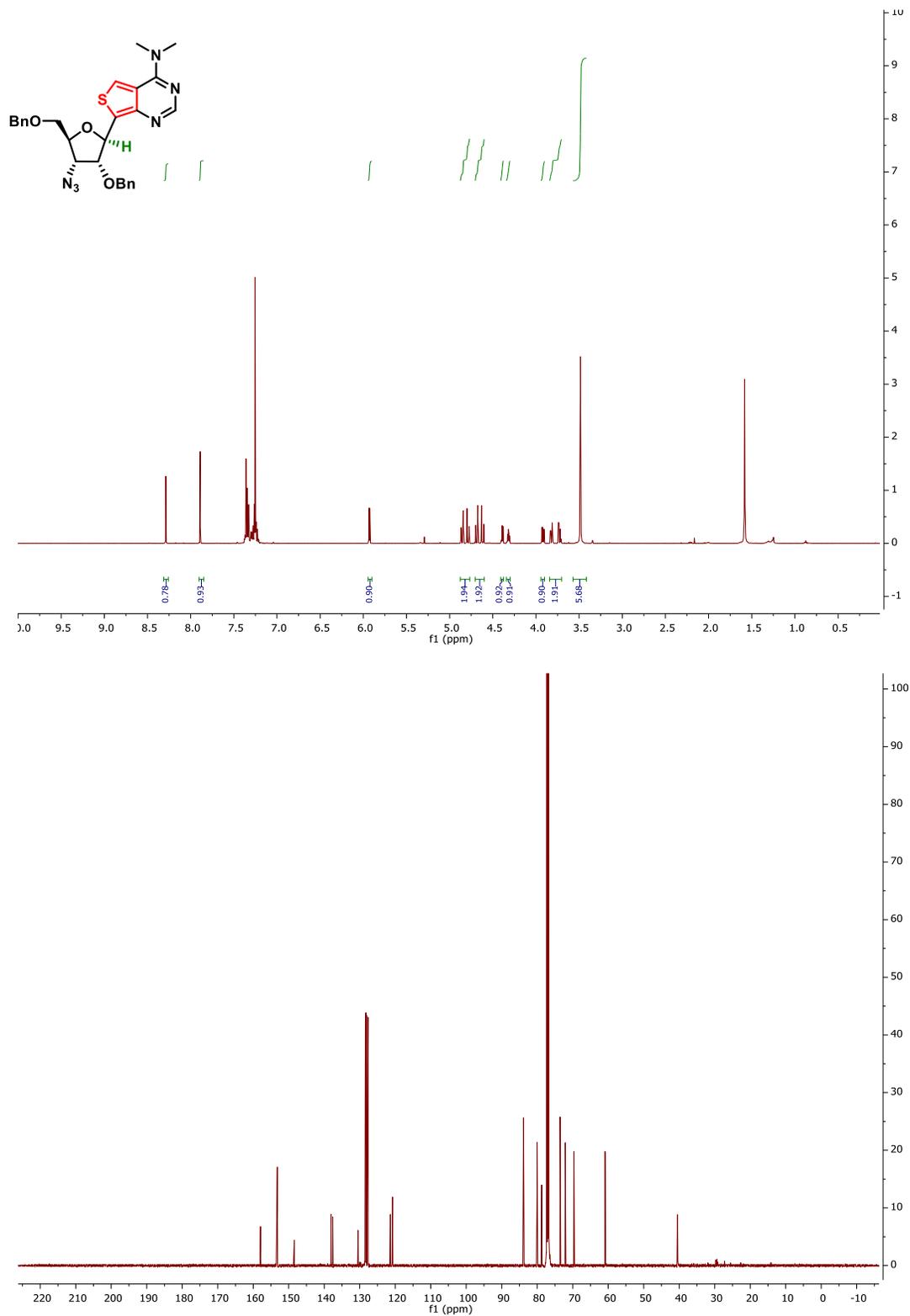
Figure 3.25. ^{1}H and ^{13}C NMR spectra of compound **59**.
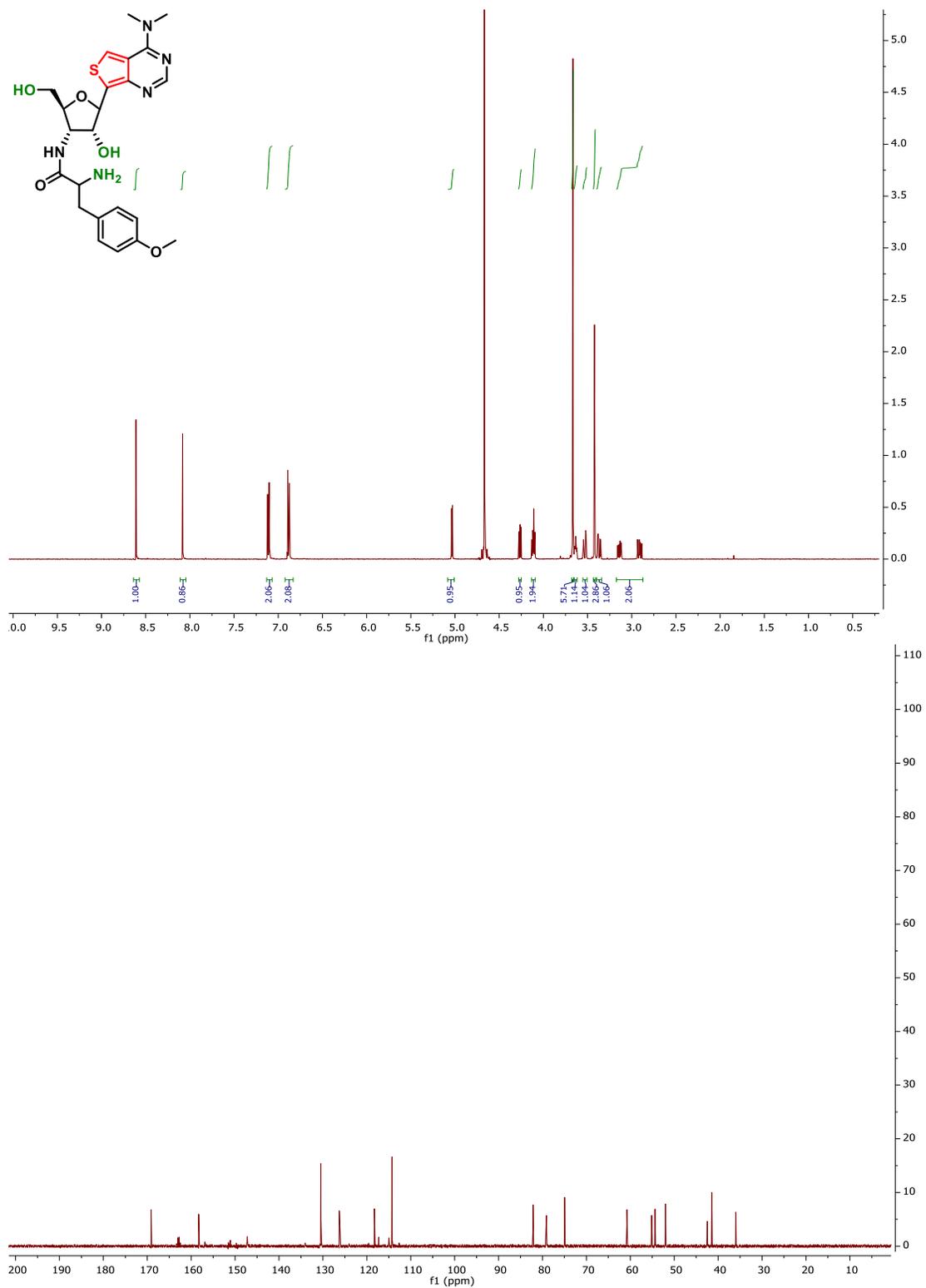
Figure 3.26. ^{1}H and ^{13}C NMR spectra of compound **33.**
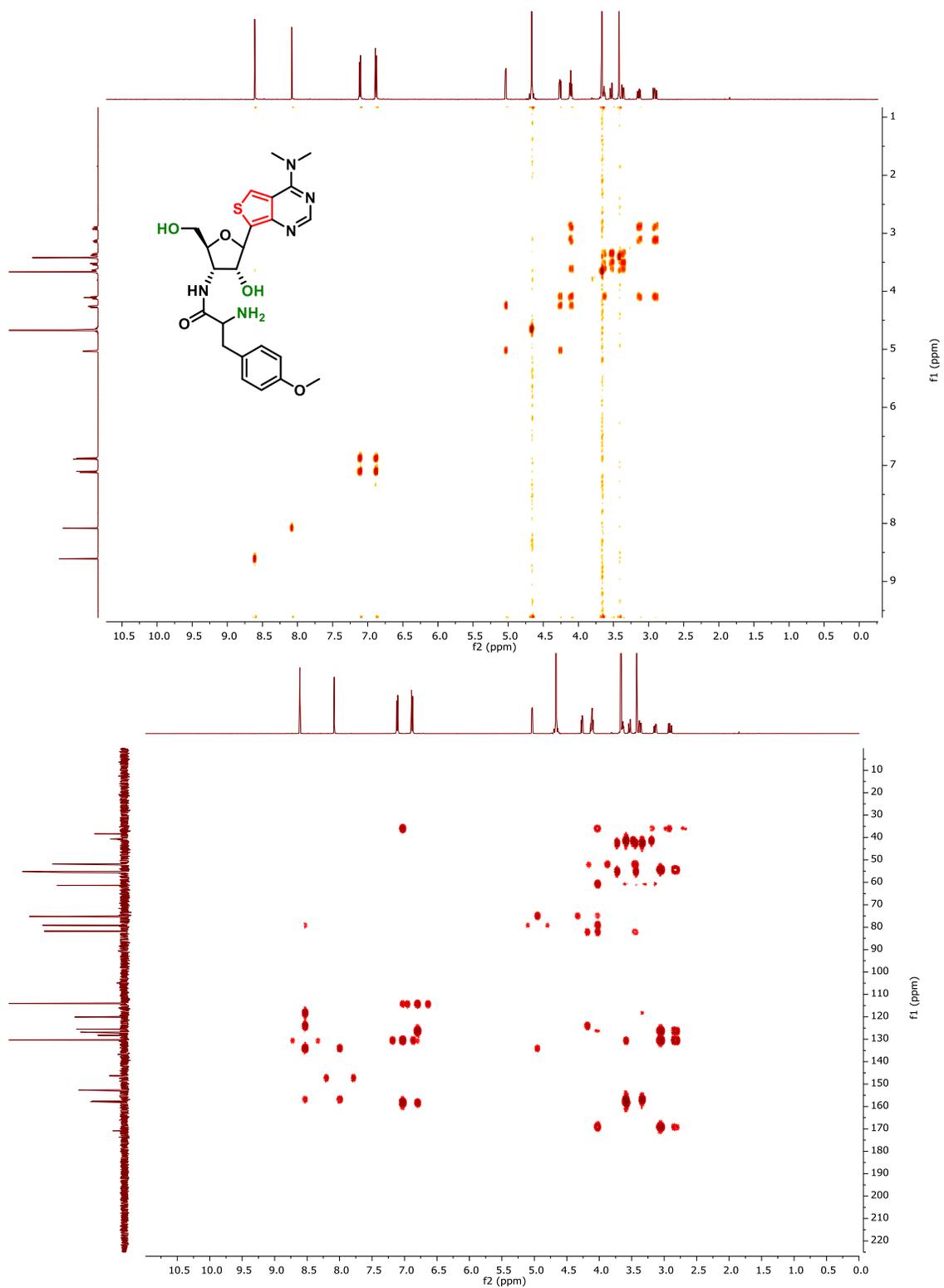
Figure 3.27. COSY and HMBC spectra of Thpuromycin **33**.
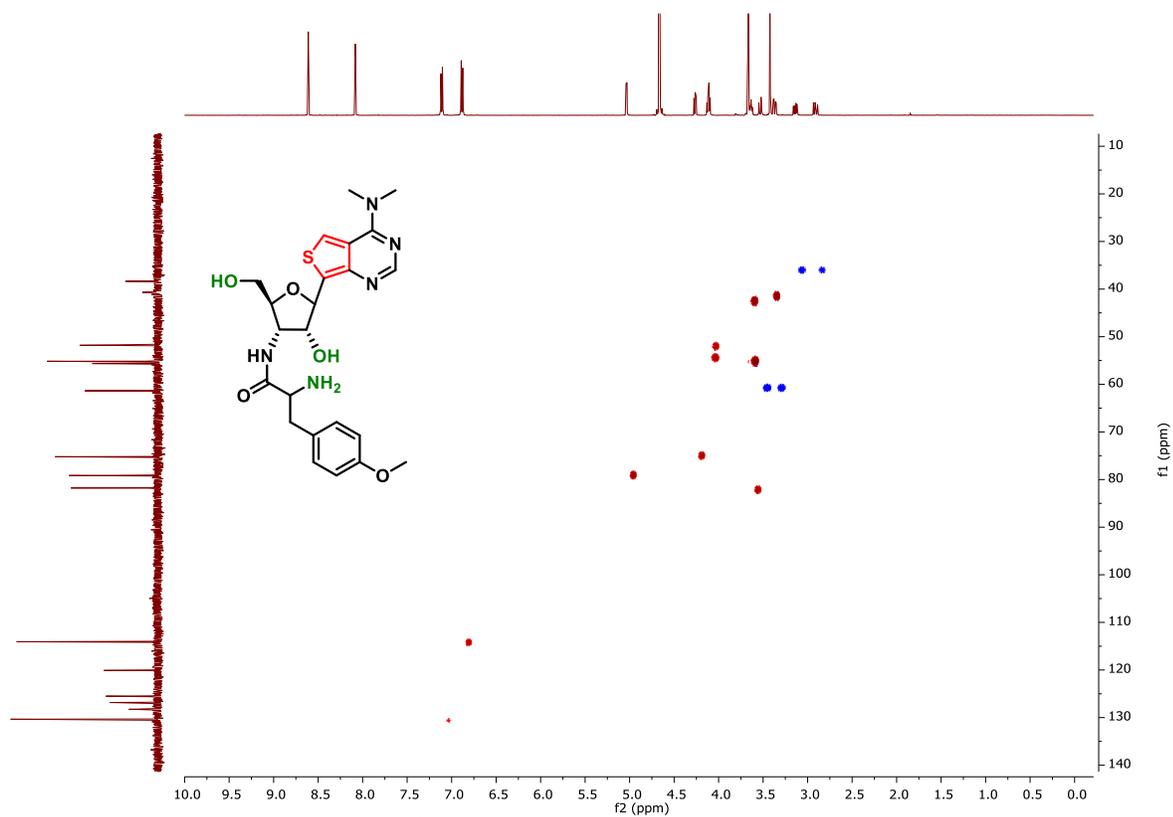
Figure 3.28. HSQC spectra of Thpuromycin **33**.
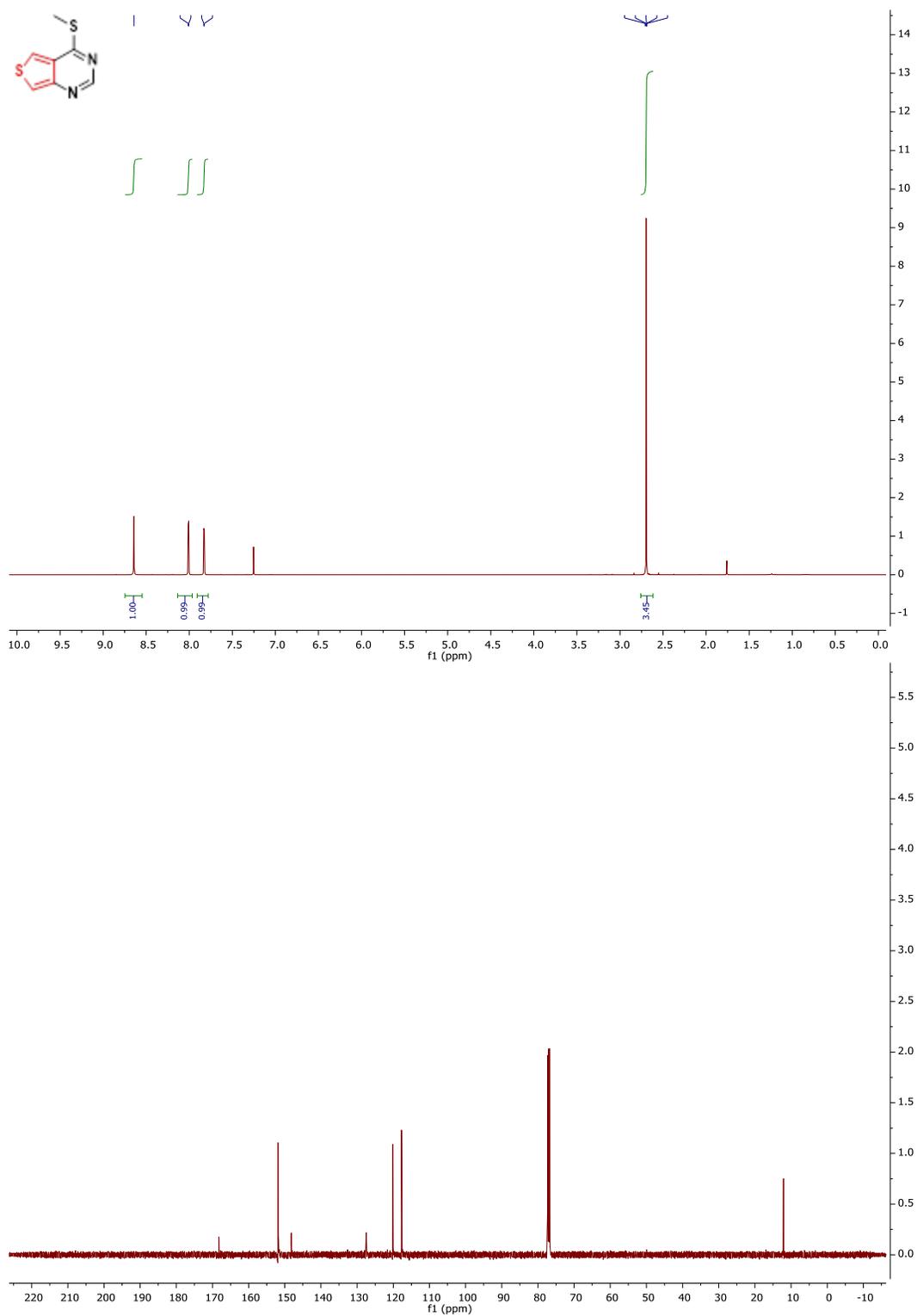
Figure 3.29. ^{1}H and ^{13}C NMR spectra of compound **61.**
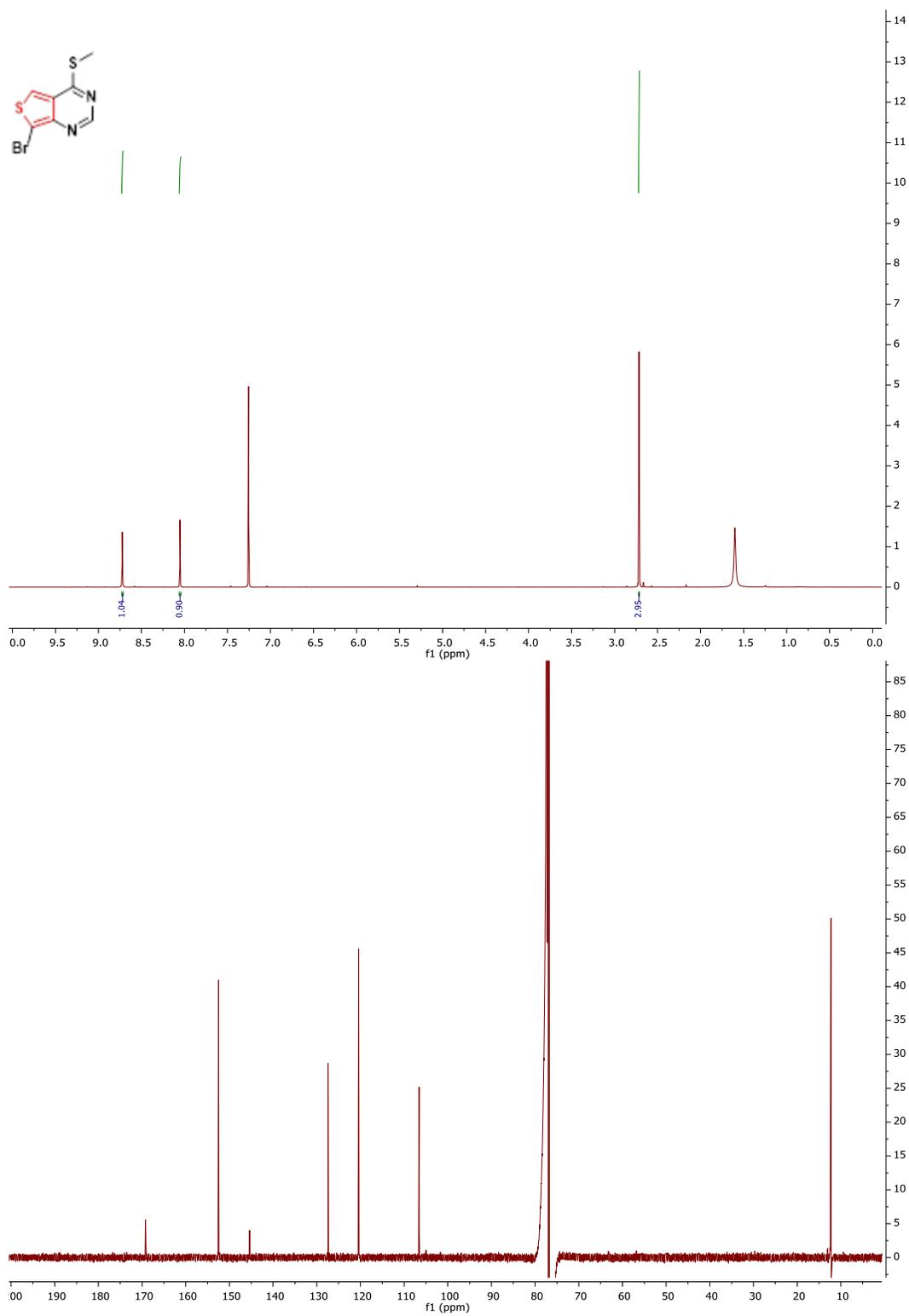
Figure 3.30. ^{1}H and ^{13}C NMR spectra of compound **62.**
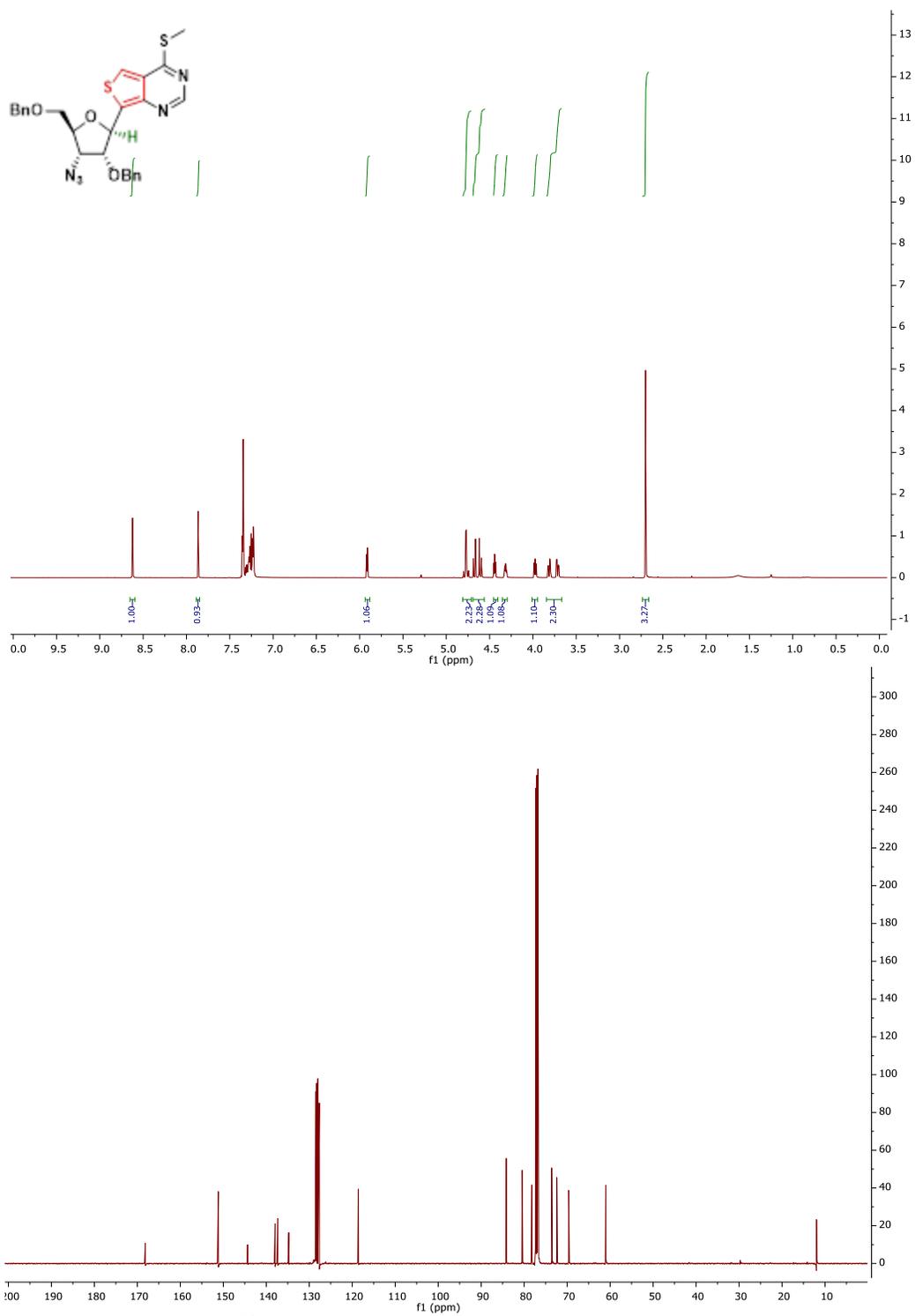
Figure 3.31. ^{1}H and ^{13}C NMR spectra of compound **64**.

