

# Spectroscopic Analysis and Synthesis of Nucleoside Analogs for Lethal Mutagenesis

by

Katherine J. Silvestre

Submitted to the Department of Chemistry  
in partial fulfillment of the requirements for the degree of

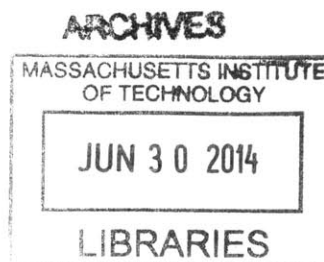
Bachelor of Science in Chemistry and Biology

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2014

© Massachusetts Institute of Technology 2014. All rights reserved.




**Signature redacted**

Author.....

Department of Chemistry  
May 9, 2013

**Signature redacted**

Certified by.....

 John M. Essigmann  
William R. & Betsy P. Leitch Professor in Residence  
Professor of Chemistry, Toxicology, and Biological Engineering  
Thesis Supervisor

**Signature redacted**

Accepted by.....

.....  
Rick L. Danheiser  
A. C. Cope Professor of Chemistry  
Undergraduate Officer, Department of Chemistry

# Spectroscopic Analysis and Synthesis of Nucleoside Analogs for Lethal Mutagenesis

by

Katherine J. Silvestre

Submitted to the Department of Chemistry  
on May 9, 2013, in partial fulfillment of the  
requirements for the degree of  
Bachelor of Science in Chemistry and Biology

## Abstract

The high mutation rate of HIV makes treatment of HIV/AIDS difficult, as the virus can develop resistance to existing therapeutics. A novel antiviral nucleoside analog, 5-Aza-5,6-dihydro-2'-deoxycytidine (KP1212), employs the mechanism of lethal mutagenesis to avoid the problem of drug resistance and use the virus's high mutation rate against it. KP1212 is incorporated into DNA by HIV reverse transcriptase (RT) but mispairs with adenine instead of guanine, increasing the rate of purine transitions. This increased mutation rate can push HIV over its error catastrophe limit, resulting in viral ablation.

To understand how KP1212 induces lethal mutagenesis, VT NMR studies were conducted on KP1212. These experiments revealed that the enol tautomers of KP1212 are more prevalent than the keto tautomers, in contrast to native 2'-deoxycytidine (dC) which is predominately found in the keto-amino tautomer. The enol tautomers are able to base-pair with adenine, causing the increase rate of purine transitions observed in tissue culture and in clinical trials of KP1212. The distribution of KP1212 tautomers measured by VT NMR support tautomerization as the molecular mechanism for KP1212-induced lethal mutagenesis.

VT NMR experiments on a variety of known mutagenic nucleoside analogs found that many of them, including 5-aza-2'-deoxycytidine (decitabine), did not show multiple tautomers and thus had other molecular mechanisms behind their mutagenicity. This suggested that breaking the aromaticity of the nucleobase in KP1212 had a significant effect on the distribution of tautomers. To understand this structural change, the hydrogenated nucleoside analog 5,6-dihydro-2'-deoxycytidine (DHdC) was synthesized via continuous-flow hydrogenation of dC. VT NMR studies show multiple tautomers for DHdC, demonstrating that hydrogenated nucleoside analogs could provide a variety of untested lethal mutagens for further study.

Thesis Supervisor: John M. Essigmann

Title: William R. & Betsy P. Leitch Professor in Residence, Professor of Chemistry, Toxicology, and Biological Engineering

## Acknowledgements

First, I would like to thank my research advisor John Essigmann for allowing me to join his lab as a freshman with no prior research experience. Working in the Essigmann lab over the past three years has taught me a lot about the challenges of research and has had a profound impact on me as a scientist.

I would also like to thank my academic advisor Tim Jamison, for allowing me to work in his lab on the synthesis part of my project. I would never been able to make my dream of this synthesis into reality without his help, and I cannot express my gratitude enough for the opportunity to collaborate with the Jamison lab.

My lab work in the Essigmann lab has been supervised by Dr. Deyu Li. He supported me throughout my years, and was open to the idea of collaborating with the Jamison group when my interests shifted from spectroscopy to synthesis. I thank him for his guidance and encouragement on all aspects of my work.

My collaboration with the Jamison lab would not have been possible without the help of Dr. Ping Zhang, who supervised me on the synthesis part of my project. She invested significant time and energy into teaching me, assisting on a project that was only tangentially related to her own work. It was a privilege to learn from her, and I am incredibly grateful to her for accepting mentorship of me on this project.

Many of the NMR techniques used for my research were taught to me by Dr. Jeffrey H. Simpson from the Department of Chemistry Instrumentation Facility (DCIF). I thank Dr. Simpson and the staff of the DCIF for their technical support over the course of many NMR experiments.

I would like to acknowledge support from the MIT Undergraduate Research Opportunities Program, the National Institutes of Health, and the National Science Foundation, without which this research would not have been possible.

Last but not least, I thank my parents for everything they have done for me. I would not be where I am today without their love and support. For this, I am grateful to them beyond words.

## **Contents**

<b>1. Introduction to Lethal Mutagenesis</b>	<b>7</b>
1.1 Modified Nucleosides for the Treatment of HIV/AIDS	7
1.2 Lethal Mutagenesis as an Antiviral Strategy	8
1.3 KP1212: a Novel Lethal Mutagen for the Treatment of HIV/AIDS	9
<b>2. NMR Spectroscopy of KP1212</b>	<b>11</b>
2.1 Introduction	11
2.2 Experimental	11
2.3 Results and Discussion	13
<b>3. Spectroscopic Exploration of Mutagenic Nucleoside Analogs</b>	<b>17</b>
3.1 Introduction	17
3.2 Experimental	17
3.3 Results and Discussion	18
<b>4. Synthesis and Spectroscopic Analysis of Dihydroneucleosides</b>	<b>24</b>
4.1 Introduction	24
4.2 Experimental	24
4.3 Results and Discussion	25
<b>5. Conclusions and Future Directions</b>	<b>31</b>
<b>References</b>	<b>34</b>
<b>Appendix</b>	<b>36</b>