169
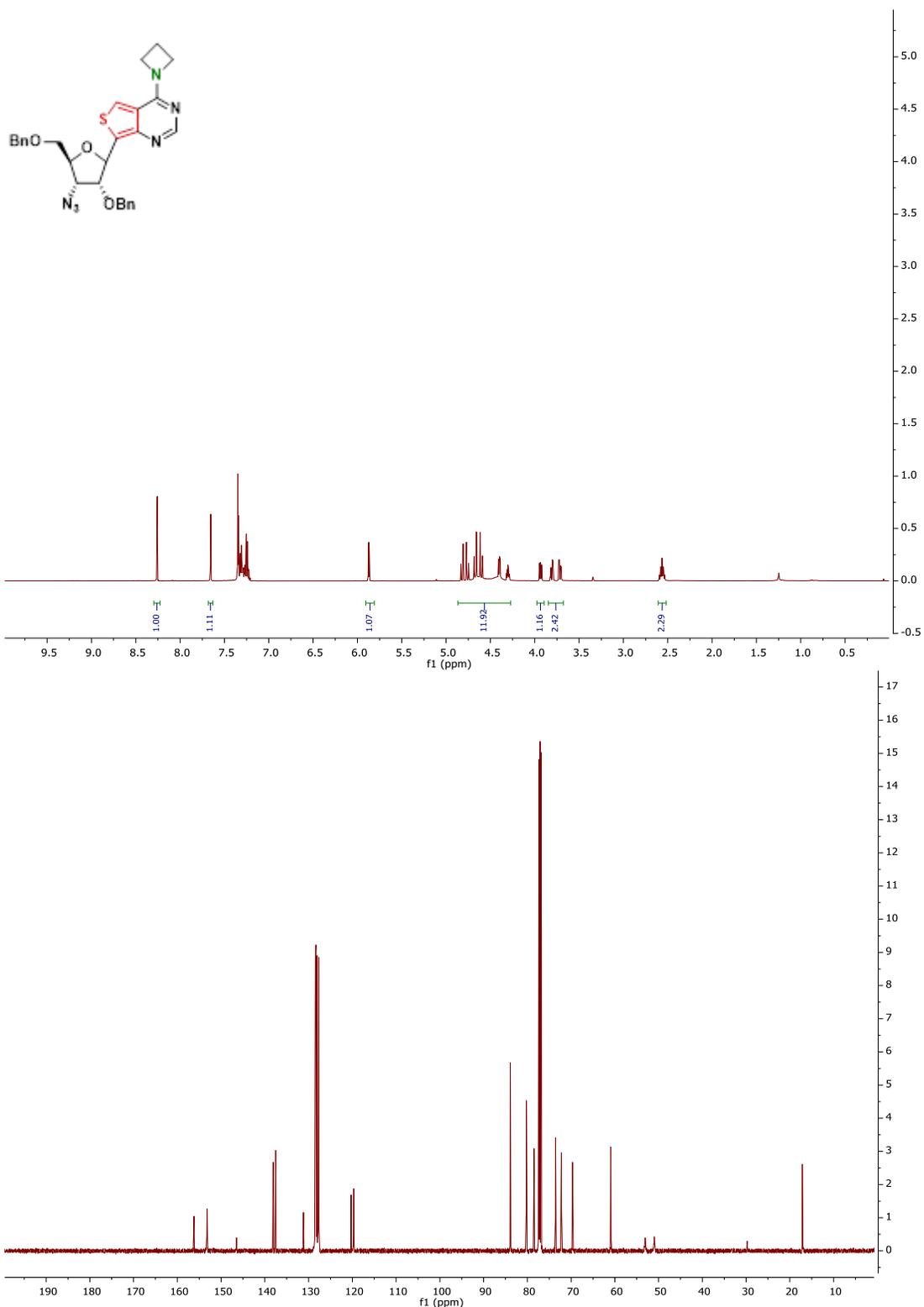
Figure 3.32. ^{1}H and ^{13}C NMR spectra of compound **65**.
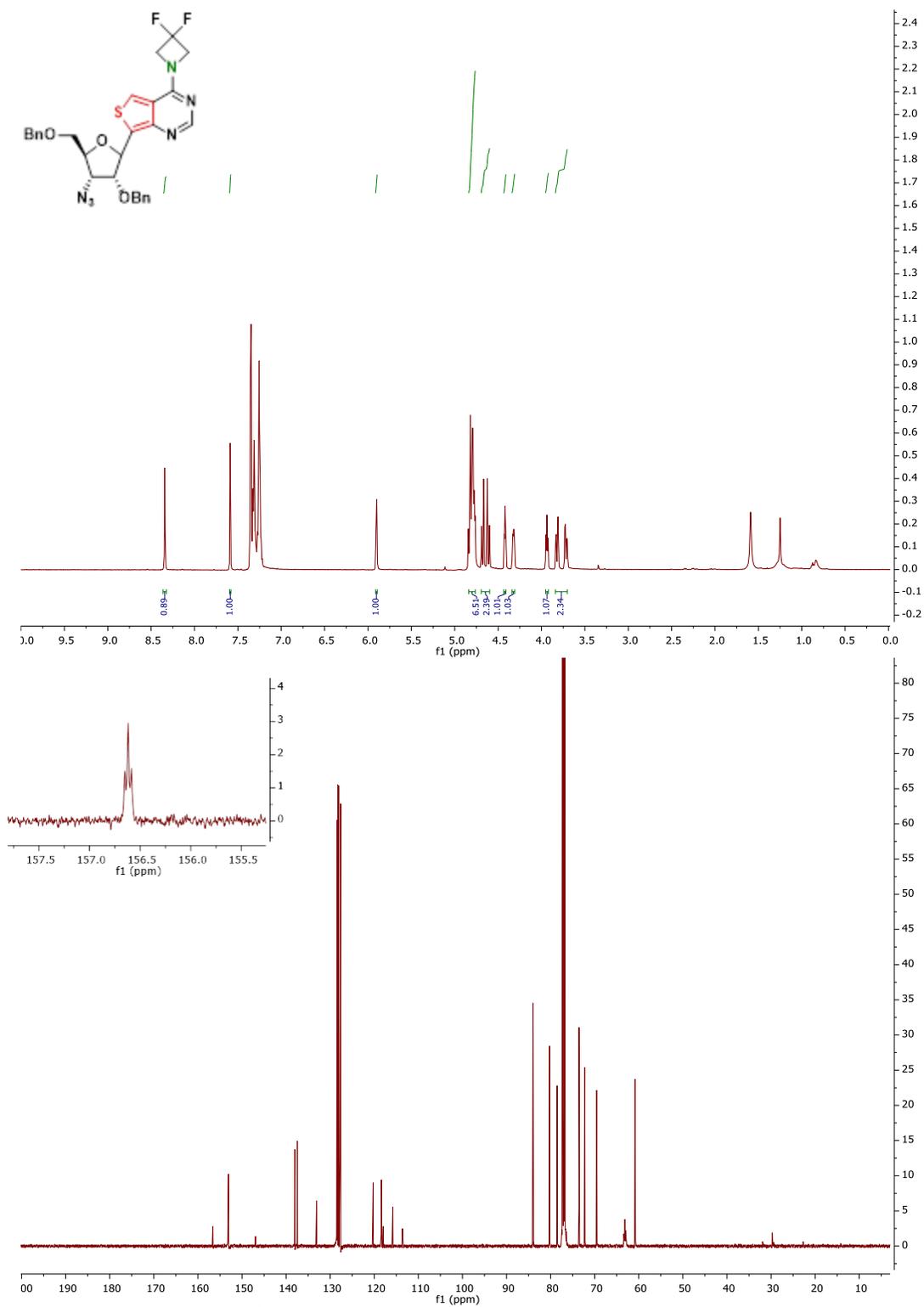
Figure 3.33. ^{1}H and ^{13}C NMR spectra of compound **66**.
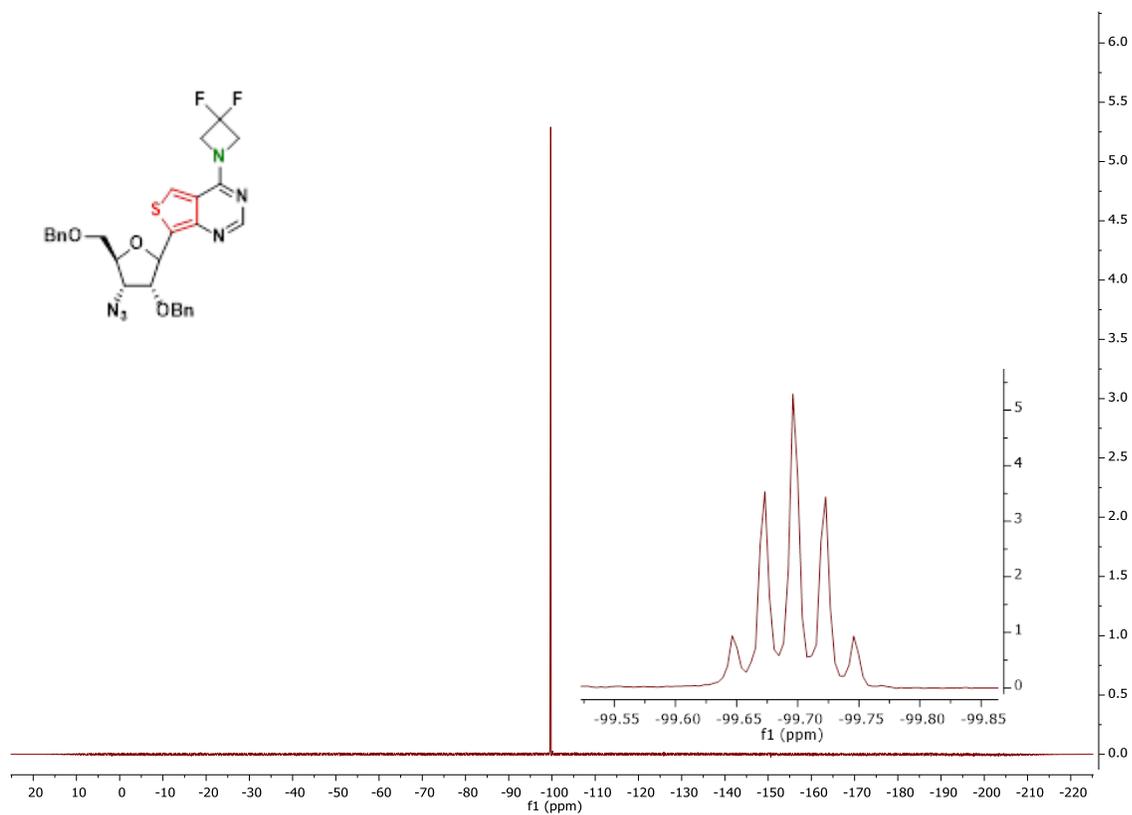
Figure 3.34. ^{19}F NMR of compound **66**.
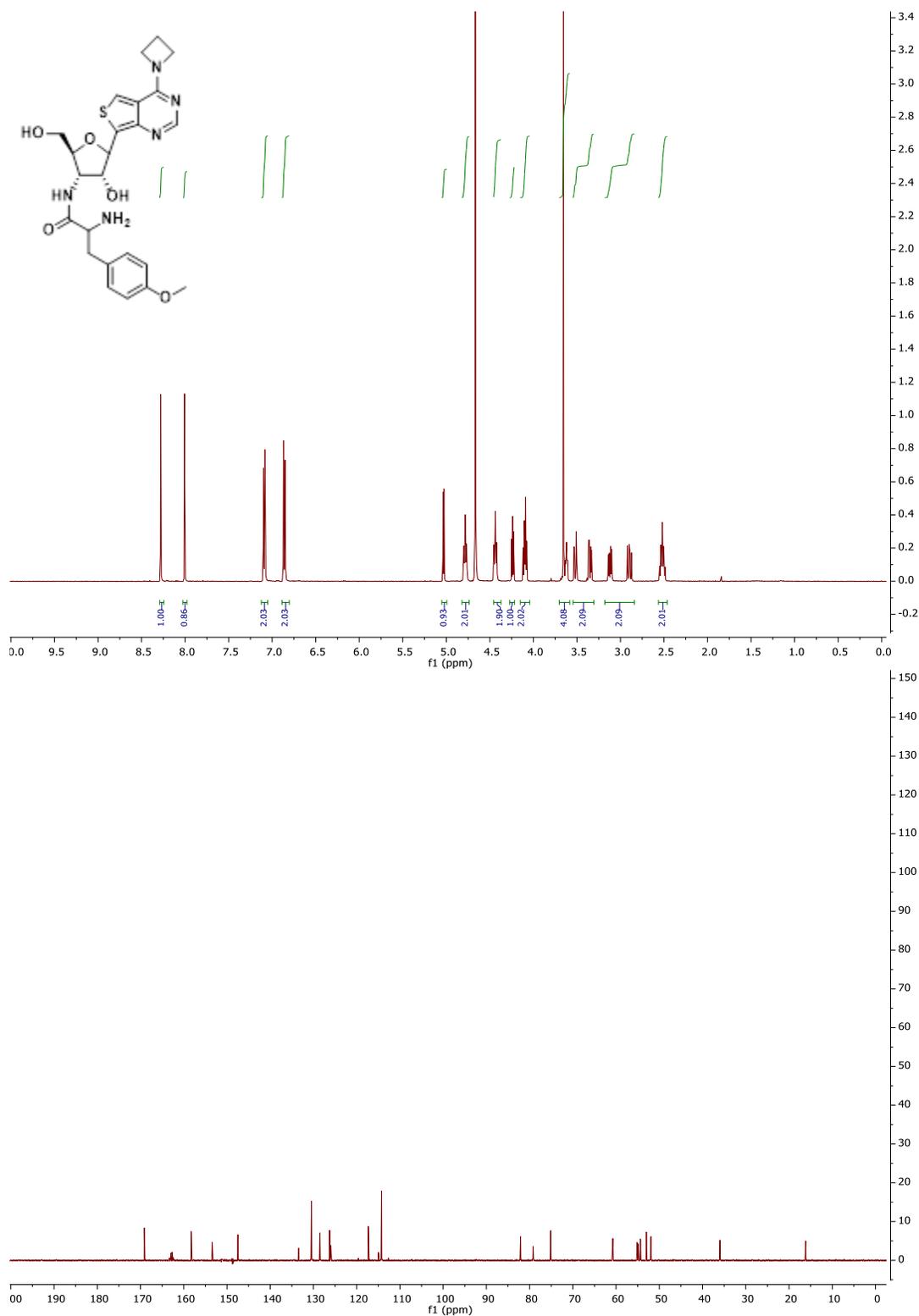
Figure 3.35. ^{1}H and ^{13}C NMR spectra of antibiotic **69**.
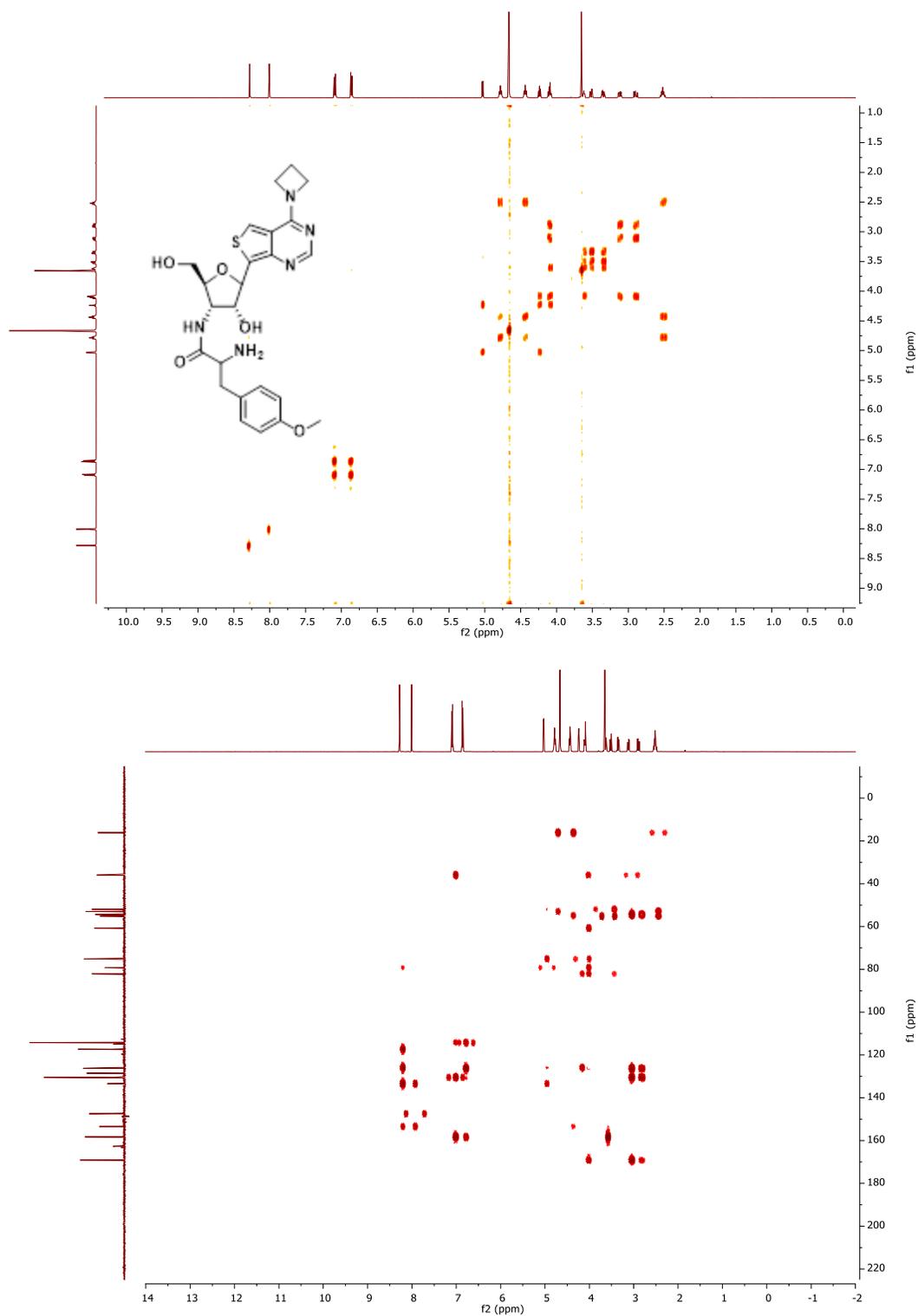
Figure 3.36. COSY and HMBC spectra of antibiotic **69**.
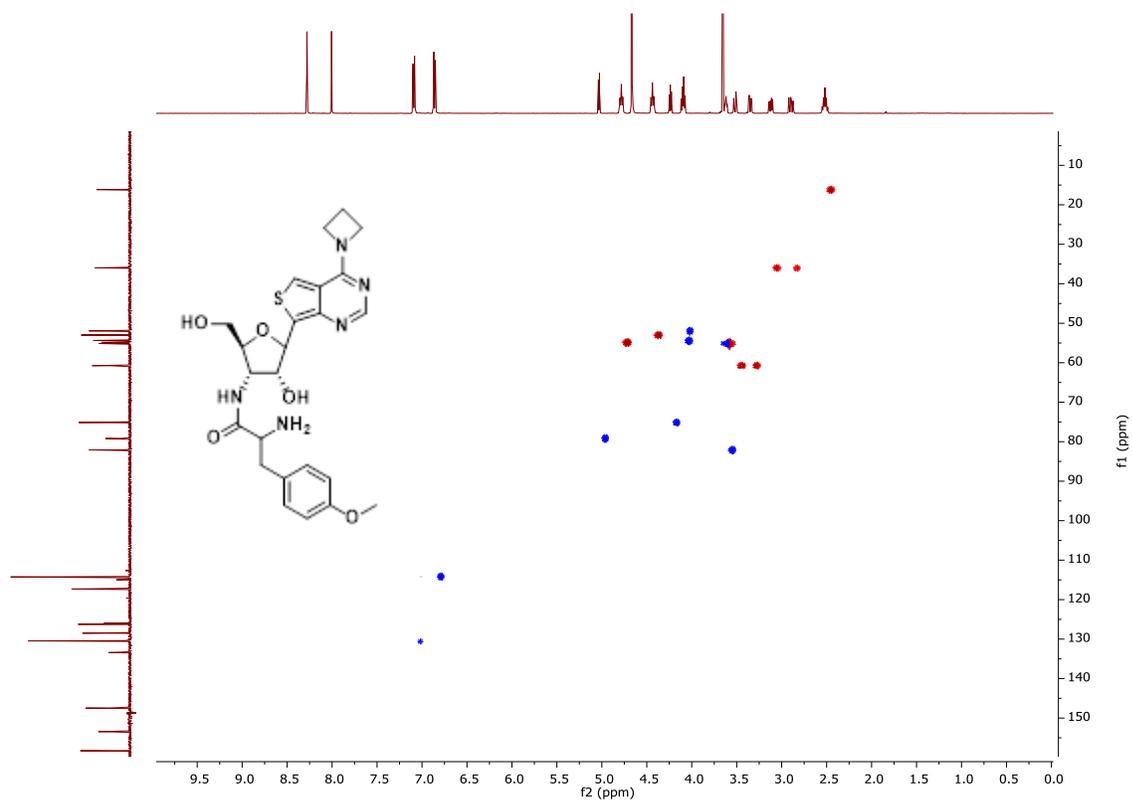
Figure 3.37. HSQC spectra of antibiotic **69**.
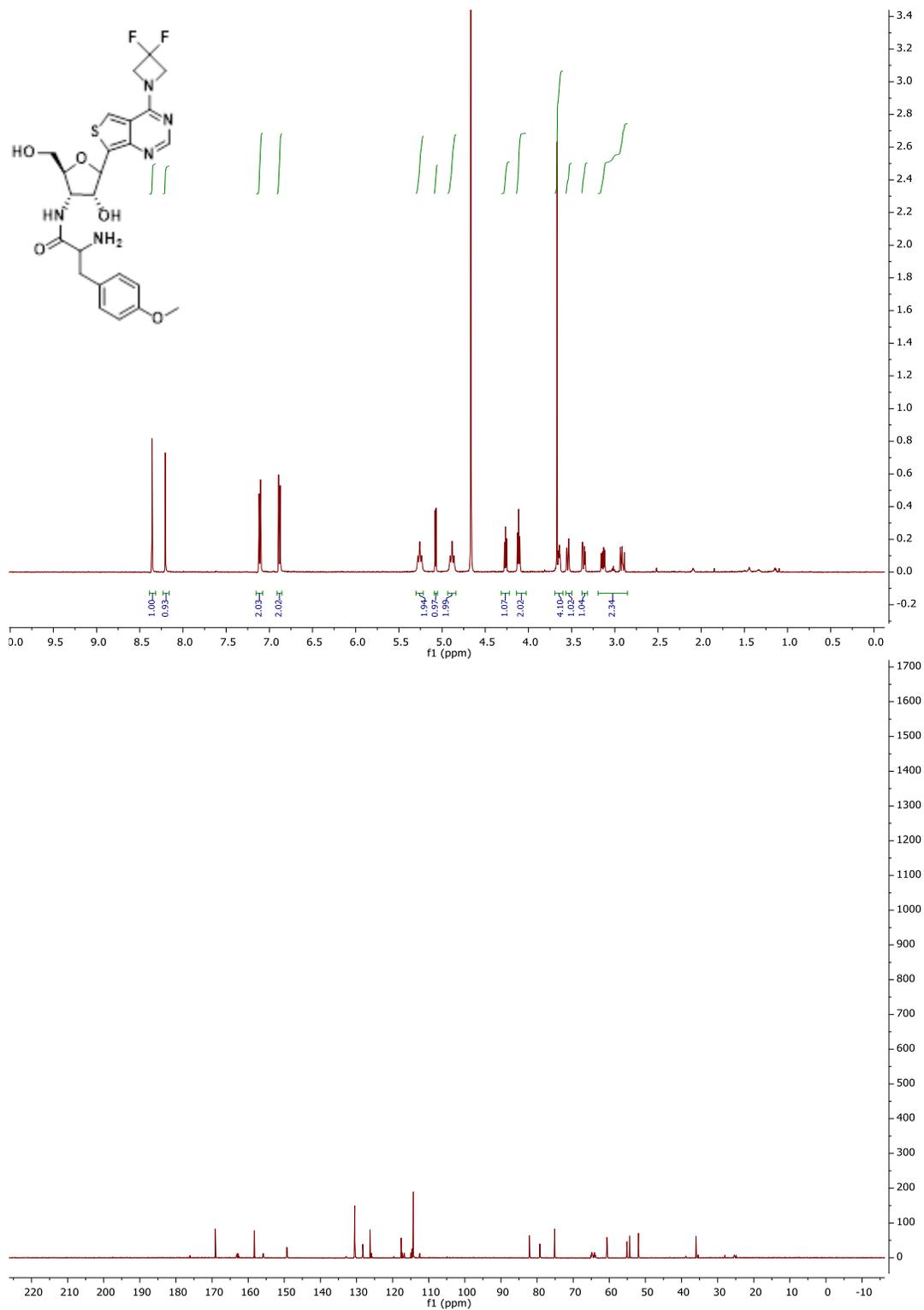
Figure 3.38. ^{1}H and ^{13}C NMR spectra of antibiotic **70**.
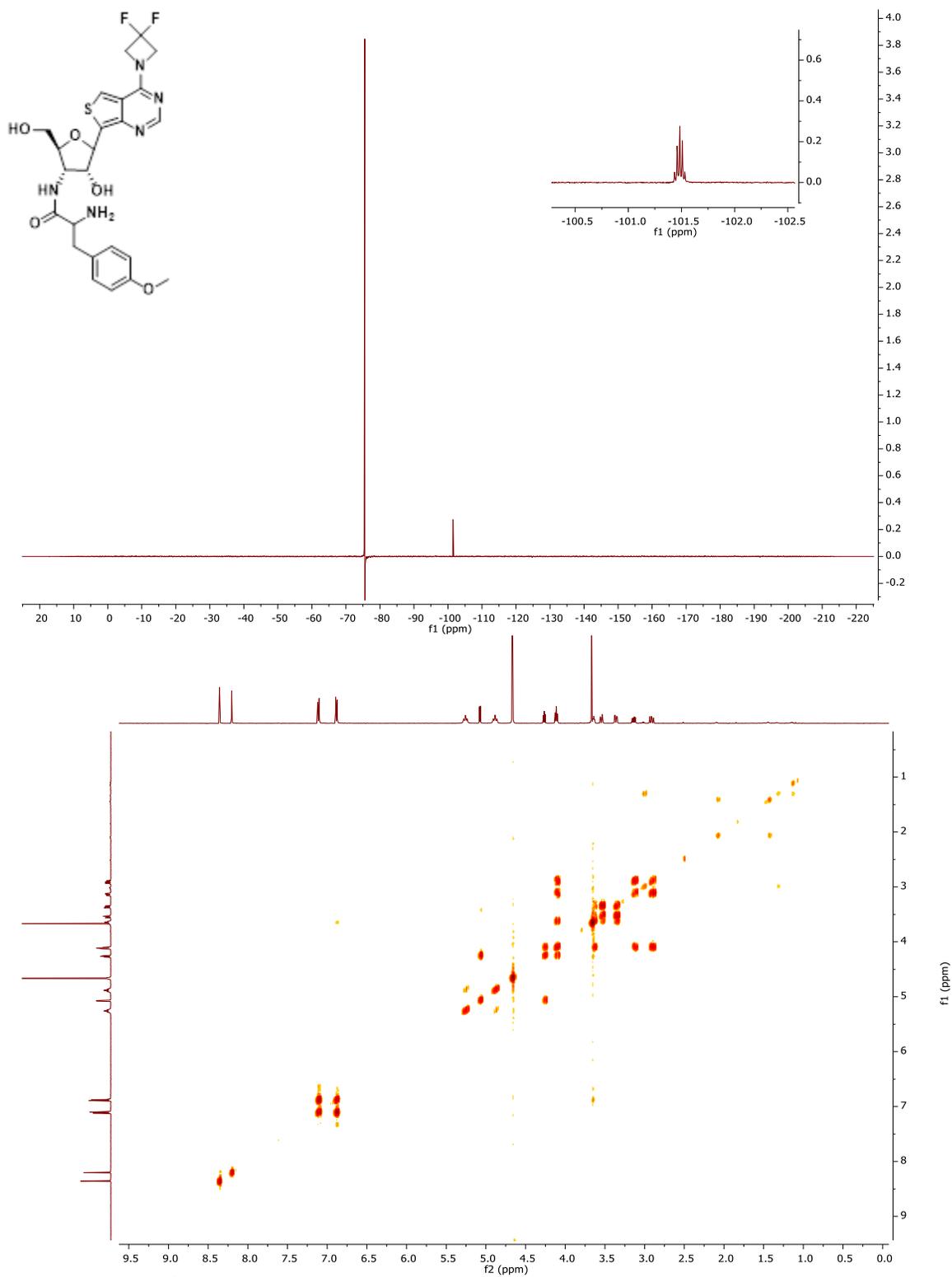
Figure 3.39. ^{19}F NMR and COSY spectra of antibiotic **70**.
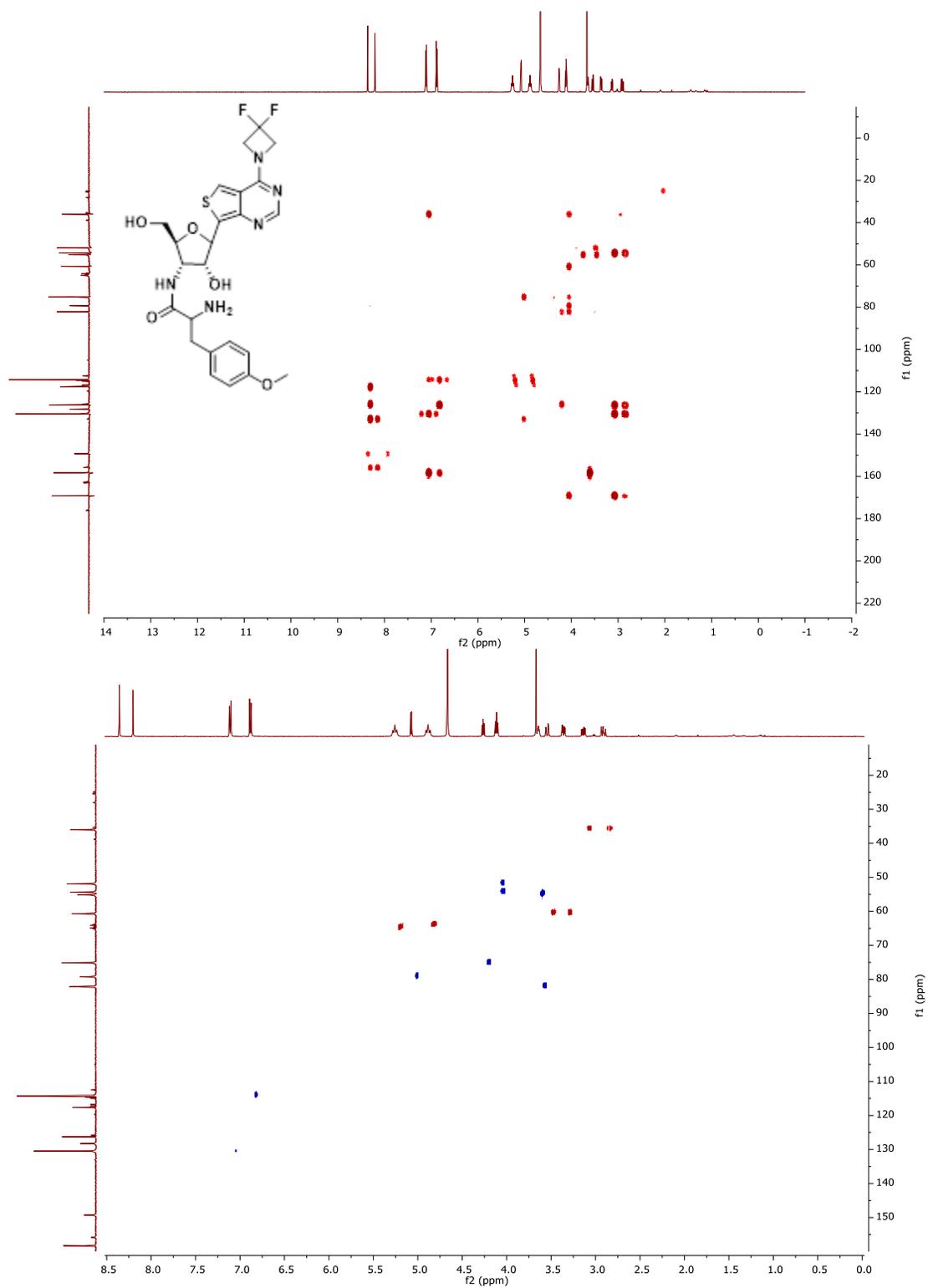
Figure 3.40. HMBC and HSQC spectra of compound **70**.