## List of Figures

1.1	Examples of nucleoside reverse-transcriptase inhibitors (NRTIs).	8
1.2	Examples of antiviral drugs believed to act through lethal mutagenesis.	9
1.3	KP1212, its prodrug KP1461, and its triphosphate form.	10
2.1	VT <sup>1</sup> H NMR spectra of KP1212 in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	12
2.2	<sup>1</sup> H NMR spectrum of KP1212 in DMF-d <sub>7</sub> at -50 °C.	13
2.3	Possible tautomers of KP1212.	13
2.4	2'-Deoxycytidine VT <sup>1</sup> H NMR spectra from 20 °C to -60 °C (top to bottom) in DMF-d <sub>7</sub> .	14
2.5	Reconstructed NMR spectra of the exchangeable nucleobase protons of KP1212 in DMF-d <sub>7</sub> at -50 °C and peak assignment based on their chemical shift.	15
2.6	Matrix equation of peak integrals, system of linear equations, and distribution of tautomers for KP1212 in DMF-d <sub>7</sub> at -50 °C.	15
2.7	Examples of KP1212 tautomers base-paired with the purine bases guanine and adenine.	16
3.1	Modified nucleoside analogs studied via VT NMR.	17
3.2	VT <sup>1</sup> H NMR spectra of 5-OH-dC in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	19
3.3	Possible tautomers of 5-OH-dC.	19
3.4	5-OH-dC tautomers base-paired with the purine bases guanine and adenine.	20
3.5	8-oxo-dG conformers base-paired with the bases cytosine and adenine.	21
3.6	Ribavirin rotamers base-paired with the pyrimidine bases cytosine and uracil.	21
3.7	5-X-dU base-paired with the purine bases guanine and adenine.	22
3.8	Chemical reaction of decitabine.	22
3.9	Decitabine base-paired with the bases guanine and cytosine.	23
4.1	Hydrogenation of dC.	26
4.2	Continuous-flow set-up for the hydrogenation of dC.	26
4.3	VT <sup>1</sup> H NMR spectra of DHdC in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	28
4.4	Hydrogenation of dU.	29
4.5	VT <sup>1</sup> H NMR spectra of DHdU in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	29
4.6	Possible hydrogenated purine nucleosides, DHdA and DHdG.	30
5.1	Conversion of the phosphoramidate prodrug sofosbuvir to the monophosphate uridine derivative.	33
5.2	Proposed synthesis of the phosphoramidate prodrug derivative of KP1212, based on the synthesis of sofosbuvir.	33
A.1	VT <sup>1</sup> H NMR spectra of 2'-deoxycytidine in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	37
A.2	VT <sup>1</sup> H NMR spectra of 8-oxo-dG in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	39
A.3	VT <sup>1</sup> H NMR spectra of ribavirin in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	40
A.4	VT <sup>1</sup> H NMR spectra of 5-F-dU in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	41
A.5	VT <sup>1</sup> H NMR spectra of 5-Cl-dU in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	43
A.6	VT <sup>1</sup> H NMR spectra of 5-Br-dU in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	44
A.7	VT <sup>1</sup> H NMR spectra of 5-I-dU in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	45
A.8	VT <sup>1</sup> H NMR spectra of 5-aza-dC in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	46
A.9	VT <sup>1</sup> H NMR spectra of 2'-deoxyuridine in DMF-d <sub>7</sub> from 20 °C to -60 °C (top to bottom).	48

## List of Tables

A.1	<sup>1</sup> H chemical shift assignments of KP1212 in DMF-d <sub>7</sub> at 20 °C.	36
A.2	<sup>13</sup> C chemical shift assignments of KP1212 in DMF-d <sub>7</sub> at 20 °C.	36
A.3	<sup>1</sup> H chemical shift assignments of 2'-deoxycytidine in DMF-d <sub>7</sub> at 20 °C.	37
A.4	<sup>13</sup> C chemical shift assignments of 2'-deoxycytidine in DMF-d <sub>7</sub> at 20 °C.	38
A.5	<sup>1</sup> H chemical shift assignments of 5-OH-dC in DMF-d <sub>7</sub> at 20 °C.	38
A.6	<sup>13</sup> C chemical shift assignments of 5-OH-dC in DMF-d <sub>7</sub> at 20 °C.	38
A.7	<sup>1</sup> H chemical shift assignments of 8-oxo-dG in DMF-d <sub>7</sub> at 20 °C.	39
A.8	<sup>13</sup> C chemical shift assignments of 8-oxo-dG in DMF-d <sub>7</sub> at 20 °C.	39
A.9	<sup>1</sup> H chemical shift assignments of ribavirin in DMF-d <sub>7</sub> at 20 °C.	40
A.10	<sup>13</sup> C chemical shift assignments of ribavirin in DMF-d <sub>7</sub> at 20 °C.	41
A.11	<sup>1</sup> H chemical shift assignments of 5-F-dU in DMF-d <sub>7</sub> at 20 °C.	42
A.12	<sup>13</sup> C chemical shift assignments of 5-F-dU in DMF-d <sub>7</sub> at 20 °C.	42
A.13	<sup>1</sup> H chemical shift assignments of 5-Cl-dU in DMF-d <sub>7</sub> at 20 °C.	43
A.14	<sup>1</sup> H chemical shift assignments of 5-Br-dU in DMF-d <sub>7</sub> at 20 °C.	44
A.15	<sup>13</sup> C chemical shift assignments of 5-Br-dU in DMF-d <sub>7</sub> at 20 °C.	44
A.16	<sup>1</sup> H chemical shift assignments of 5-I-dU in DMF-d <sub>7</sub> at 20 °C.	45
A.17	<sup>1</sup> H chemical shift assignments of 5-aza-dC in DMF-d <sub>7</sub> at 20 °C.	46
A.18	<sup>13</sup> C chemical shift assignments of 5-aza-dC in DMF-d <sub>7</sub> at 20 °C.	47
A.19	<sup>1</sup> H chemical shift assignments of DHdC in DMF-d <sub>7</sub> at 20 °C.	47
A.20	<sup>13</sup> C chemical shift assignments of DHdC in DMF-d <sub>7</sub> at 20 °C.	47
A.21	<sup>1</sup> H chemical shift assignments of DHdU in DMF-d <sub>7</sub> at 20 °C.	48
A.22	<sup>1</sup> H chemical shift assignments of 2'-deoxyuridine in DMF-d <sub>7</sub> at 20 °C.	49
A.23	<sup>13</sup> C chemical shift assignments of 2'-deoxyuridine in DMF-d <sub>7</sub> at 20 °C.	49