CHAPTER 4: Biochemical Utility of Fluorescent Puromycin Analogues

4.1 Biological Activity

4.1.1 *In Vitro* Translation Assays

Before probing the biological ultility of the new antibiotics (figure 4.1), we opted to measure their photophysical properties (figure 4.2, table 4.1). Compounds **33**, **69**, and **70** mirrored nucleobases **2a**, **3a**, and **4a** in the overall shape of their emission bands. However, their quantum yields differed drastically compared to their aglycon counterparts. The quantum yield of antibiotic **33** improved ten-fold compared to nucleobase **2a** (0.004 vs 0.0005), whilst **69**'s
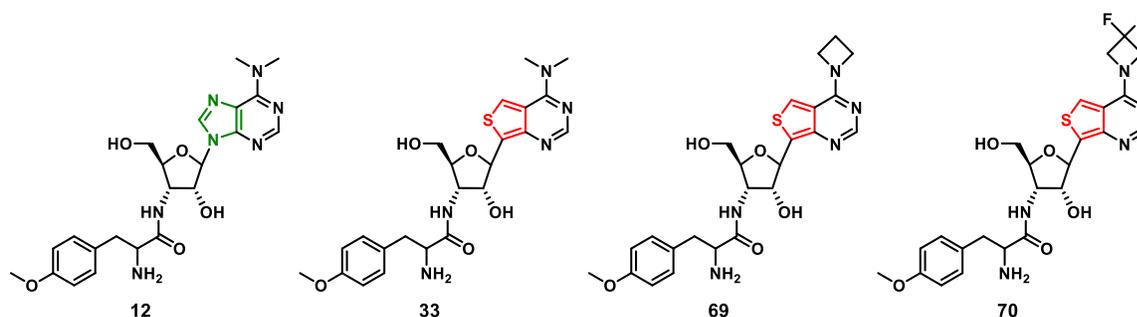
Figure 4.1. Chemical structures of antibiotics studied in this work.

quantum improved three-fold compared to **3a** (0.21 vs 0.06). A drop in quantum efficiency between **70** and **4a** was observed (0.31 vs 0.64), which was also observed previously with nucleoside **11** (0.30 vs 0.64). A potential explanation for this behavior lies in the hydrophobicity of the dimethyl and azetidinyl substituents, as opposed to the more "hydrophilic" characteristics imparted by the two fluorine atoms on the azetidine ring. Given hydration significantly impacts these chromophores, the dipoles imparted by the two fluorine atoms may provide an avenue for the fluorophores to non-radiatively relax through interactions with water in the excited state once the ribose moiety (and subsequently the full antibiotic skeleton) are introduced.
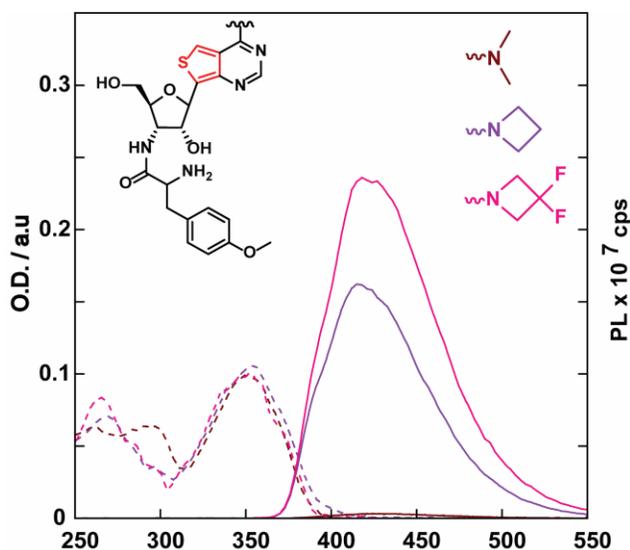
Figure 4.2. Absorption (dashed line) and emission (solid line) spectra of antibiotics **33**, **69**, and **70** in PBS. The absorption and emission spectra were normalized to 0.1 intensity at the excitation wavelength used to measure the emission spectra (350 nm).

Table 4.1 Photophysical Properties of Antibiotics **33**, **69**, and **70**

Antibiotic	Solvent	λ_{abs}	λ_{em} $(\Phi)^a$	Stokes Shift[a]
33	PBS	354	424 (0.004)	4.67±0.16
69	PBS	354	417 (0.21)	4.21±0.07
70	PBS	353	419 (0.31)	4.44±0.07

[a] λ_{abs}, λ_{em} and Stokes shift are reported in nm, nm and 10^3 cm^{-1} respectively. All the photophysical values reflect the average over three independent measurements.

To demonstrate the utility of the new puromycin derivatives (figure 4.1) and analyze their biological activity, we opted to probe the dose-dependent inhibition of GFP translation. Other puromycin analogues synthesized indicate that no significant loss of activity is observed upon modification of the puromycin skeleton.[1-3] However, changes to the O-methyl tyrosine arm can drastically affect activity if the α-amino group or amino acid identity is affected.[3,4] Given that the modifications in our analogues were focused on the nucleobase, we hypothesized that little to no different in translation inhibition would be observed compared to native puromycin. To test this, we opted to utilize a rabbit reticulocyte translation system as the expression medium for a 5'-capped mRNA encoding the GFP protein, with the intention of disrupting synthesis of the fluorescent protein through the addition of antibiotic. Additionally, the ability to monitor the

fluorescence of both GFP and the fluorescent analogues provides a unique opportunity to analyze the system through two different ways. The dose-dependent response of GFP translation inhibition by puromycin and its analogues monitoring GFP expression are given in (figure 4.3). All derivatives exhibited similar termination of GFP expression, with near total inhibition at 8 µM of antibiotic. However, it appears that the native antibiotic reaches the terminus with slightly lower concentrations than the other analogues. IC_{50} values extrapolated through Sigmoidal-Hill plots of endpoint GFP fluorescence vs antibiotic dose indicates that native puromycin is indeed the most effective (< 1 µM), followed by antibiotic **69** (2 µM), antibiotic **33** (2.3 µM) and lastly **70** (2.8 µM) (figure 4.3). Intriguingly, there appears to be a slight structure-activity relationship given that **70** is the least effective yet deviates the most from the classical structure. Whilst monitoring fluorescence of GFP over time proved useful, unfortunately the fluorescence of the nucleoside antibiotics did not change significantly over time, indicating that more detailed experiments such as single-molecule spectroscopy and ribosome synchronization would be necessary to both monitor protein synthesis and analyze signal changes upon ribosome binding.
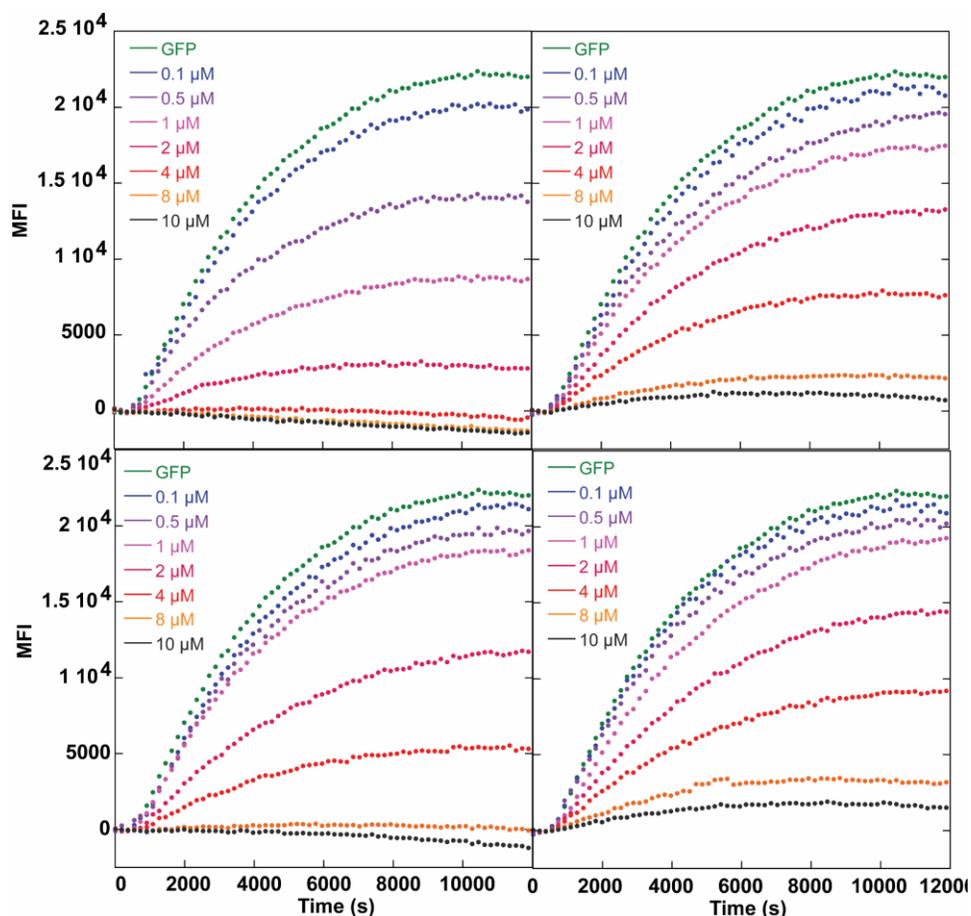
Figure 4.3. Dose-dependent inhibition of *in vitro* GFP expression with (top left) puromycin (**12**), (top right) thpuromycin (**33**), (bottom left) azetitidine-modified Thpuromycin (**69**), and (bottom right) difluoroazetidine-modified Thpuromycin (**70**).

4.2 Antibacterial Activity

To further analyze the antibiotic activity of the new derivatives, various gram-positive and gram-negative bacteria were subjected to the puromycin analogues to confirm the structural modifications were benign. MIC values from an MTT assay are provided in table 4.2. Native puromycin is indiscriminate towards prokaryotic and eukaryotic cells, yet a striking difference in bactericidal activity is evident between gram-positive and gram-negative bacteria (16 µg/ml vs 128 µg/ml in *staph aureus* and *serratia marcescens*, respectively). In general, the puromycin analogues retained similar antibiotic activity compared to the native compound, indicating that

both the introduction of the thiophene heterocycle and the azetidine modifications posed no significant disruption to bioactivity. Intriguingly, [Th]puromycin **33** exhibited higher bactericidal activity in gram-negative bacteria compared to the native compound. A potential explanation for this observation could reside in the slight improvement on cellular uptake instigated by the substitution of a polar imidazole for a more hydrophobic thiophene heterocycle. However, a decrease in antibiotic activity was observed upon introduction of the azetidine modification, which was further substantiated by the difluoroazetidine modification on antibiotic **70**. These results are in line with the IC_{50} values measured for the dose-dependent obstruction of *in vitro* GFP translation, where **70** exhibited the lowest activity, providing additional evidence of a structure-activity relationship between the various derivatives.

Table 4.2. Antibacterial activity of puromycin (**12**), [Th]puromycin (**33**), azetidino- and difluoroazetidine analogues **69** and **70**

Bacterial Strains	Differentiation	Tested compounds (μg/ml)			
		12	33	69	70
Staphylococcus aureus ATCC 29213	Gram-positive	16	16	16	32
Bacillus cereus ATCC 14579	Gram-positive	8	8	8	16
Serratia marcescens ATCC 8100	Gram-negative	128	64	128	128
Escherichia coli ATCC 51739	Gram-negative	128	64	128	>128
Escherichia coli ATCC 51739	Gram-negative	64	32	64	64

4.3 Cellular Imaging

4.3.1 HEK293T live-cell experiments

Following the activity experiments, we decided to incubate human embryonic kidney (HEK293T) cells with the fluorescent antibiotics to assess multiple factors: 1) whether the fluorescent properties of the chromophores could be utilized in a cellular setting to label proteins, 2) how the systems change over time, 3) proof of mechanism of action, and 4) If the analogues localize with cellular ribosomes to suggest localization of protein synthesis. It is important to note that the antibiotics must be dosed below the IC_{50} to minimize cell death and morphological changes.

183

Following 30 minute, 1 hr, and 4 hr incubation with 10 μM antibiotic solutions, HEK293T cells were visualized using fluorescence microscopy and the images are shown in figures 4.4-4.7. As expected, puromycin exhibited no fluorescence across all time points measured, and a similar trend was observed in cells treated with **33** and azetidine-modified antibiotic **69**, indicating the "on-off" emission observed in different solvents does not aid in cellular imaging, and intense aqueous emission intensity is necessary. This point is supported by images of HEK293T cells treated with antibiotic **70** incorporating the heterocycle with the highest fluorescence intensity in water. Although treatment after 30 minutes yielded no discernable fluorescence within cells, fluorescence was detected within cells following 2- and 4-hour exposure times. Intriguinigly, while weak fluorescence can be detected across the cytoplasm of the cells, higher-intensity fluorescence was evident in localized "spots" within the cells. We postulated these fluorescent regions were nucleoli, the progenitors of ribosome biogenesis. This phenomenon has literature precedence – puromycin labelling of nucleoli within tissue cultures has been previously observed,[5] and the most intense fluorescence signal within *C. Elegans* treated with click-reactive O-propargyl puromycin was localized within nucleoli.[6] Such behavior may reflect the trafficking or diffusion of puromyclated peptides into the nucleolus,[6] and the fluorescent signal observed with treatment of difluoroazetidine-modified antibiotic **70** supports this observation.
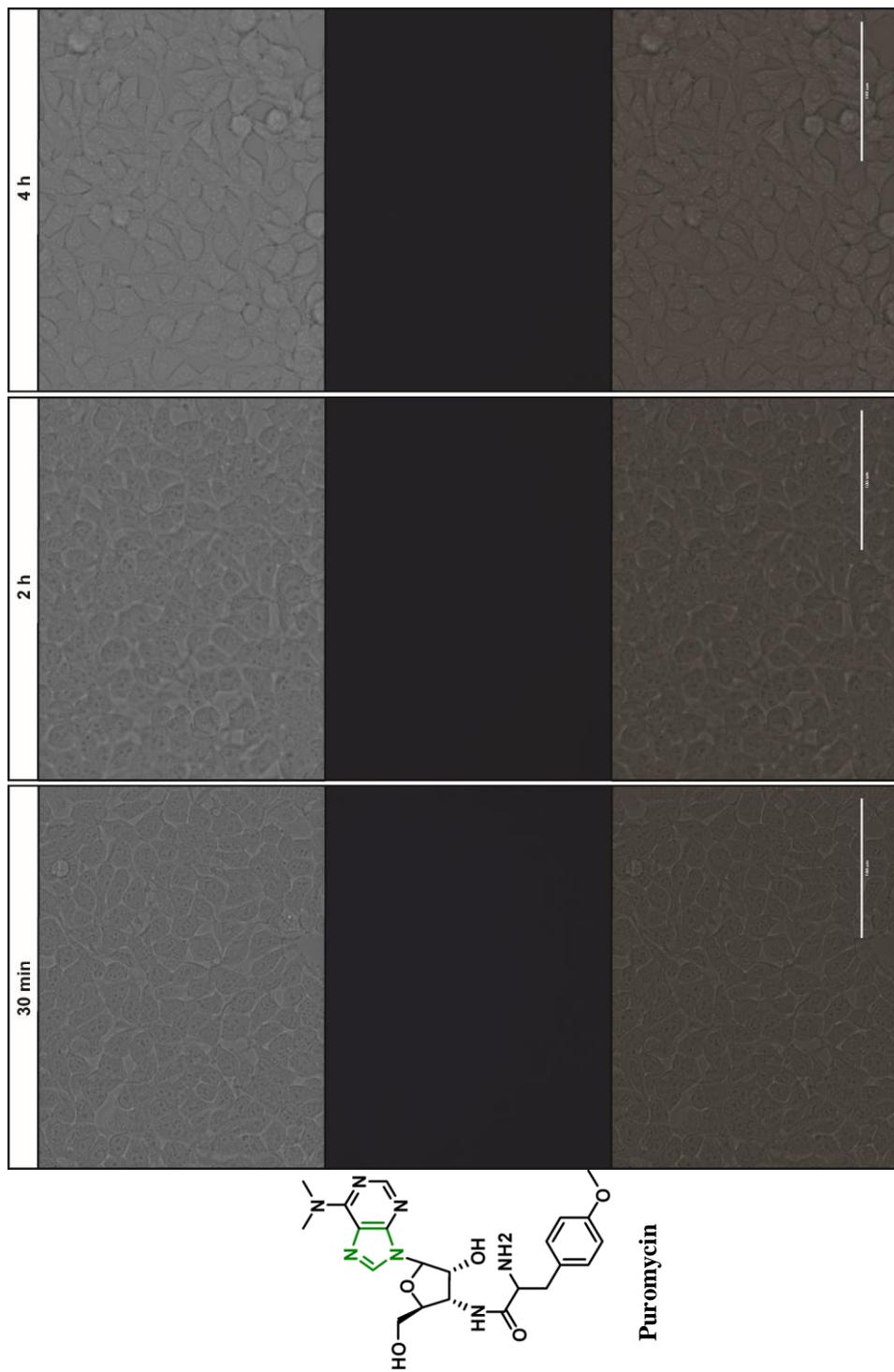
Figure 4.4. Brightfield (top row), DAPI channel (360 nm, middle row) and overlay (bottom row) images taken of HEK293T cells incubated withpuromycin at 30 min (left column), 2 hours (middle column) and 4 hours (right column).