## Chapter 1

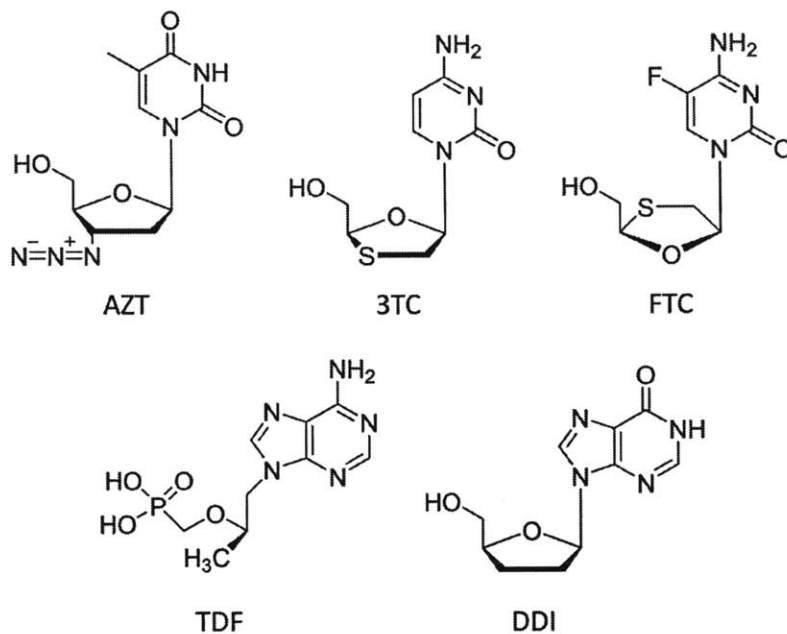
### Introduction to Lethal Mutagenesis

#### 1.1 Modified Nucleosides for the Treatment of HIV/AIDS

Acquired immunodeficiency syndrome (AIDS), the disease caused by human immunodeficiency virus (HIV), has become a serious pandemic since it was first reported in 1981. In 2012, 35.3 million people worldwide were living with HIV, and 1.6 million people died from AIDS. Despite efforts to control the spread of HIV through safe sex education and needle-exchange programs, 2.3 million people were newly infected with HIV in 2012.<sup>1</sup> Antiretroviral treatment (ART) uses a combination of multiple antiretroviral (ARV) drugs to suppress the HIV virus in an infected person and prevent progression into AIDS. While ART can give HIV-positive individuals a near-normal life expectancy, there remains no vaccine or cure for HIV/AIDS.<sup>2</sup>

First-line ART typically consists of two nucleoside reverse-transcriptase inhibitors (NRTIs) and one non-nucleoside reverse-transcriptase inhibitors (NNRTIs), while second-line ART uses a protease inhibitor in addition to NRTIs. NRTIs are analogs of natural deoxynucleosides used to synthesize viral DNA by reverse transcriptase (RT).<sup>3</sup> In the cell, NRTIs are phosphorylated by cellular kinases to the corresponding triphosphate, which is an irreversible inhibitor of RT. Once the NRTI is incorporated into growing viral DNA, the RT cannot continue DNA synthesis, as NRTIs lack the 3'-hydroxyl group required to form the next 5'-3'-phosphodiester bond. This results in chain-termination, preventing viral replication.<sup>4</sup> Commonly used NRTIs are zidovudine (azidothymidine, AZT), a thymidine-analog; lamivudine (2',3'-dideoxy-3'-thiacytidine, 3TC) and emtricitabine (FTC), cytidine-analogs; tenofovir (tenofovir disoproxil fumarate, TDF), an adenine-analog; and didanosine (2',3'-dideoxyinosine, DDI) (Figure 1.1).<sup>3</sup> Though ART can successfully

suppress HIV in many patients, the high mutation rate of HIV makes drug resistance a significant problem.



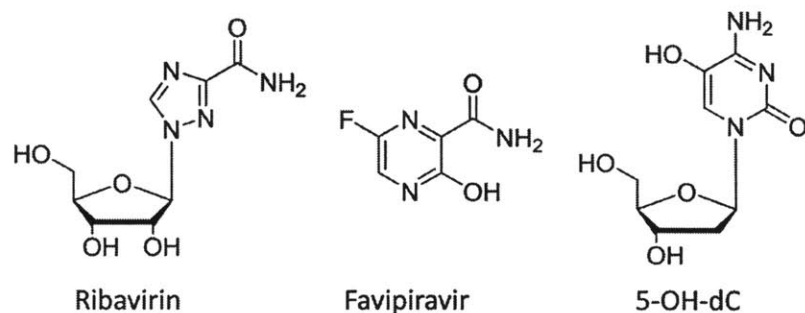
**Figure 1.1** Examples of nucleoside reverse-transcriptase inhibitors (NRTIs).

## 1.2 Lethal Mutagenesis as an Antiviral Strategy

While the high mutation rate of HIV presents a problem to traditional therapeutic approaches, it makes the mechanism of lethal mutagenesis a possibility. Lethal mutagenesis occurs when the mutation rate in DNA replication is increased above the error catastrophe limit. At this point, the load of deleterious mutations is too high and the population is ablated.<sup>5</sup> Due to a variety of proofreading and DNA repair mechanisms, the human genome has a relatively low mutation rate and is thus less susceptible to lethal mutagenesis. On the other hand, viruses like HIV with a high mutation rate are already near the error catastrophe limit, and a slight increase in the mutation rate can cause population collapse.<sup>6</sup> For a lethal mutagen to be a successful antiviral, it must affect the viral genome while sparing the host. Lethal mutagenesis has been proposed as the therapeutic mechanism for multiple antiviral drugs, including ribavirin against hepatitis C virus (HCV)<sup>7-10</sup>,



favipiravir (T-705) against influenza A viruses<sup>11-12</sup>, and 5-hydroxy-2'-deoxycytidine (5-OH-dC) against HIV<sup>13</sup>(Figure 1.2).



**Figure 1.2** Examples of antiviral drugs believed to act through lethal mutagenesis.

### 1.3 KP1212: a Novel Lethal Mutagen for the Treatment of HIV/AIDS

5-Aza-5,6-dihydro-2'-deoxycytidine (KP1212) is a novel nucleoside analog designed to increase the mutation rate of HIV and cause lethal mutagenesis (Figure 1.3). Unlike existing NRTIs, KP1212 is non-chain-terminating, as the 2'-deoxyribose portion of the nucleoside is unmodified. Instead, the nucleobase portion is altered with a nitrogen heteroatom at the 5-position and saturation of the 5,6-double bond. These changes allow KP1212 to alternately base-pair with guanine and adenine via tautomerization, inducing purine transitions. Serial passaging with KP1212 has been shown to increase the viral mutation rate of HIV, causing viral ablation.<sup>14</sup> Once converted into the corresponding nucleotide triphosphate, KP1212 is a substrate for HIV-1 RT and is incorporated into the growing DNA. *In vitro* experiments show that HIV-1 RT and human mitochondrial DNA polymerase (Pol  $\gamma$ ) incorporate KP1212 with significant efficiency, while mouse DNA polymerase  $\beta$  (Pol  $\beta$ ) does not recognize KP1212 as a substrate.<sup>15</sup> However, *in vivo* assays demonstrate that KP1212 has a low genotoxicity profile, is not toxic to mitochondria, and does not inhibit mitochondrial DNA synthesis.<sup>14</sup> These results suggest that KP1212 could provide a novel therapeutic approach to treating HIV with low genomic and mitochondrial toxicity.



## Chapter 2

### NMR Spectroscopy of KP1212

The work presented in this chapter is adapted from the following manuscript:

Li, D.\*; Fedeles, B. I.\*; Singh, V.\*; Peng, C. S.\*; Silvestre, K. J.; Simi, A. K.; Simpson, J. H.; Tokmakoff, A.; Essigmann, J. M. Tautomerism provides a molecular explanation for the mutagenic properties of the anti-HIV nucleoside analog 5-aza-5,6-dihydro-2'-deoxycytidine. [submitted].

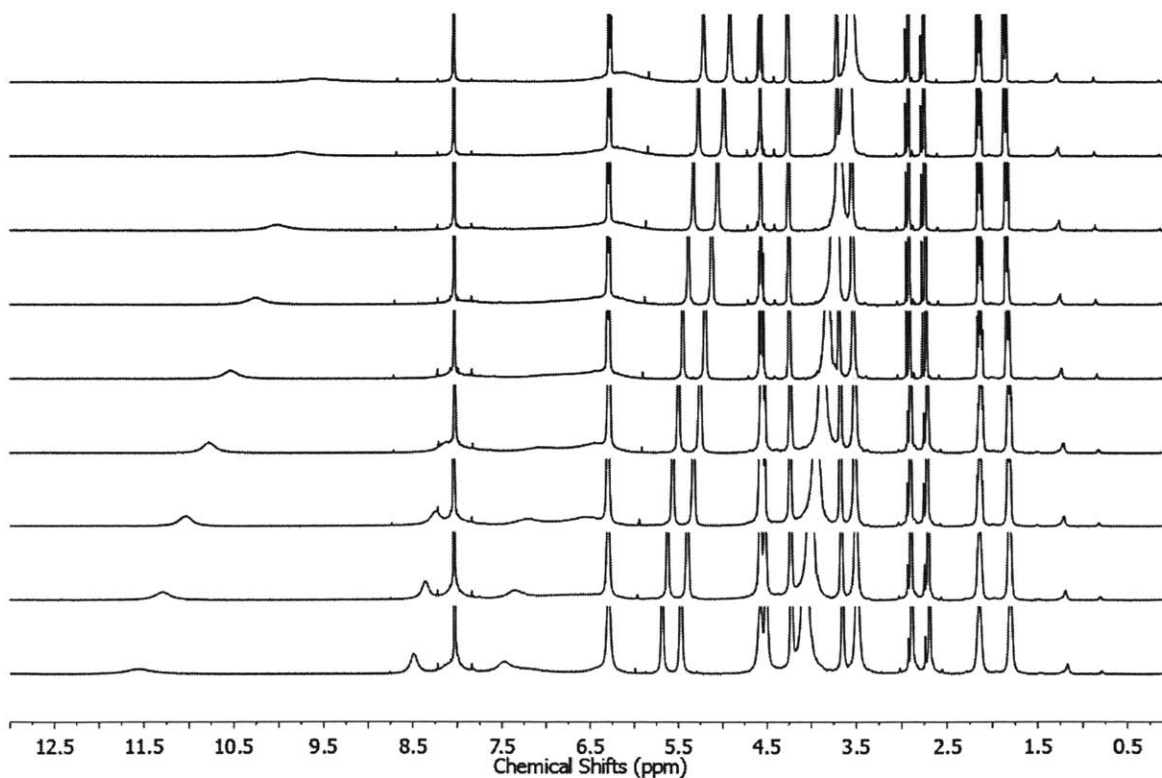
#### 2.1 Introduction

The purpose of the work described here was to determine if the mutagenic properties of KP1212 seen in biological systems could be attributed to the ability of the compound to adopt multiple tautomers in solution. Though tautomerism of KP1212 has been suggested to be the mechanism of its clinical activity, there are no data that directly support this hypothesis. Nuclear magnetic resonance (NMR) spectroscopic methods were used to identify and quantify the array of tautomers exhibited by KP1212. Previous work has demonstrated the ability of NMR spectroscopy to distinguish between different tautomers of small molecules, including nucleosides.<sup>17,18</sup> In proton NMR, the resonant frequency of a proton is determined by its chemical environment. As a result, different tautomers show different chemical shifts for the protons involved in tautomerization. In addition, the proportions of different tautomers in solution can be calculated using integration values for the active protons. NMR spectroscopy is thus uniquely suited to identify and quantify multiple tautomers of KP1212.

#### 2.2 Experimental

The NMR experiments on KP1212 were carried out on a Varian Inova-500 NMR spectrometer. DMF-d<sub>7</sub> was selected as the solvent because it afforded good solubility of the compound, was aprotic to prevent deuterium exchange with the active protons, and had a low freezing point (ca. -61 °C) to permit low temperature NMR experiments. In order to identify the active protons involved in

tautomerization of KP1212, a full chemical assignment was completed using 1D ( $^1\text{H}$ ,  $^{13}\text{C}$ ) and 2D (gCOSY, HSQC, gHMBC) NMR experiments. These spectra provided a definitive assignment for all proton and carbon atoms in KP1212, aside from the active protons on the nucleobase. Once all other protons were accounted for, variable temperature (VT) NMR was used to study the active nucleobase protons (Figure 2.1). Low temperature NMR was used to slow the rate of tautomerization and observe distinct tautomers on the NMR time scale. At 20 °C, the exchangeable nucleobase protons gave rise to three broad peaks. When the temperature was decreased to -50 °C, these three broad peaks divided into six distinct peaks (Figure 2.2).



**Figure 2.1** VT  $^1\text{H}$  NMR spectra of KP1212 in  $\text{DMF-d}_7$ , from 20 °C to -60 °C (top to bottom).

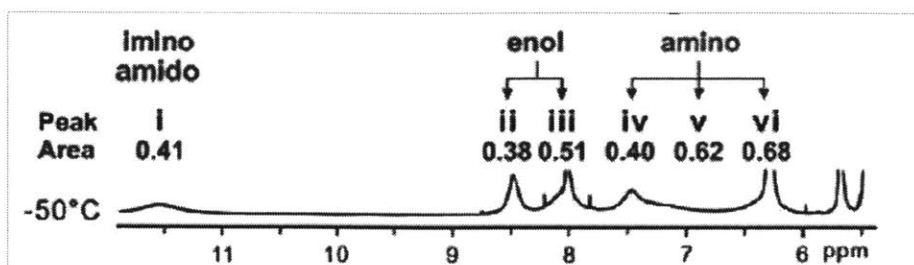


Figure 2.2  $^1\text{H}$  NMR spectrum of KP1212 in  $\text{DMF-d}_7$  at  $-50\text{ }^\circ\text{C}$ .