185
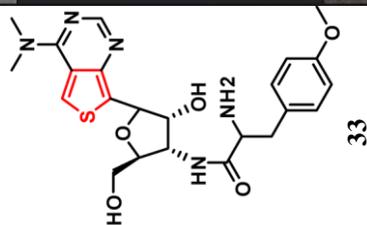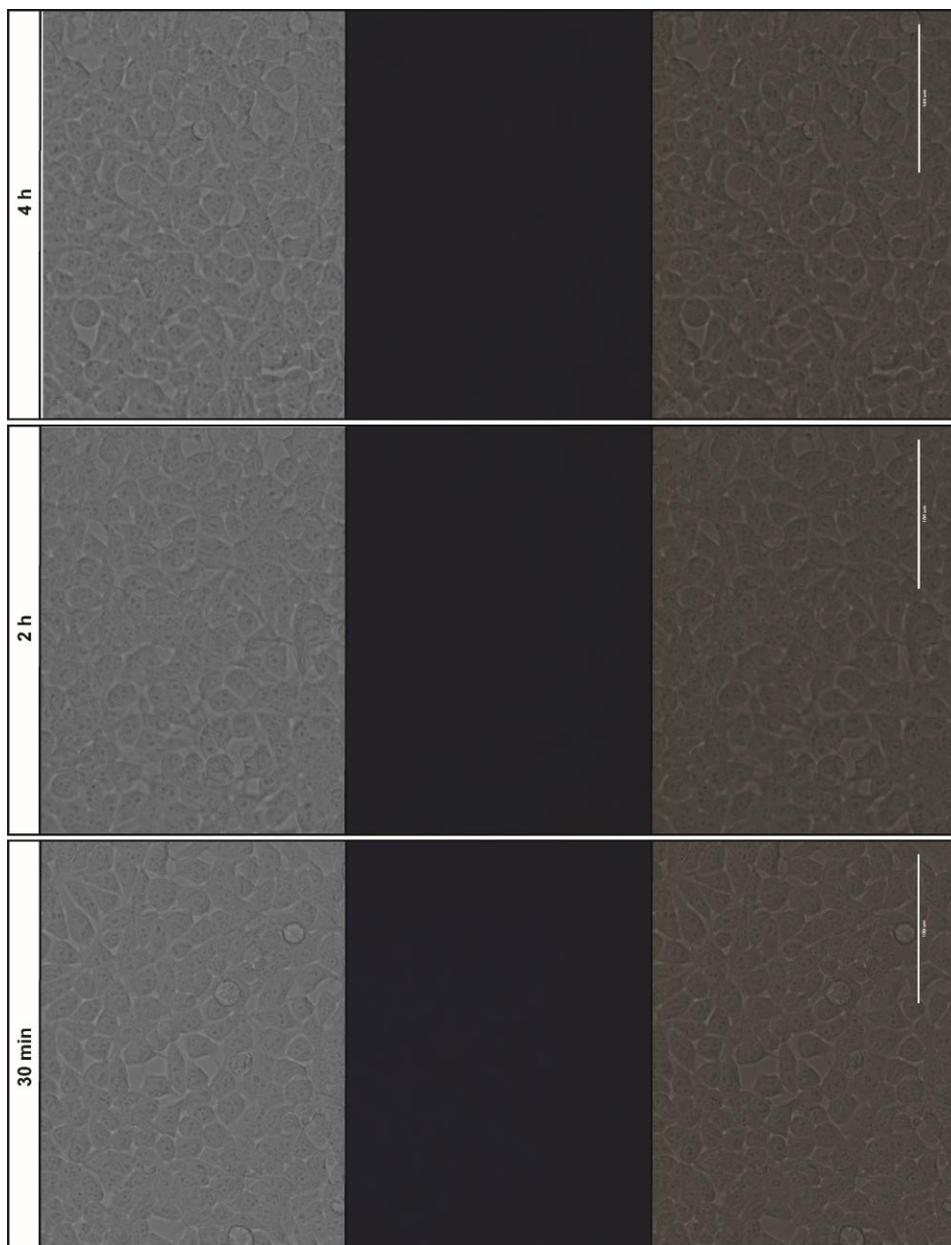
Figure 4.5. Brightfield (top row), DAPI channel (middle row) and overlay (bottom row) images taken of HEK293T cells incubated with [Th]puromycin **33** at 30 min (left column), 2 hours (middle column) and 4 hours (right column).
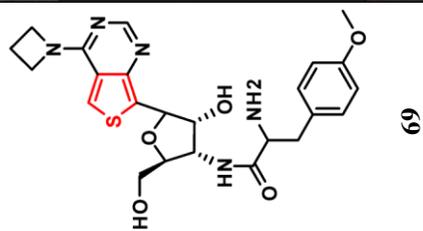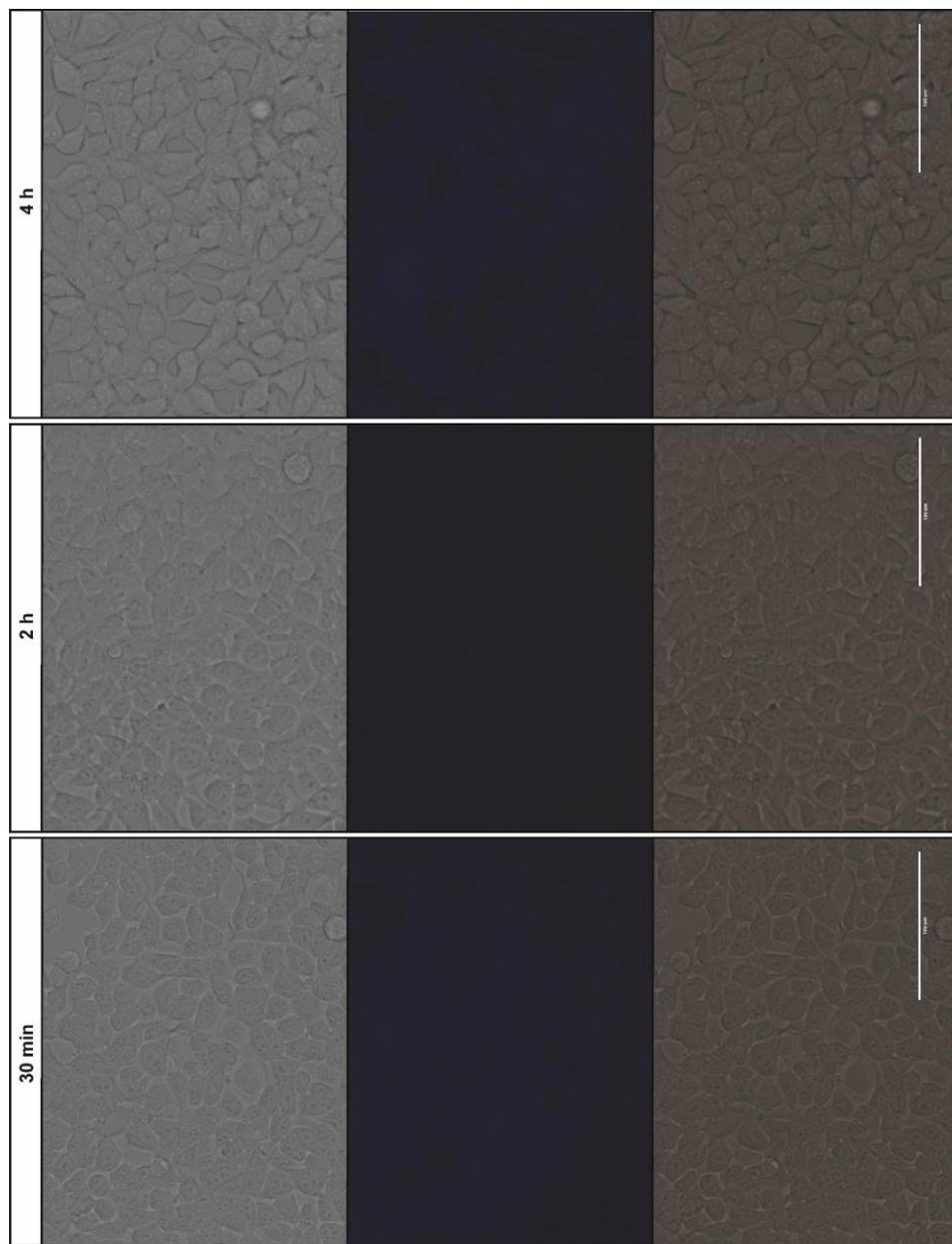
Figure 4.6. Brightfield (top row), DAPI channel (middle row) and overlay (bottom row) images taken of HEK293T cells incubated with azetidine-modified antibiotic **69** at 30 min (left column), 2 hours (middle column) and 4 hours (right column).
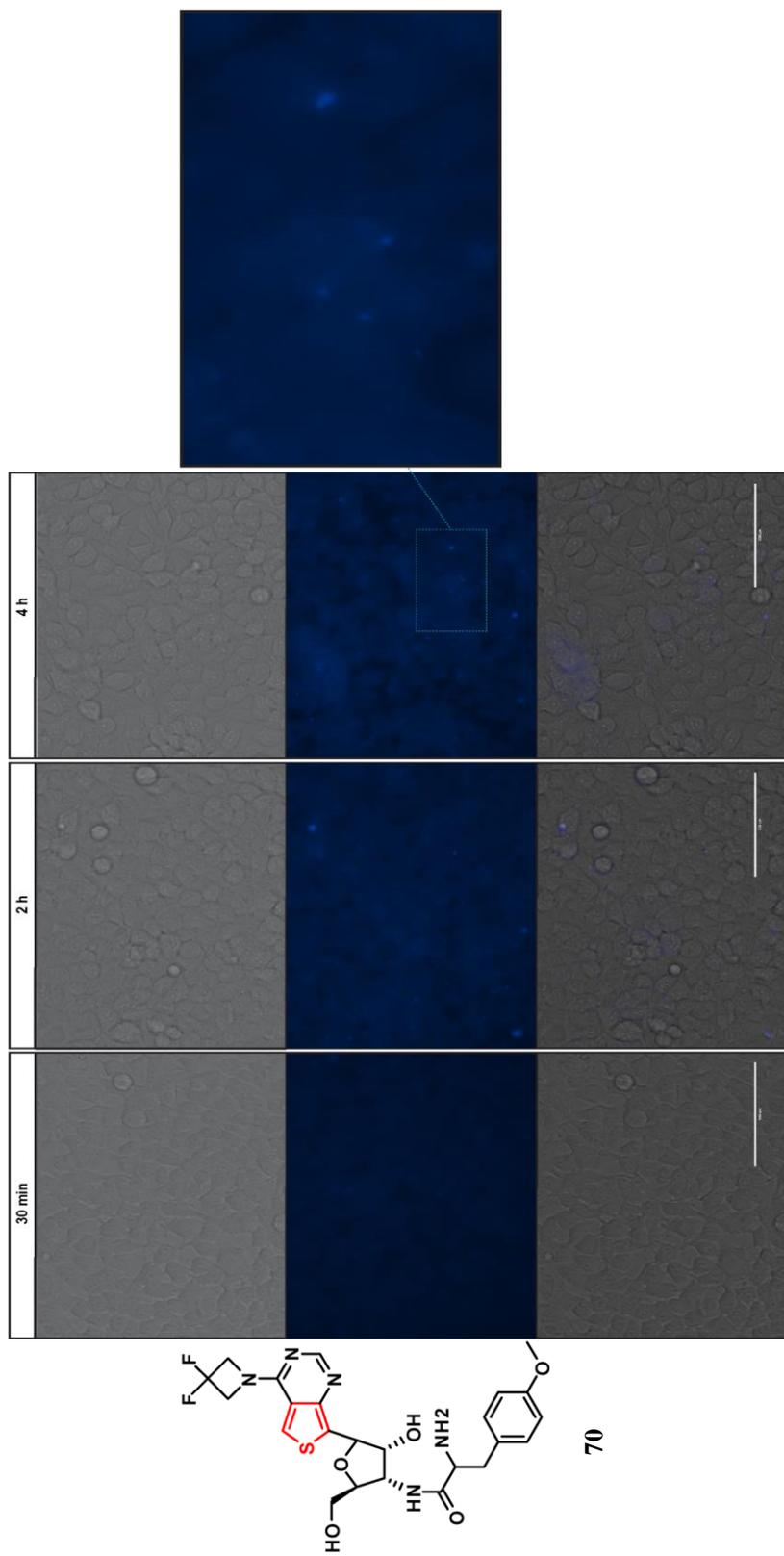
Figure 4.7. Brightfield (top row), DAPI channel (middle row) and overlay (bottom row) images taken of HEK293T cells incubated with difluoroazetidine-modified antibiotic **70** at 30 min (left column), 2 hours (middle column) and 4 hours (right column).

4.3.2 Puromycin Immunodetection

To ascertain mechanistic information about the activity of the fluorescent puromycin derivatives within cells, HEK293T cells were similarly treated with native puromycin or the fluorescent analogues. However, cells were treated in the presence and absence of cycloheximide (0, 5 and 50 µg/ml), a protein synthesis inhibitor that also impedes puromycin incorporation. Following specific treatment times, the cells were harvested, and isolated total protein was subjected to anti-puromycin western blott analysis (figure 4.8).

Western blott of cellular total protein following puromycin treatment showed bands of many different sized proteins that were labelled with puromycin within respective lanes. The signal intensity from labelled peptides drops significantly in cycloheximide treated samples, indicating that the puromycylation of proteins was likely performed through the established mechanism of action. Additionally, this suggests anti-puromycin antibodies were able to recognize the modified puromycin derivatives. The Thpuromycin **33** exhibited the highest signal intensity following primary and secondary antibody binding to puromycylated peptides following native puromycin. The azetidine- and difluoroazetidine- derivatives **69** and **70** exhibited the lowest signal intensity, respectively.
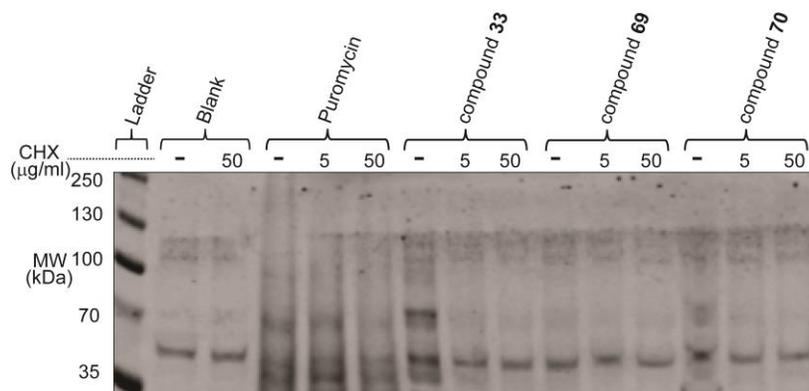
Figure 4.8. Western blott analysis of total protein isolated from HEK293T cells following antibiotic treatment.

4.3.3 Puromycin Immunofluorescence in HEK 293T cells

Puromycin and various analogues are reported to accurately localize at sites of translation at the subcellular level in neruons, and to a lesser extent, other types of mammalian cells due to the metabolic rates of clearance of puromycin-terminated peptides. Given that we had produced fluorescent analogues of puromycin, we opted to determine whether our fluorescent analogues were capable of similar behavior. To analyze this aspect, we performed immunofluorescence experiments on treated HEK293T cells. Following 30 min, 2h, and 4h treatments of either puromycin or any of the modified analogues in the presence/absence of cycloheximide, the cells were fixed and permeabilized to allow for the penetration of an anti-puromycin and anti-ribosoma ntibodies, along with the corresponding fluorescently tagged secondary antibodies. Results are given in figure 4.9-4.13. We found that the anti-puromycin antibodies bound puromycin and the analogues nicely, according to the fluorescence from an AlexaFluor-tagged goat anti-mouse secondary antibody. No fluorescence from anti-puromycin immunofluorescence treatments was detected in the blank, indicating that the incorporation of the antibiotic was necessary for visualization. Native puromycin incorporated into peptides the fastest, with fluorescence from anti-puromycin immunofluorescence present at 30 minutes and intensifying at 2 and 4 hours. Similar behavior was observed in the fluorescent analogues, although their incorporation into cellular proteins appeared slower with very little signal observed at 30 minutes and is likely due to both slow incorporation and less efficient antibody binding as indicated in section 4.2.2. Additionally, the pretreatment of HEK cells with 50 µg/ml resulted in decreased anti-puromycin fluorescence across all time points for every antibiotic testing, which strongly supports the idea that anti-puromycin antibodies are binding truncated peptides (as advertised).

In addition to being treated with anti-puromycin, the same cells were also treated with DRAQ5, a far-red nuclear stain, and HPO-0100, an anti-ribosomal P antigen antibody. Antibody binding was visualized with a Cy3-labelled donkey anti-human secondary antibody. No effect on immunofluorescence was observed with cycloheximide pretreatment. Additionally, the fluorescent signal from the anti-ribosomal antibody appeared both on the periphery of the nucleus and in the cytoplasm.