### 2.3 Results and Discussion

There are only three exchangeable protons on the nucleobase of KP1212. Thus, the presence of six distinct resonances corresponding to these protons at low temperature indicates that multiple tautomers are present in solution. KP1212 has a total of five possible tautomers (Figure 2.3) These peaks were assigned to imino, amido, enol, and amino protons on the basis of chemical shift. In contrast, when the VT NMR experiment was repeated on the native nucleoside 2'-deoxycytidine (dC), only the common keto-amino tautomer was observed (Figure S1). The predominance of the keto-amino tautomer for dC is expected, as the rarer tautomers of dC can mispair with adenine and cause spontaneous mutations.<sup>19,20</sup>

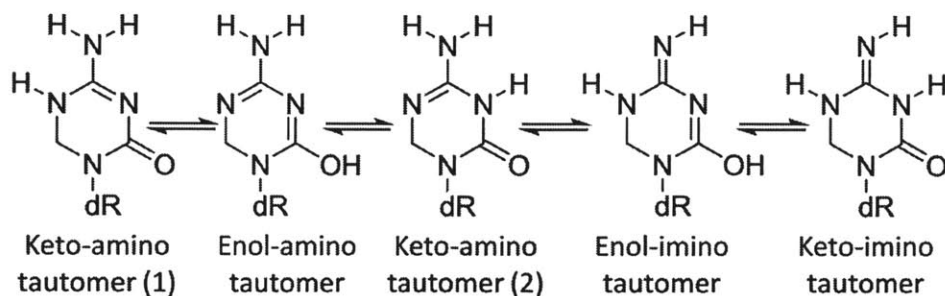
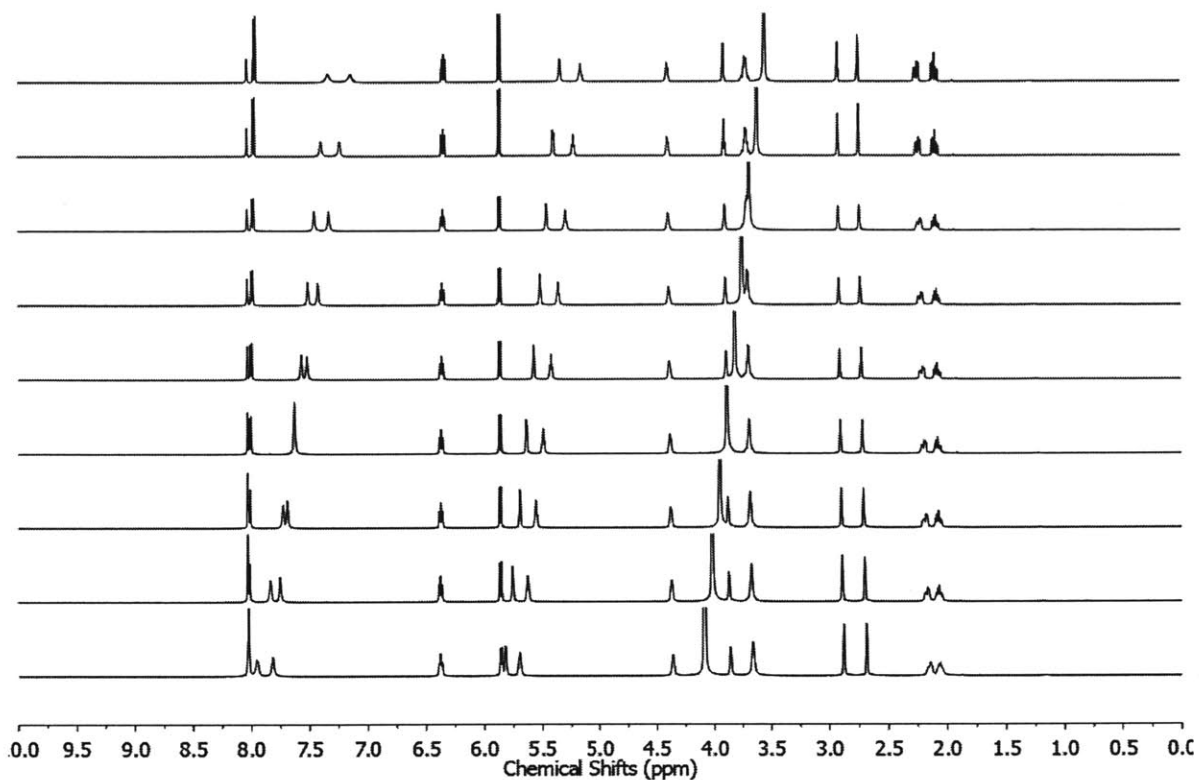
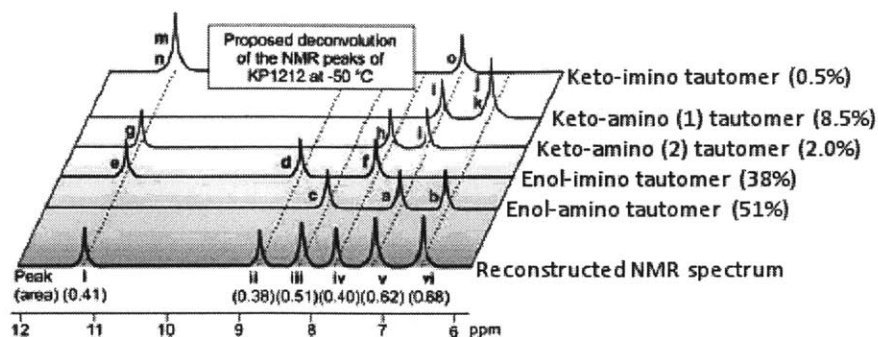


Figure 2.3 Possible tautomers of KP1212.



**Figure 2.4** 2'-Deoxycytidine VT  $^1\text{H}$  NMR spectra from 20 °C to -60 °C (top to bottom) in  $\text{DMF-d}_7$ .

In the proton NMR spectrum at -50 °C, multiple active nucleobase proton resonances overlapped, making it difficult to determine integral values for each peak. To integrate these peaks, the spectrum was reconstructed using Varian NMR simulation software. The reconstructed spectrum was integrated and normalized to give relative amounts of each proton resonance (Figure 2.5). The integral values, normalized to three active protons, were used in a system of linear equations to calculate the relative amounts of tautomers. A comprehensive analysis of possible assignments provided a unique solution to the spectroscopic data. This solution gave the relative ratios of the KP1212 tautomers in  $\text{DMF-d}_7$  at -50 °C (Figure 2.6).

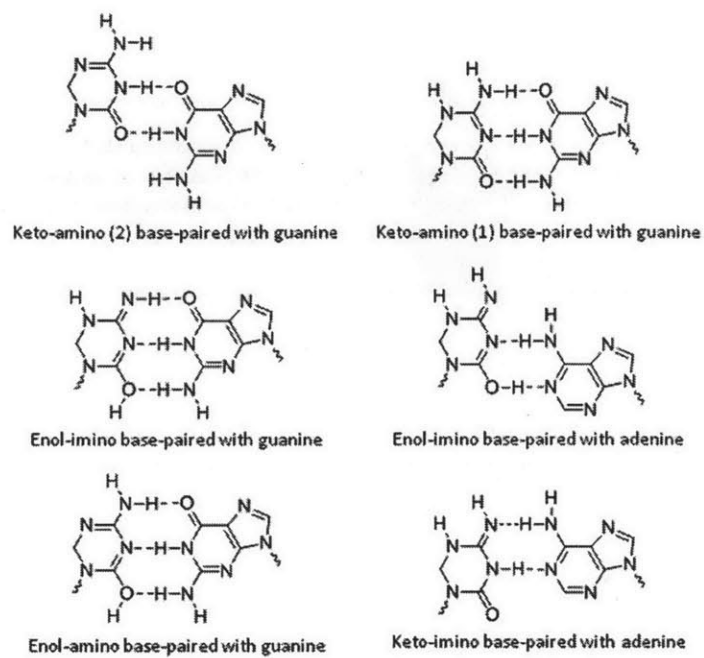


**Figure 2.5** Reconstructed NMR spectra of the exchangeable nucleobase protons of KP1212 in DMF-d<sub>7</sub> at -50 °C and peak assignment based on their chemical shift.