Co-localization of anti-puromycin and anti-ribosome fluorescence (green and pink fluorescence channels, respectively) appears evident in 2-hour antibiotic treatment times for all antibiotics tested. After 4 hours, however, the anti-puromycin fluorescence is much more disperse throughout the cell. Although diffusion through the cytoplasm may provide a potential explanation, it is known that puromycylated peptides are shuttled to lysosomes for degradation, and these shuttling events provide an avenue for terminated peptides to transverse the crowded cellular environment. Unfortunately, no fluorescence signal was observed in the DAPI channel originating from compound **70**. While anti-puromycin staining indicates that the labelled peptides are present, they likely reside in concentrations too low for fluorescence detection from the small molecule itself.
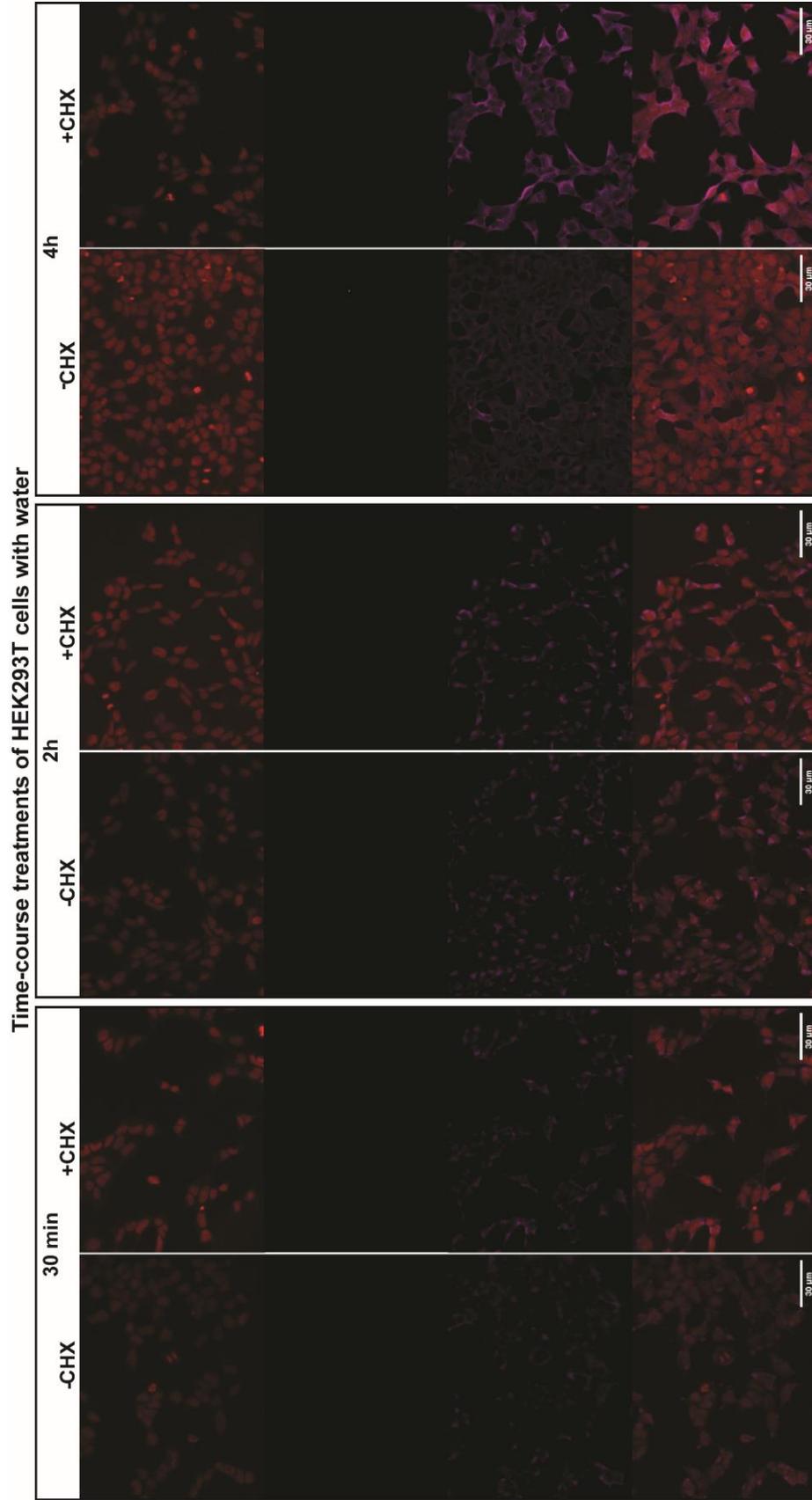
Figure 4.9. DRAQ5 (top row), anti-puromycin antibody/GAM-AF488 (second row), HPO-0100/DAH-Cy3 (third row) and overlay images (bottom row) of fixed HEK293T cells treated with water (negative control) at 30 minutes without cycloheximide (first column) and with cycloheximide (second column), 2 hours without cycloheximide (third column) and with cycloheximide (fourth column) and 4 hours without cycloheximide (fifth column) and with cycloheximide (sixth column).

192
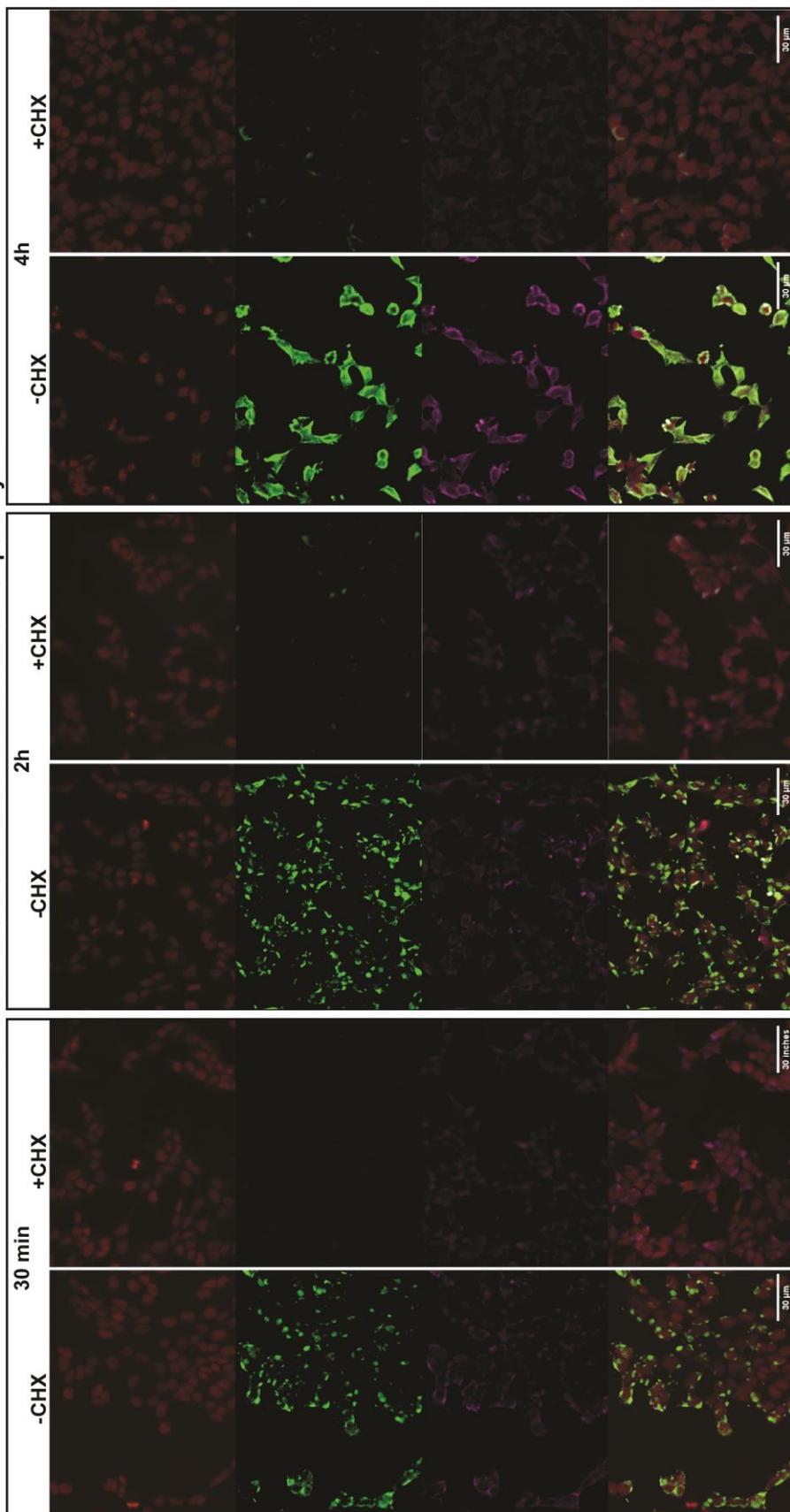
Figure 4.10. DRAQ5 (top row), anti-puromycin antibody/GAM-AF488 (second row), HPO-0100/DAH-Cy3 (third row) and overlay images (bottom row) of fixed HEK293T cells treated with puromycin at 30 minutes without cycloheximide (first column) and with cycloheximide (second column), 2 hours without cycloheximide (third column) and with cycloheximide (fourth column) and 4 hours without cycloheximide (fifth column) and with cycloheximide (sixth column).
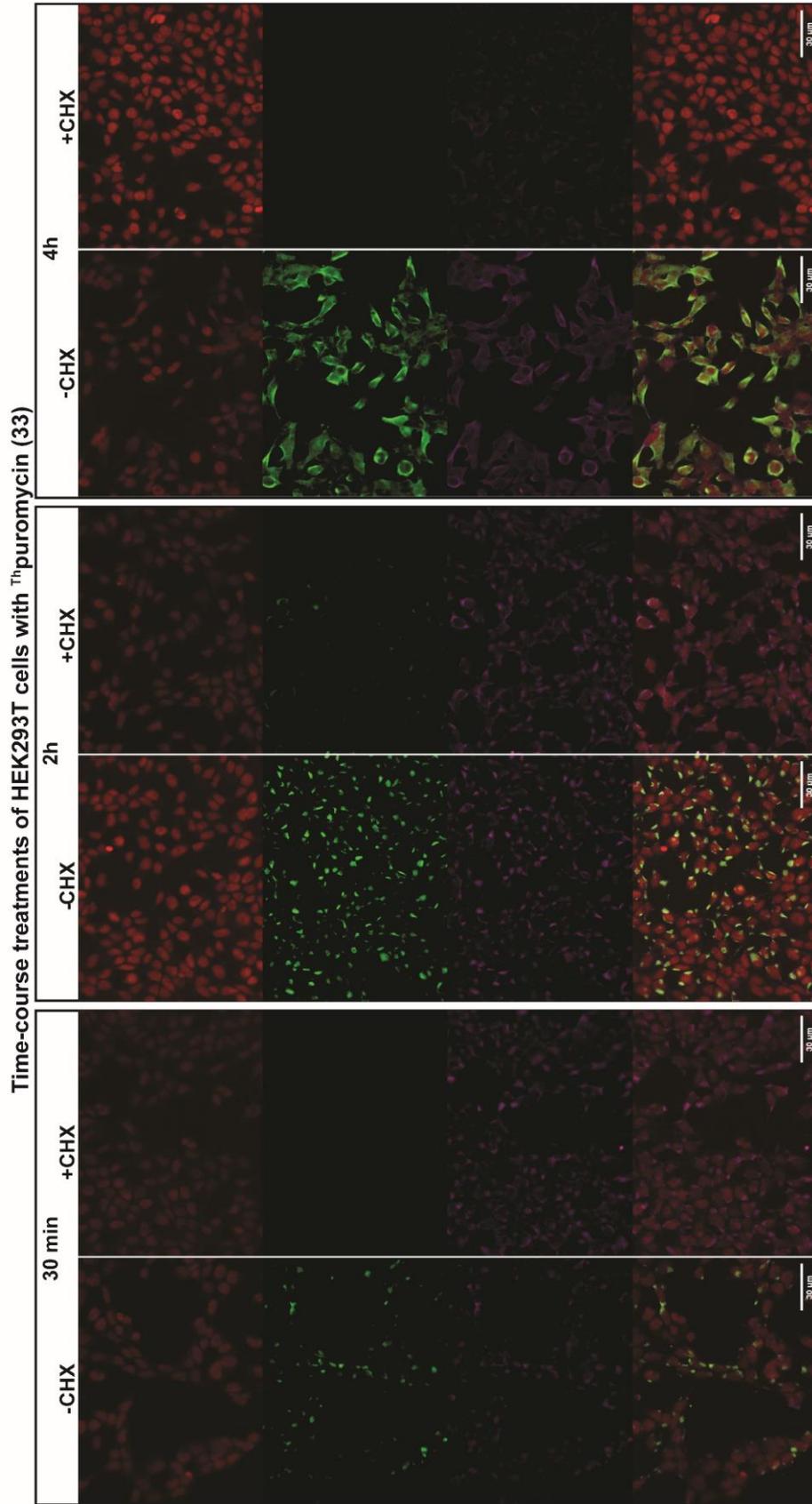
Figure 4.11. DRAQ5 (top row), anti-puromycin antibody/GAM-AF488 (second row), HPO-0100/DAH-Cy3 (third row) and overlay images (bottom row) of fixed HEK293T cells treated with ᵀʰpuromycin **33** at 30 minutes without cycloheximide (first column) and with cycloheximide (second column), 2 hours without cycloheximide (third column) and with cycloheximide (fourth column) and 4 hours without cycloheximide (fifth column) and with cycloheximide (sixth column).

194
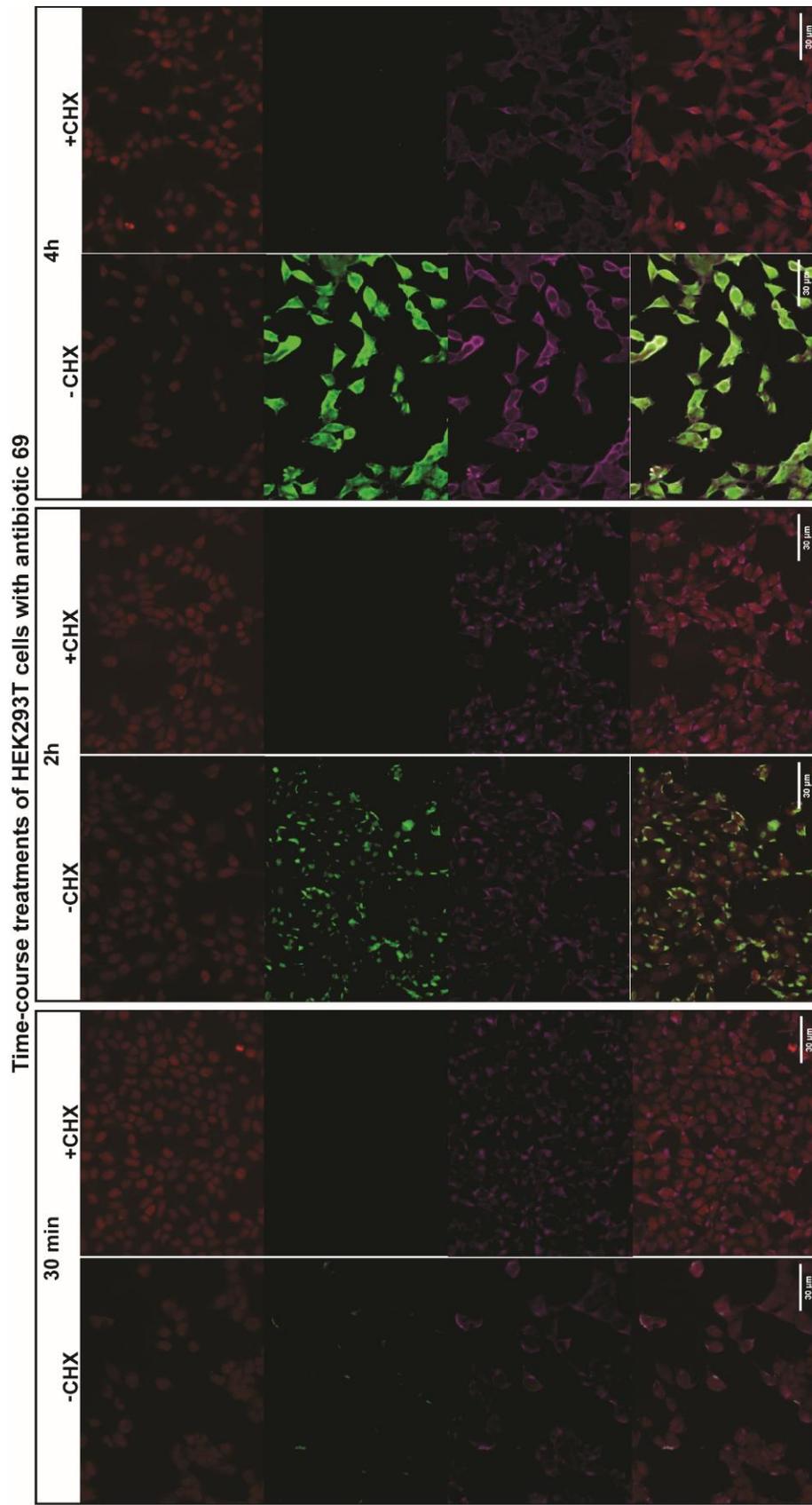
Figure 4.12. DRAQ5 (top row), anti-puromycin antibody/GAM-AF488 (second row), HPO-0100/DAH-Cy3 (third row) and overlay images (bottom row) of fixed HEK293T cells treated with azetidine-modified antibiotic **69** at 30 minutes without cycloheximide (first column) and with cycloheximide (second column), 2 hours without cycloheximide (third column) and with cycloheximide (fourth column) and 4 hours without cycloheximide (fifth column) and with cycloheximide (sixth column).
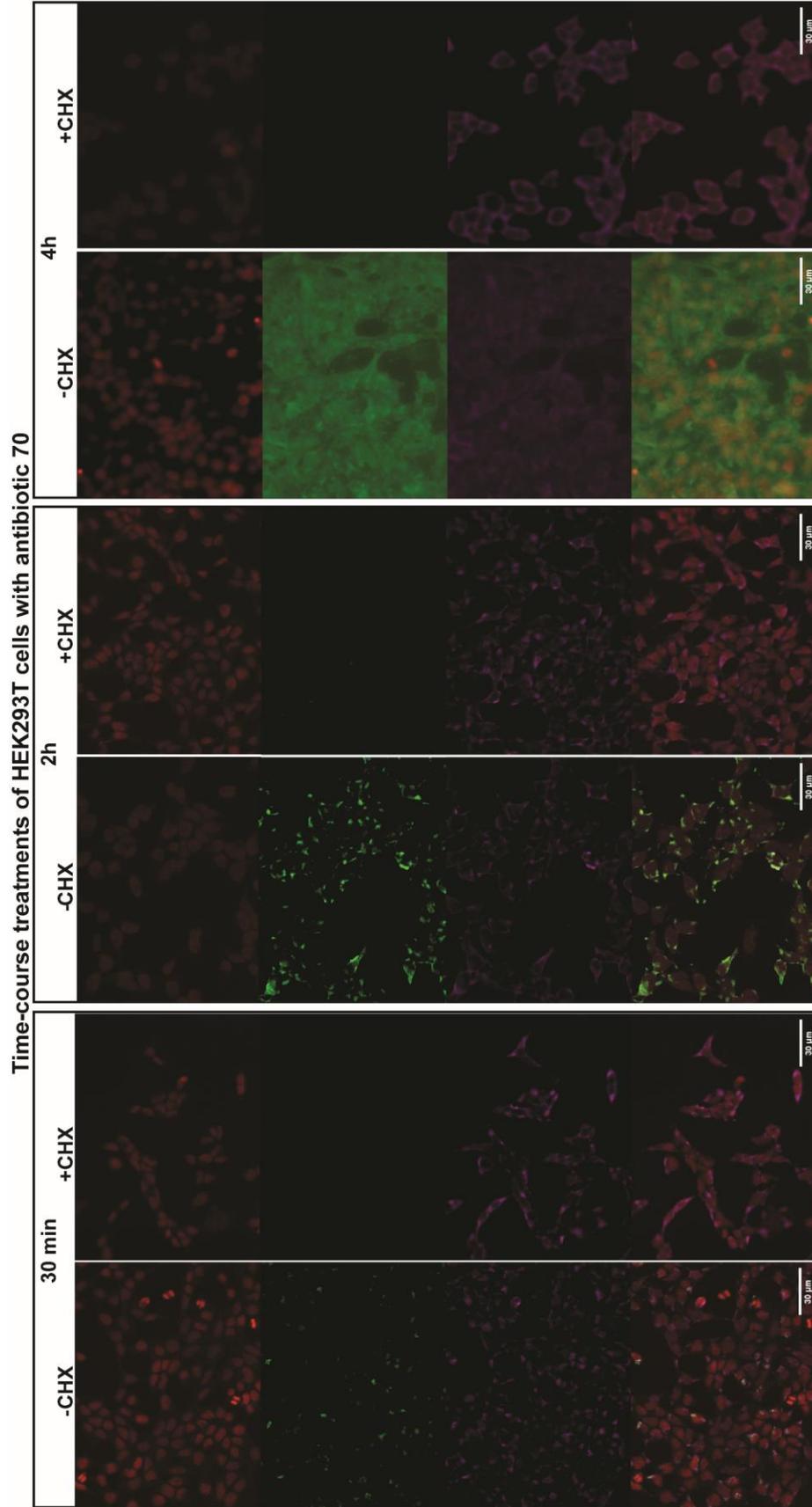
Figure 4.13. DRAQ5 (top row), anti-puromycin antibody/GAM-AF488 (second row), HPO-0100/DAH-Cy3 (third row) and overlay images (bottom row) of fixed HEK293T cells treated with difluoroazetidine-modified antibiotic **70** at 30 minutes without cycloheximide (first column) and with cycloheximide (second column), 2 hours without cycloheximide (third column) and with cycloheximide (fourth column) and 4 hours without cycloheximide (fifth column) and with cycloheximide (sixth column).