	Matrix A	Matrix X	Matrix B							
Peak i	$\begin{pmatrix} 0 & 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 2 & 0 \end{pmatrix}$	$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{pmatrix}$	$\begin{pmatrix} 0.41 \\ 0.38 \\ 0.51 \\ 0.40 \\ 0.62 \\ 0.68 \end{pmatrix}$	$\Rightarrow$	$\begin{aligned} X_2 + X_3 + 2X_5 &= 0.41 \\ X_2 &= 0.38 \\ X_1 &= 0.51 \\ X_2 + X_3 &= 0.40 \\ X_1 + X_3 + X_4 + X_5 &= 0.62 \\ X_1 + 2X_4 &= 0.68 \end{aligned}$					
Peak ii						$\Rightarrow$	$\begin{aligned} X_1 &= 0.51 \\ X_2 &= 0.38 \\ X_3 &= 0.02 \\ X_4 &= 0.085 \\ X_5 &= 0.005 \end{aligned}$			
Peak iii								$\Rightarrow$		
Peak iv									$\Rightarrow$	
Peak v										$\Rightarrow$
Peak vi										

**Figure 2.6** Matrix equation of peak integrals, system of linear equations, and distribution of tautomers for KP1212 in DMF-d<sub>7</sub> at -50 °C.  $X_1$  is enol-amino,  $X_2$  is enol-imino,  $X_3$  is keto-amino (2),  $X_4$  is keto-amino (1), and  $X_5$  is keto-imino.

The proportions of tautomers calculated from the NMR spectral data support the proposed mechanism of lethal mutagenesis. Under these conditions, the enol tautomers of KP1212 were the dominant species, while the keto tautomers were the minor species. In contrast, the natural DNA nucleosides like dC are found almost entirely in their keto tautomers and the enol tautomers are rarely observed.<sup>19,20</sup> The different keto and enol tautomers of KP1212 can base-pair with either guanosine or adenosine, inducing purine transitions (Figure 2.7). Thus, the tautomer distribution of KP1212 provided by VT NMR experiments supports the putative mechanism for lethal mutagenesis.



**Figure 2.7** Examples of KP1212 tautomers base-paired with the purine bases guanine and adenine.

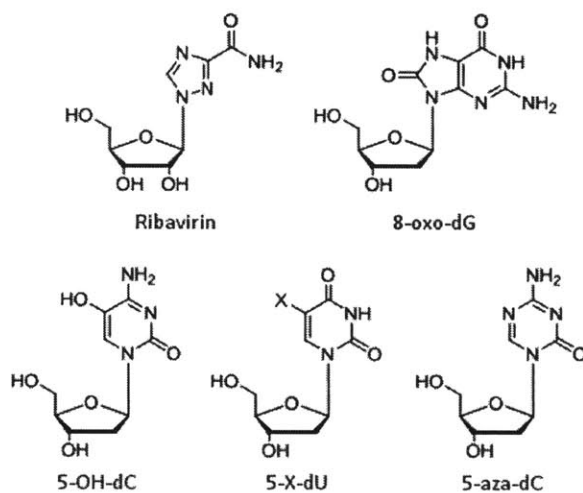


## Chapter 3

### Spectroscopic Exploration of Mutagenic Nucleoside Analogs

#### 3.1 Introduction

Having confirmed that multiple tautomers are observable for the mutagenic nucleoside analog KP1212 via VT NMR, an effort was undertaken to identify other potentially mutagenic modified nucleosides on the basis of their NMR spectra. These nucleosides could be the basis for next-generation antiviral agents that work by the principle of lethal mutagenesis. A variety of nucleoside and nucleobase analogs have demonstrated mutagenic effects due to their ability to mispair, including 5-bromo-2'-deoxyuridine, 5-fluoro-2'-deoxyuridine, and 5-azacytidine.<sup>21</sup> To look for the presence of multiple tautomers, 1D, 2D, and VT NMR experiments were carried out on a variety of modified nucleoside analogs, several of which are associated with DNA damage and mutations (Figure 3.1).



**Figure 3.1** Modified nucleoside analogs studied via VT NMR. X = F, Cl, Br, I.

#### 3.2 Experimental

As with KP1212, the NMR experiments on these modified nucleoside analogs were carried out on a Varian Inova-500 NMR spectrometer. DMF- $d_7$  was used as the solvent for all of the compounds. For each nucleoside, a full chemical assignment was completed using 1D ( $^1\text{H}$ ,  $^{13}\text{C}$ ) and 2D (gCOSY, HSQC, gHMBC) NMR experiments. Once all other protons were accounted for, variable temperature (VT) NMR was used to study the active nucleobase protons as the temperature was decreased from 20 °C to -60 °C.

### 3.3 Results and Discussion

5-Hydroxy-2'-deoxycytidine (5-OH-dC) is a product of oxidative DNA damage that increases the frequency of cytidine to thymidine mutations *in vivo*.<sup>22</sup> The observed mutagenic properties of 5-OH-dC against HIV started the search for other mutagenic nucleosides like KP1212.<sup>13</sup> VT NMR identified four distinct peaks corresponding to the three active protons on the nucleobase, indicating the presence of multiple tautomers (Figure 3.2). The keto-amino tautomer of 5-OH-dC is expected to base-pair with guanosine, while the keto-imino tautomer will base-pair with adenosine (Figures 3.3, 3.4). Thus, the presence of both tautomers in the NMR spectrum supports the mechanism of tautomerization-induced lethal mutagenesis. These results are supported by other spectroscopy techniques that show a higher frequency of the rare imino tautomer for 5-OH-dC than for dC.<sup>22</sup> In addition, the 5-OH moiety has a  $\text{p}K_a = 7.4$ , which is close to physiological pH.<sup>23</sup> Deprotonation of the nucleobase may affect the distribution of tautomers under physiological conditions. VT NMR experiments support the putative mechanism for lethal mutagenesis caused by 5-OH-dC.

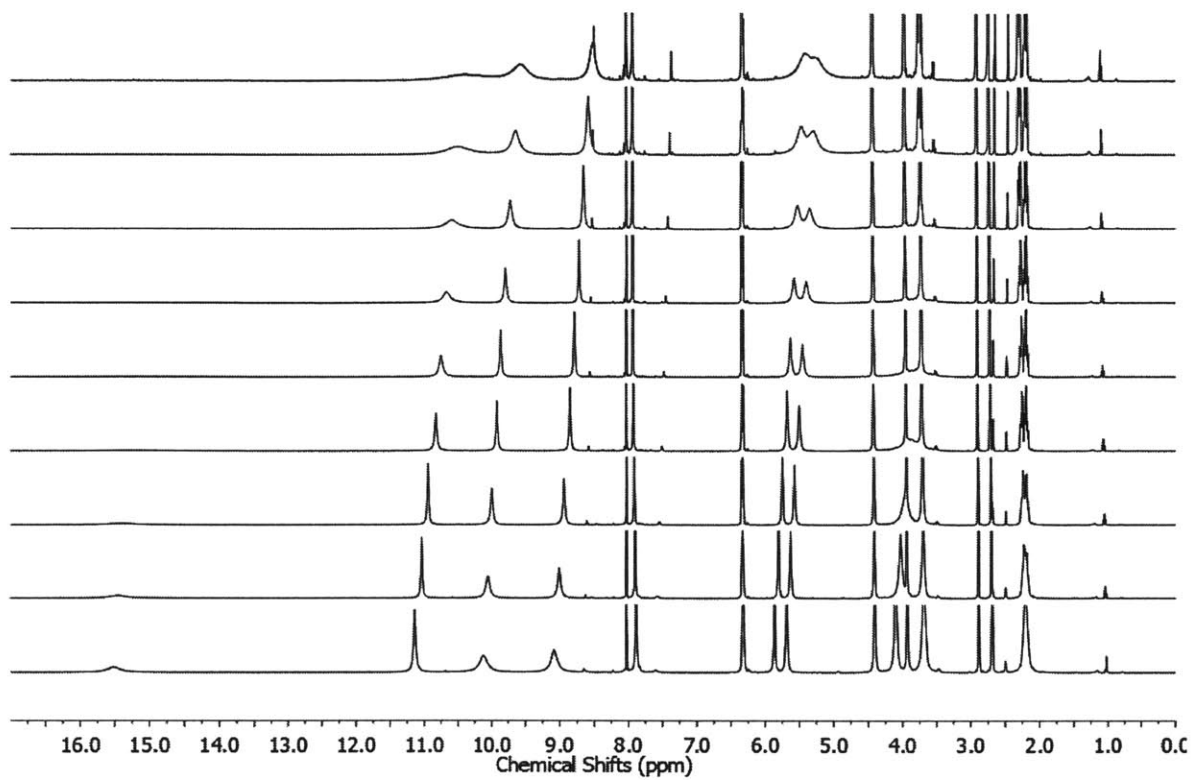


Figure 3.2 VT  $^1\text{H}$  NMR spectra of 5-OH-dC in  $\text{DMF-d}_7$  from  $20\text{ }^\circ\text{C}$  to  $-60\text{ }^\circ\text{C}$  (top to bottom).

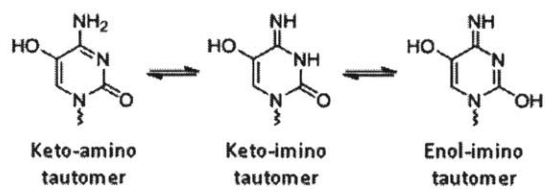
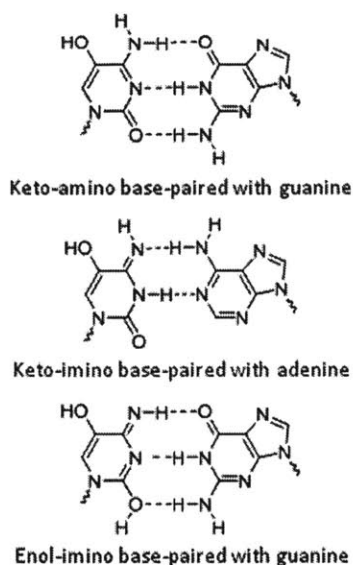
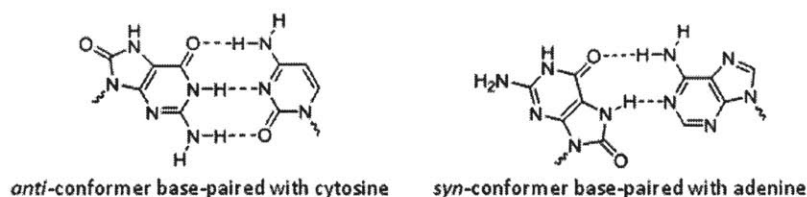


Figure 3.3 Possible tautomers of 5-OH-dC.



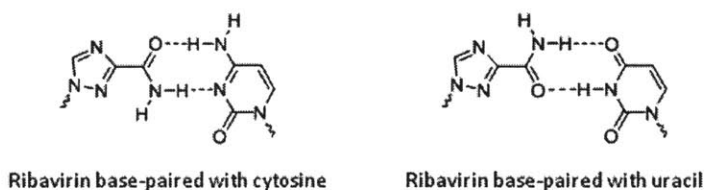
**Figure 3.4** 5-OH-dC tautomers base-paired with the purine bases guanine and adenine.

8-Oxo-7,8-dihydro-2'-deoxyguanosine (8-oxo-dG) is a biomarker of oxidative stress induced by free radicals.<sup>22</sup> The 8-oxo-dG lesion can mispair with adenine, causing guanine to thymine transversions during DNA replication.<sup>23</sup> One possible explanation for the mutagenicity of 8-oxo-dG is mispairing by rare tautomers. However, VT NMR experiments on 8-oxo-dG showed no evidence of multiple tautomers, indicating that this is not likely the mechanism. An alternative explanation for the mutagenic properties of 8-oxo-dG is the use of low-fidelity bypass polymerases. The normal replicative DNA polymerases are inefficient at extending past an 8-oxo-dG lesion, and a bypass polymerase is recruited for translesion DNA synthesis. However, the bypass polymerases used to extend past the 8-oxo-dG lesion have low fidelity and often incorporate the wrong base opposite 8-oxo-dG.<sup>24</sup> In addition, the ability of 8-oxo-dG to base-pair with cytidine requires that both nucleosides are in their *anti*-conformations. If 8-oxo-dG rotates around the glycosidic linkage, the resulting *syn*-conformation of 8-oxo-dG can base-pair with the *anti*-conformation of adenosine (Figure 3.5).<sup>25</sup> Thus, the use of low fidelity bypass polymerases and rotation around the glycosidic linkage are likely the mechanisms behind the mutagenicity of 8-oxo-dG.