196

4.3.4 Puromycin immunofluorescence in mouse hippocampal neurons

After visualizing the puromycin analogues in HEK293T cells, we decided to probe whether we can detect puromycylated peptides in cultured neurons. With the aid of the Patrick lab at UCSD, we took our most promising antibiotic analogue, the difluoroazetidine-modified antibiotic **70**, and incubated neuronal cultures with the compound for 4 hours. Following antibiotic treatment, cells were fixed, permeabilized, and subjected to anti-puromycin/anti-ribosome primary antibodies. After treating with the corresponding secondary antibodies and DRAQ5 nuclear stain, the cells were imaged on a fluorescent microscope.

Anti-puromycin antibody immunofluorescence of neurons treated with difluoroazetidine antibiotic **70** indicated the presence of puromycylated peptides throughout the neuronal cell body, axons, and dendrites (figure 4.14). Anti-ribosomal antibody (HPO-0100) immunofluorescence was observed in the main cell body, although weaker fluorescence was present in the thicker axons. Strikingly, peptides labelled with antibiotic **70** exhibited an intense fluorescent signal within neurons that was absent in HEK293T cells. This result corroborates
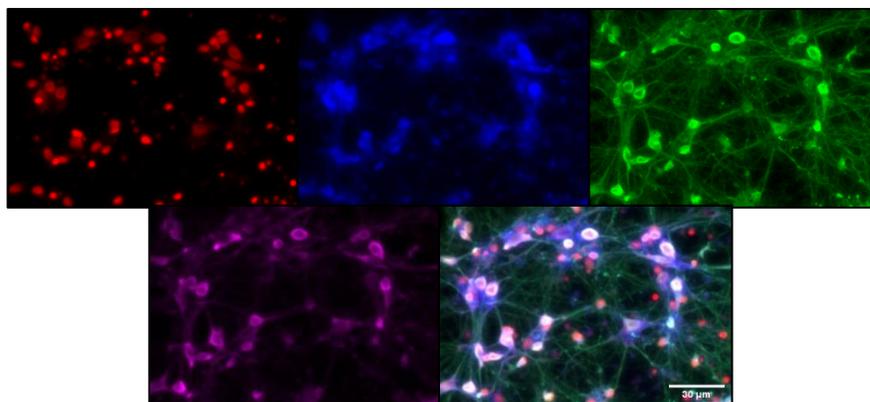
Figure 4.14. Primary mouse hippocampal neuron images containing DRAQ5 (nuclear stain, top left), compound **70** (top middle), anti-puromycin (top right), and anti-ribosome (bottom left). Images are overlayed at the bottom right.

observations reported in the literature regarding rates of diffusion of puromyclated peptides, in which higher rates are observed within mammalian cells as opposed to neuronal compartments.[2,6] However, while the fluorescence from antibiotic **70** colocalized nicely with the anti-puromycin fluorescence, the antibiotic itself was only visible primarily in the soma. Although dim fluorescence is visible in the axon branches, it appears the compound is not bright enough for visualization through fluorescent microscopy alone, and more powerful imaging techniques will be necessary to enhance any signal in smaller regions of interest, such as dendrites. This point is further evidenced by the fact that anti-puromycin immunofluorescence was able to visualize the antibiotic in the smaller neuronal structures such as axons and dendrites.

4.3.5 Live-cell imaging in mouse hippocampal neurons

Lastly, live neurons were treated difluoroazetidine-modified antibiotic **70** to ascertain the ability of compound **70** to fluorescently tag newly synthesized proteins in a non-cytotoxic manner. Following a 4h treatment with antibiotic **70**, neurons were imaged using fluorescent microscopy and the images are shown in figure 4.15. As observed with the neuron immunofluorescence experiments, the antibiotic is primarily visible in the soma. Although we are interested in fluorescence within axons and dendrites, the compound doesn't seem to provide a strong enough signal in neuronal structures with low spatial flexibility.
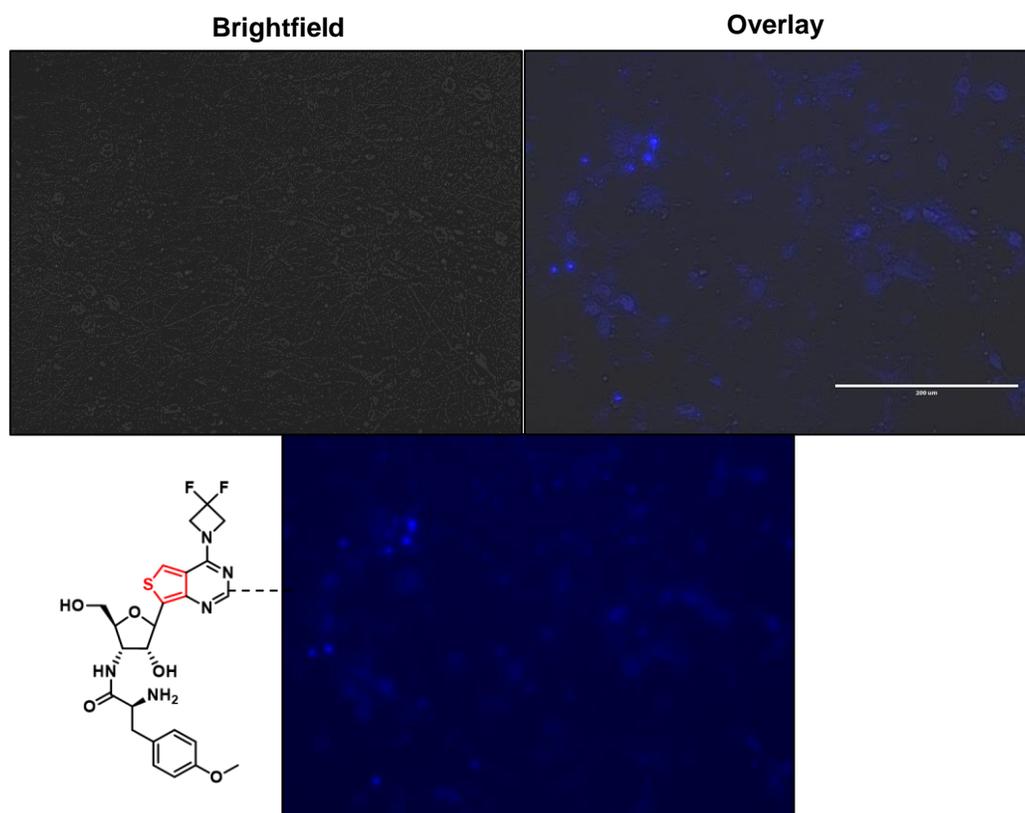
Figure 4.15. Brightfield (top left), DAPI channel (360 nm, bottom), and overlay of brightfield and DAPI channels (top right) of live neurons treated with difluoroazetidine-modified antibiotic **70**.

4.4 Summary

In summary, the thiopheno-modified puromycin analogues display isofunctionality to their native counterpart yet exhibit photophysical properties that provide them with much more emissive properties than their native counterpart in solvents of different polarity. Although only aqeuous conditions were necessary in the context of cellular imaging, the dilfuoroazetidine-modified antibiotic **70** still produced intense fluorescence in HEK293T cells and neurons whilst still undergoing the same mechanism of action as native puromycin given inhibition by cycloheximide, indicating that newly synthesized peptides were fluorescently tagged within live-cells in real-time without the need for extraneous treatments.

4.5 Acknowledgements

4.6 References

1. Robins, M. J., Miles, R. W., Samano, M. C. & Kaspar, R. L. Syntheses of Puromycin from Adenosine and 7-Deazapuromycin from Tubercidin, and Biological Comparisons of the 7-Aza/Deaza Pair1. *J. Org. Chem.* **66**, 8204–8210 (2001).

2. Ge, J. *et al.* Puromycin Analogues Capable of Multiplexed Imaging and Profiling of Protein Synthesis and Dynamics in Live Cells and Neurons. *Angew. Chem. Int. Ed.* **55**, 4933–4937 (2016).

3. Aviner, R. The science of puromycin: From studies of ribosome function to applications in biotechnology. *Comput. Struct. Biotechnol. J.* **18**, 1074–1083 (2020).

4. Nathans, D. & Neidle, A. Structural Requirements for Puromycin Inhibition of Protein Synthesis. *Nature* **197**, 1076 (1963).

5. David, A. *et al.* Nuclear translation visualized by ribosome-bound nascent chain puromycylation. *J. Cell Biol.* **197**, 45–57 (2012).

6. Enam, S. U. *et al.* Puromycin reactivity does not accurately localize translation at the subcellular level. *eLife* **9**, e60303 (2020).

4.7 Methods

4.7.1 Materials

Puromycin dihydrochloride was purchased from Sigma-Aldrich. DMEM, phenol-red free DMEM, and neurobasal medium (21103049) were purchased from Thermo-Fisher. B-27 plus neuronal culture system (A3653401) was purchased from Fisher. Fetal bovine serum (FB-01) was purchased from Omega Scientific. Penicillin:streptomycin (Pen-strep, 10,000 U/ml) was purchased from Gibco. CleanCap® EGFP mRNA (L-7601) was purchased from Trilink, and rabbit reticulocyte cell lysate was purchased from (L4960) was purchased from Promega. 24-well plates (10062-896) were purchased from VWR and TC coverslips (NC9964441) were purchased from Fisher. Poly-D-Lysine was purchased from Neta Scientific. 1X Phosphate-buffered saline (PBS) with calcium and magnesium (21-030-CM) was purchased from Corning, and 4% Paraformaldehyde solution (PFA) in PBS (AAJ19943K2) was purchasesd from Fisher. Triton-X 100 (X100-5ML) was purchased from Sigma-Aldrich. Bovine-serum albumin was purchased from sigma-aldrich. Trysin-EDTA was purchased from Life Technologies. 10X RIPA buffer was purchased from Cell-Signaling Technologies and cOmplete Mini, EDTA-free proteinase inhibitor cocktail tablets were purchased from Sigma-Aldrich. Cycloheximide was purchased from AG scientific. Pierce BCA Protein Assay kits and NuPAGE LDS sample buffer were purchased from Fisher. Bolt 4 to 12%, Bis-Tris, 1.0 mm, Mini Protein Gels were purchased from Fisher. DRAQ5 5 mM solution was purchased from Thermo-Fisher. Anti-puromycin [3RH11] was purchased from Kerafast, and Human Antibody Against Ribosomal P Antigen (HPO-0100) was purchased from Immunovision. Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 was purchased from Fisher, and Donkey anti-human Cy3 was purchased from Jackson. Immunoresearch. Seablock was purchased from Thermo-Fisher.

4.7.2 Photophysical measurements on antibiotics 33, 69, and 70

For detailed protocols on photophysical experiments, see chapter 2.

4.7.3 HEK293T cell culture

HEK293T cells were a generous gift from the lab of Jeffrey D. Esko. Cells were grown under a 5% CO_2 atmosphere in air with 100% relative humidity at 37 °C, using DMEM with 10% v/v fetal bovine serum and 1% Pen-strep as a growth medium. All experiments use PBS with calcium and magnesium unless otherwise stated.

4.7.4 Neuronal cell culture

Mouse hippocampal neurons on were a generous gift from the lab of Gentry Patrick at UCSD. Cells were grown under a 5% CO_2 atmosphere in air with 100% relative humidity at 37 °C, using 2% Gibco B-27 supplements in neurobasal medium. Following seeding in 24-well plates (on TC coverslips), neurons were passaged every 2-3 days by removing half of the growth medium and replacing it with fresh medium. After 14 days *in vitro* (DIV), the cells were subjected to experimentation.

4.7.5 Bacterial cell culture and MIC determination

Compounds were tested using the broth double-dilution method in 96-well plates (Corning). For inoculum preparation, overnight starter cultures incubated at 37 °C were diluted into fresh Luria Broth (LB) and allowed to grow to the exponential phase ($O.D_{600}$ = 0.5-0.7). Cultures were then diluted again in LB to obtain an optical density of 0.008 (OD_{600}). Tested compounds were added to LB to form mother liquors (128 μL of stock solution in 1,122 μL of LB) at a concentration of 256 μg/mL. Next, 100 μL serial double dilutions of the tested compounds (256, 128, 64, 32, 16, 8, and 4 μg/mL) were prepared in LB in flat-bottomed 96-well microplates. An equal volume (100 μL) of bacterial suspension in LB was added to each well for

a final volume of 200 μL and a final OD$_{600}$ of 0.0004. Control wells containing only LB (blanks) or only bacteria in LB were also prepared. After 18 h, MTT (50 μL of a 1 mg/mL solution in H$_2$O) was added to each well followed by additional incubation at 37 °C for 2 h. The MIC value was defined as the lowest concentration of compound in which no bacterial growth was observed. Each drug concentration was tested in triplicate and results were reproduced in two independent experiments.

4.7.6 Dose-dependent inhibition of GFP translation *In Vitro*.

The dose-dependent inhibition of GFP was assessed using a rabbit reticulocyte cell lysate system, RNAsin ribonuclease inhibitor (40 u/ml) and CleanCap eGFP mRNA (trilink). Reactions were performed in Corning Costar black 96-well plates, non-sterile, flat-bottom with a total volume of 25 μl, in the presence of 0, 0.1, 0.5, 1, 2, 4, or 10 μM of puromycin or antibiotics **33**, **69**, or **70**. Reactions were initiated through the addition of 1 μg of mRNA to each reaction solution, and quickly transferred to a pre-warmed TECAN spark plate reader at 30 °C. Fluorescence at 410 nm (**33**, **69**, **70**) and at 535 nm (GFP) were measured by exciting samples at 350 nm or 488 nm, respectively. Measurements were taken every 3 minutes for 200 minutes.

4.7.7 Live-cell imaging of HEK293T cells

HEK293T cells were seeded into sterile 24-well culture plates at a density of 200,000 cells per well and allowed to grow to confluence overnight. The growth medium was aspirated and the cells were then washed with PBS. 200 μl of a 10 μM solution of the corresponding antibiotic (puromycin, **33**, **69**, or **70**) in growth medium was added to the wells. Cells were incubated with the solutions for either 30 minutes, 2 h, or 4 h, at which point the solutions were aspirated off, and the cells were washed with PBS (with calcium and magnesium). The cells were suspended in phenol-red free DMEM and imaged using an EVOS cell imaging system.

4.7.8 HEK293T Immunofluorescence experiments

TC coverslips were added to the wells of a 24-well tissue culture plate, and 500 µl of PDL solution (50 mg/ml) was added. Following incubation for 45 minutes, the PDL solution was removed and the plate was left to sterilize under UV light for 20 minutes. The wells were then washed with PBS and HEK293T cells were seeded at a density of 200,000 cells/well. Cells were allowed to grow to confluence overnight. The growth medium was aspirated, and the cells were washed with PBS. 200 µl of cycloheximide solution (0, 5, or 50 µg/ml) diluted in growth medium was added to the cells. The cells were incubated with the translation inhibitor for 1 hour at 37 °C with a 5% CO_2 atmosphere, then aspirated and washed with PBS. 200 µl of a 10 µM solution of the corresponding blank (water) or antibiotic (puromycin, **33**, **69**, or **70**) in growth medium was added. The cells were incubated with the antibiotic solutions for 30 minutes, 2 h or 4 h at 37 °C with a 5% CO_2 atmosphere, after which the cells were washed with PBS and treated with 500 µl fixation solution (4% PFA, 4% sucrose, and 5 mM $MgCl_2$ in PBS) for 15 minutes. The cells were washed with PBS and permeabilized with 500 µl 0.1% Triton X-100 in PBS for 1 minute, then blocked with 500 µl 1% BSA in PBS for 1 hr at room temperature with shaking. The solution was aspirated and a 500 µl solution of anti-puromycin [3RH11] and HPO0100 (1:1000 and 1:3000 in 0.1% BSA in PBS) was added to the cells and incubated overnight at 4 °C with shaking. The cells were washed with PBS and a 500 µl solution of Goat anti-mouse AlexaFluor 488 and Donkey anti-human Cy3 (1:1000 and 1:500 in 0.1% BSA in PBS) was added. Following incubation for 2 hours at room temperature with shaking, the cells were washed with PBS and treated with a 500 µl solution of 5 µM DRAQ5. The cells were shaken with the solution for 30 minutes at room temperature protected from light. The DRAQ5 solution was removed and the cells were washed with PBS. The coverslips were then extracted and

mounted onto microscope slides using prolong gold antifade mountant. Once dried, the slides were imaged on a Keyence BZ-X700 microscope.

4.7.9 Western Blott immunostaining

HEK293T cells were seeded into 24-well tissue culture plates at a density of 200,000 cells/well and incubated at 37 °C with a 5% CO_2 atmosphere overnight. The growth medium was aspirated, and the cells were washed with PBS. 200 µl of cycloheximide solution (0, 5, or 50 µg/ml) diluted in growth medium was added to the cells. The cells were incubated with the translation inhibitor for 1 hour at 37 °C with a 5% CO_2 atmosphere, then aspirated and washed with PBS. 200 µl of a 10 µM solution of the corresponding blank (water) or antibiotic (puromycin, **33**, **69**, or **70**) in growth medium was added. The cells were incubated with the antibiotic solutions for 30 minutes, 2 h or 4 h at 37 °C with a 5% CO_2 atmosphere, after which the cells were washed with PBS. 100 µl of trypsin-EDTA was added and the cells were trypsinzed for 5 minutes at 37 °C with a 5% CO_2 atmosphere. The lifted cells were diluted with 600 µl of growth medium and transferred to microcentrifuge tubes and centrifuged for 1 minute at 14,000 rpm. The supernatant was discarded, and the cells were washed with PBS and pelleted again using centrifugation. The PBS was removed and the cell pellets were suspended in 100 µl 1X RIPA buffer to lyse the cells. 100 µl 1X RIPA buffer with an EDTA-free proteinase inhibitor cocktail tablet was added and the solutions were centrifuged to remove insoluble cellular debris. The concentration of proteins within the cell lysates was assayed using the BCA assay, and appropriate aliquots of the cell lysates were concentrated via speedvac to obtain 15 µg of total protein. The pellet was then reconstituted in 8 µl of water, and 4 µl of 3:1 NuPAGE LDS sample buffer: β-mercaptoethanol was added to resolubilize everything remaining. The solutions were boiled in a heat block set to 115 °C for 15 minutes. After the samples had cooled, they were

loaded onto 4-12% Bis-Tris gels and ran at 100 V for 2 hours. The gels were removed and washed with water for 5 minutes, then blotted onto PVDF membranes using a power blotter (Thermo). The membrane was washed with water for 5 minutes, then blocked with seablock fish serum for 1 hour. The membrane was then incubated with anti-puromycin [3RH11] (1:1000 in 3% BSA in TBST) overnight at 4 °C. The membrane was washed with TBST buffer for 5 minutes at room temperature, then incubated with alexafluor 488-tagged goat anti-mouse secondary antibody (1:15,000 in 3% BSA in TBST) for 1 hour at room temperature. Lastly, the membrane was washed with TBST for 5 minutes and imaged.

4.7.10 Neuron immunofluorescence experiments

Half of the growth media of neuronal cell cultures was removed and replaced with an equivalent volume of 20 μM solution of difluoroazetidine-modified antiobiotic **70** diluted in growth medium (for an active concentration of 10 μM. The neurons were then incubated with the antibiotic for 4 hours at 37 °C with a 5% CO_2 atmosphere, then washed with growth medium and PBS (quickly) and fixed with fixation solution (4% PFA, 4% sucrose, and 5 mM $MgCl_2$ in PBS) for 15 minutes. The cells were washed with PBS and permeabilized with 500 μl 0.1% Triton X-100 in PBS for 1 minute, then blocked with 500 μl 1% BSA in PBS for 1 hr at room temperature with shaking. The solution was aspirated and a 500 μl solution of anti-puromycin [3RH11] and HPO0100 (1:1000 and 1:3000 in 0.1% BSA in PBS) was added to the cells and incubated overnight at 4 °C with shaking. The cells were washed with PBS and a 500 μl solution of Goat anti-mouse AlexaFluor 488 and Donkey anti-human Cy3 (1:1000 and 1:500 in 0.1% BSA in PBS) was added. Following incubation for 2 hours at room temperature with shaking, the cells were washed with PBS and treated with a 500 μl solution of 5 μM DRAQ5. The cells were shaken with the solution for 30 minutes at room temperature protected from light. The DRAQ5

solution was removed and the cells were washed with PBS. The coverslips were then extracted and mounted onto microscope slides using prolong gold antifade mountant. Once dried, the slides were imaged on a Keyence BZ-X700 microscope.

4.7.11 Neuron live-cell experiments

Half of the growth media of neuronal cell cultures was removed and replaced with an equivalent volume of 60 μM solution of difluoroazetidine-modified antiobiotic **70** diluted in growth medium (for an active concentration of 30 μM). The neurons were incubated with compound **70** for 4 hours at 37 °C with a 5% CO_2 atmosphere, then washed with growth medium and PBS (quickly). The cells were then quickly suspended in phenol-red free DMEM and imaged using an EVOS cell imaging system.

CHAPTER 5 – Reflections and Future Directions

5.1 Reflections

When I initially began my PhD, I had just completed my MS degree and was trying to decide on what I wanted to research from my doctoral degree. For my master's, I had been originally trained as a synthetic chemist, and ended up performing quite a bit of cell culture work. Given that I had already been exposed to various disciplines of chemistry and biology at the time, I wanted a project that could challenge me to truly become a better scientist and scholar yet expand my expertise within the disciplines that I was exposed to and beyond. I think it's safe to say that I got what I asked for (and more…) with this project. It was a very difficult question to ask, in terms of feasibility and practicality – can we actually get a small chromophore to compete on the level of classical biological dyes, and actually *do more*? We essentially wanted to design a fluorophore that behaved just like the native compound, was nearly as bright as some of the best chromophores in molecular biology, that was responsive to the microenvironment, exhibited some form of tunability and could be used in real-time analysis for live-cell protein labelling *in vitro*. Did I mention we wanted it to be non-cytotoxic as well?

Although we didn't hit all the checkmarks off the to-do list, we got very close. For the first time, we were able to tune the photophysical properties of isomorphic fluorescent nucleosides *in a predictive manner*, something that is very difficult to do even with the best computational methods these days. In addition to that, the synthesis of the fluorescent antibiotics was a landmark in terms of synthetic access and variability. Not only can we now produce any puromycin-like skeleton we desire, but the scope extends beyond puromycin. Through the common precusor **64**, not only can we introduce various modifications at the 6-position of the thienopurine (native purine numbering), but we can also utilize other chemical reactions such as

208

the click reaction and Staudinger ligation to introduce novel functionalities at the 3'-position of the ribose ring and produce new purine derivatives. This synthesis, although only used for puromycin analogues in this work, opens the door to producing fluorescent purine-based probes to study various other biological systems, whatever that may be. And let's not forget that the difluoroazetidine-modified antibiotic **70** was able to do what the design entailed – it not only behaved similarly to native puromycin, but it fluorescently tagged newly synthesized proteins in real-time in live cells. Given it's an antibiotic, it will be cytotoxic to some extent, but the fact that it fluoresced while being treated in doses under the IC_{50} value of puromycin was also a major success.

5.2 Project shortcomings and potential solutions

Unfortunately, where we fell short was in the neuronal treatments-sadly, we couldn't visualize the compounds in axons and dendrites, and that was due to several variables. The most crippling one centered around the antibiotics being incompatible with confocal microscopy. The excitation of the probe is performed at 350-360 nm, and while fluorescent microscopy can accommodate this parameter, confocal cannot, with excitations starting at 400 nm. Additionally, subjecting the chromophore to intense light sources used in fluorescence or confocal microscopy begin to photobleach the compound much faster compared to spectrophotometry. Red-shifting the emission, along with the overall brightness of the fluorophore and its photostability, can potentially be improved by examining other azetidine modifications at 3-position of the 4-membered ring. However, according to the Hammett inductive constants of various functional groups vs. the quantum yield of the azetidine-modified fluorophores (figure 2.5), the quantum yield of the thieno[3,4-*d*]pyrimidine and isothiazolo[4,3-*d*]pyrimidine heterocycles increases based on the sum of the azetidine substituents' Hammett inductive constants. Therefore, the only

209

combination left that can produce a total hammett value higher than two fluorine atoms (and therefore produce a hypothetical quantum yield higher than that provided by the difluoroazetidine modification) is that of two nitro groups (σ_1=1.26 vs 1.0, respectively) with a predicted quantum yield of 0.70 in water (figure 5.1). However, 3-dinitroazetidine is not commercially available, and is likely difficult to synthesize.
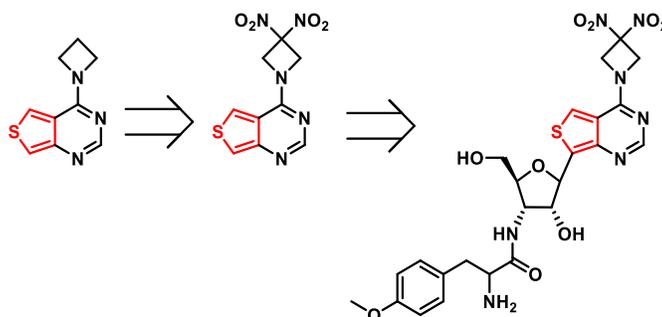
Figure 5.1. Hypothetical dinitroazetidine modification and dinitroazetidine-modified puromycin analog.

5.3 Potential future applications.

5.4 Determination of the mechanism of ribosomal catalysis.

The ribosome's catalysis of peptide bond formation is among one of the most significant processes in biology, yet little of its detailed mechanism is known. Fluorescent analogues of puromycin can aid in deconvoluting this process given that puromycin and its analogues participate in the process of peptide bond formation. However, peptide bond formation is a rapid process—once the corresponding tRNAs have bound their respective sites in the ribosome, peptide-bond formation is initiated and finishes virtually instantly.[1] Approximately 15-20 amino acids are added to a nascent peptide chain in *E. coli* every second, and the rate of elongation is much slower in eukaryotes (~2 amino acids per second).[2] Combined with the speed of peptide bond formation itself, analyzing any fluorescent signal obtained from ribosomal A-site binding or peptide bond formation from a population of ribosomes is especially challenging, and likely

necessitates more complicated photophysical experiments. One such experiment that may provide the information desired is single-molecule spectroscopy on an immobilized substrate, which removes the issues introduced by attempting experimentation on a population of ribosomes and fluorophores. However, this would require immobilized ribosomes, along with a fluorescent analogue of puromycin compatible with the technique, and it remains unclear whether any of the derivatives synthesized are fitting in terms of their brightness or stability to photobleaching.

5.4.1 Cancer cell identification
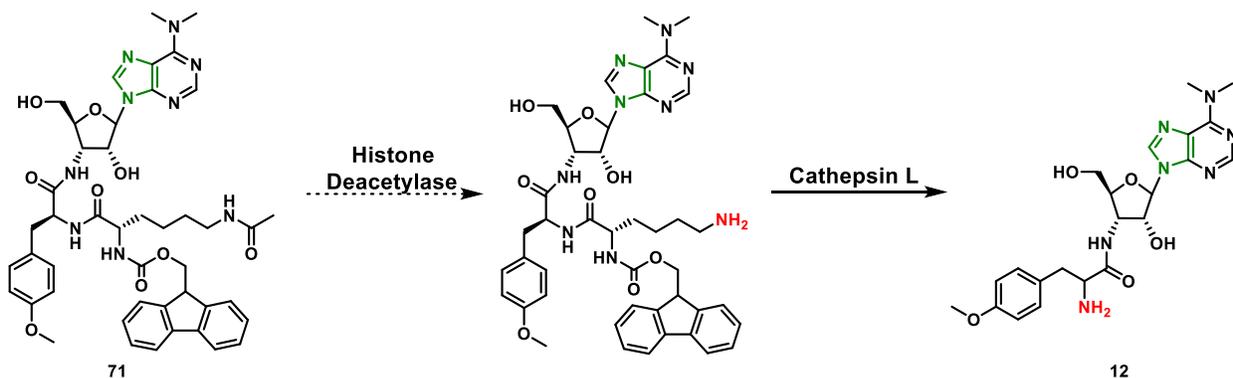
Figure 5.2. Deprotection of puromycin prodrug **71** into puromycin by overexpressed histone deacetylase and cathepsin L in cancerous cells.

Recently, a puromycin prodrug **71** (figure 5.2)[3] was reported to kill cancerous cells whilst remaining benign to healthy cells through utilizing overexpressed histone deacetylases and endogenous cathepsin L to remove the acetyl group and the unacetylated lysine moiety, respectively. In addition to antineoplastic behavior, we postulated that an analogue employing a fluorescent adenosine surrogate would be quenched through potential π-π stacking of the nucleobase and the aromatic ring of the proximal fmoc protecting group. However, removal of these protecting groups in cancerous cells would theoretically separate this quencher from the nucleobase, resulting in increased emission that would not occur in healthy cells.
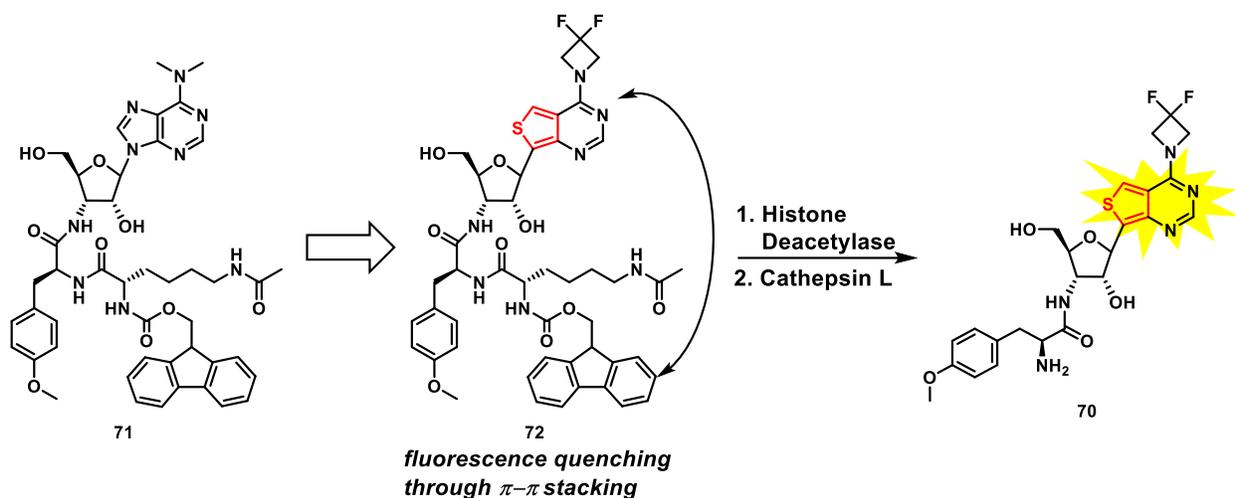
Figure 5.3. Anticancerous puromycin analogue **71** and proposed derivative **72**. Removal of the fmoc protecting group on the lysine moeity of **72** produces antibiotic **70**.

Based on these ideas, prodrug **72** incorporating antibiotic **70** could be employed in this context given the intense fluorescence of the nucleobase. Synthesis of prodrug **72** could proceed through the pathway established for difluoroazetidine-modified antibiotic **70**. However, using dipeptide **77** instead of the "classical" tyrosine analogue Fmoc-O-methyl-L-tyrosine could provide a direct route to prodrug **72** following deprotection. The synthesis of dipeptide **77** was already elucidated in the lab according to (figure 5.4), and the coupling reaction between compound **66** and dipeptide **77** was subsequently performed, however difficulties regarding purification and scale of synthesis are currently hindering further progress.

5.5 Impact of the Fluorescent Antibiotics

In general, the development of the fluorescent antibiotics and more specifically compound **70** has created unique possibilities to probe and exploit protein synthesis within various cell types. Fluorescently tagging proteins for real-time analysis of proteomic spatiotemporal changes is a key parameter necessary for understanding the reasoning behind the translation of specific proteins, what they do, and where they go. Although the analogues in this

work just barely fell short of providing answers to these major questions, we have clearly uncovered a new approach to tackling these problems through tunable fluorescent probe design. The development of both a flexible synthesis and a modular photophysical design for fluorescent puromycin-like antibiotics provides a feasible method of designing, creating, and implementing novel derivatives for applications both within and beyond the domain of protein synthesis.
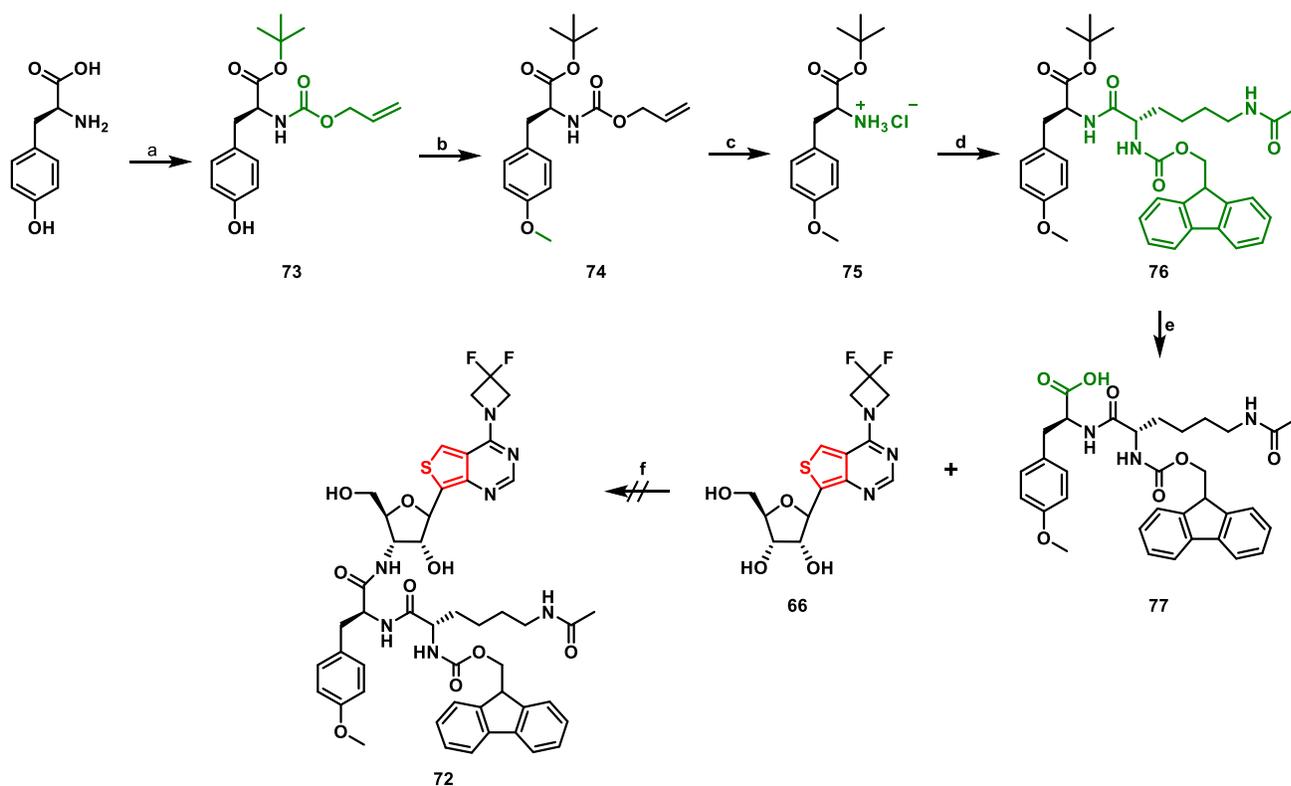
Figure 5.4. Chemical synthesis of precursor dipeptide **77**: (a) i. Allyl chloroformate, 4M NaOH; ii. *t*-butyl 2,2,2-trichloroacetimidate, EtOAc, 22%. (b) K₂CO₃, MeI, DMF, 80%. (c) i. Pd(PPh3)4, phenylsilane, DCM; ii. HCl, Et2O, 66%. (d) Fmoc-Lys(Ac)-OH, EDC, HOBt, DIPEA, DCM, 45%. (e) TFA, DCM, 63%. (f) DIC, HOBt, (*n*-Bu)₃P, THF.

5.6 References

1. Prabhakar, A., Choi, J., Wang, J., Petrov, A. & Puglisi, J. D. Dynamic basis of fidelity and speed in translation: Coordinated multistep mechanisms of elongation and termination. *Protein Sci. Publ. Protein Soc.* **26**, 1352–1362 (2017).

2. Cannarozzi, G. *et al.* A Role for Codon Order in Translation Dynamics. *Cell* **141**, 355–367 (2010).

3. Ueki, N., Lee, S., Sampson, N. S. & Hayman, M. J. Selective cancer targeting with prodrugs activated by histone deacetylases and a tumour-associated protease. *Nat. Commun.* **4**, 2735 (2013).