**Figure 3.5** 8-oxo-dG conformers base-paired with the bases cytosine and adenine.

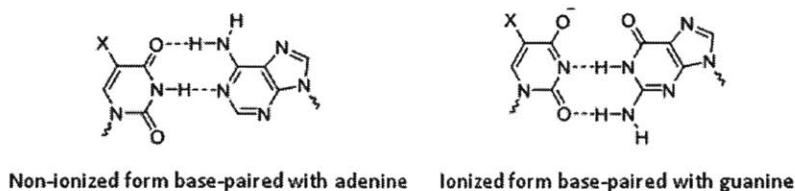
Ribavirin is mutagenic nucleoside analog used in the clinical treatment of hepatitis C. Though there are multiple proposed mechanisms for its antiviral properties, one theory is that ribavirin causes lethal mutagenesis of HCV.<sup>7-10</sup> VT NMR experiments on ribavirin showed a single tautomer at all temperatures, meaning that tautomerization is unlikely to cause mispairing and lethal mutagenesis. Unlike natural purines, the nucleobase-like portion of ribavirin contains a triazole heterocycle with a formamide attached to the C<sup>3</sup> carbon. Rotation around the C<sup>3</sup>-carbonyl bond forms two conformers with different hydrogen bonding patterns. One conformer resembles guanosine and is able to base-pair with cytidine, while the other conformer resembles adenosine and can base-pair with uridine (Figure 3.6). This unique ability to rotate allows ribavirin to base-pair with either pyrimidine and increases the mutation rate of HCV, thereby causing lethal mutagenesis.



**Figure 3.6** Ribavirin rotamers base-paired with the pyrimidine bases cytosine and uracil.

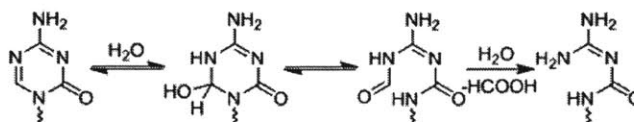
The halogenated analogs 5-X-2'-deoxyuridine and 5-X-uridine, where X is fluorine, bromine, or iodine, induce mutations in a variety of biological systems.<sup>25</sup> However, evidence of multiple tautomers is not observed for the halogenated 2'-deoxyuridine analogs via VT NMR spectroscopy. An alternative explanation is that the electron-withdrawing halogen lowers the  $pK_a$  of the nucleobase, allowing the N<sup>6</sup> position to be deprotonated at physiological pH. For example, while the

$pK_a$  of uridine is 9.25, the  $pK_a$  of 5-fluorouridine is 7.57.<sup>26</sup> Resonance structures of the conjugate base can pair with either adenosine or guanosine, increasing the rate of purine transitions (Figure 3.7). Ionization induced by electron-withdrawing groups is most likely the mechanism behind the mutagenic properties of halogenated nucleoside analogs.

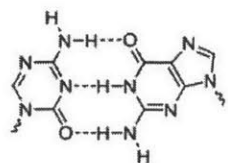


**Figure 3.7** 5-X-dU base-paired with the purine bases guanine and adenine.

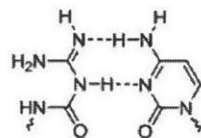
5-Aza-2'-deoxycytidine (decitabine, 5-aza-dC) is a nucleoside analog used clinically in the treatment of myelodysplastic syndrome (MDS). It acts as a hypomethylating agent that inhibits DNA methyltransferases and is highly cytotoxic.<sup>29,30</sup> However, both decitabine and its ribose analog azacytidine also induce lethal mutagenesis at lower cellular concentrations.<sup>31,32</sup> Unlike KP1212, decitabine shows a single tautomeric species in VT NMR experiments, suggesting that tautomerization does not explain its mutagenic properties. An alternative mechanism is chemical rearrangement of the nucleobase. Once decitabine or azacytidine are incorporated into the corresponding nucleic acid, hydration occurs across the N<sup>5</sup>-C<sup>6</sup> double bond. The ring opens and formic acid is lost, allowing the remainder of the structure to base-pair with cytidine (Figure 3.8).<sup>32,33</sup> As a result, decitabine increase the rate of guanosine to cytidine transversions (Figure 3.9).<sup>32</sup> Though similar in their chemical structure, decitabine and KP1212 use very different mechanisms to cause lethal mutagenesis.



**Figure 3.8** Chemical reaction of decitabine.



Decitabine base-paired with guanine



Open ring base-paired with cytosine

**Figure 3.9** Decitabine base-paired with the bases guanine and cytosine.

## Chapter 4

### Synthesis and Spectroscopic Analysis of Dihydronucleosides

#### 4.1 Introduction

From a structural standpoint, there are two differences between the natural nucleoside dC and KP1212. First, the carbon atom in position 5 is replaced by a nitrogen atom. Second, the 5,6-double bond in dC is reduced to a single bond in KP1212. In order to understand why KP1212 is able to readily tautomerize, while dC shows no evidence of tautomerization, each of these changes was substituted independently. As previously stated, decitabine, the unsaturated analog with a nitrogen atom substituted on the ring, shows no evidence of tautomerization via VT NMR experiments. This suggests that the effect of substituting a nitrogen atom at position 5 is not enough to explain the multiple tautomers observed with KP1212.

The unsubstituted, saturated analog, 5,6-dihydro-2'-deoxycytidine (DHdC), was not commercially available and there was no existing literature on its medicinal properties against viruses. Hydrogenation of the carbon-carbon double bond in dC may contribute significantly to the efficacy of KP1212 as a mutagenic nucleoside analog. Hydrogenation breaks the aromaticity of the cytosine base moiety, causing the electron density on the ring to become more localized. In addition, the heterocyclic ring may become puckered at the C<sup>5</sup> and C<sup>6</sup> positions when its aromaticity is broken. This could contribute to the low cytotoxicity of KP1212 relative to decitabine, as human DNA polymerases may reject KP1212 due to its unnatural puckered shape. Based on the dissimilar results from NMR studies of KP1212 and decitabine, as well as the structural differences between KP1212 and dC, DHdC had the potential to display multiple tautomers, similar to KP1212, and thus be a lethal mutagenic agent.

#### 4.2 Experimental



2'-Deoxycytidine (284.8 mg, 1.253 mmol) was dissolved in 25 mL MeOH via sonication to form a 0.05 M solution. Three stainless steel cartridges were packed with ~400 mg of Rh/alumina (5%) each and connected in series. The solution was loaded into an 8 mL stainless steel syringe in a syringe pump and connected to the cartridges and a hydrogen gas tank via a Y-linker. The solution and hydrogen gas were pumped through the connected cartridges at a flow rate of 80  $\mu$ L/min with hydrogen gas pressure of 175 psi. The outflow was connected in ~5 mL fractions and checked via TLC (30% MeOH-NH<sub>3</sub> in DCM on silica, starting material R<sub>f</sub> = 0.29, product R<sub>f</sub> = 0.14). All fractions showed similar conversion rates and were combined and concentrated via rotary evaporation to 282 mg of a white, sticky solid. Proton NMR showed ~80% conversion to product with anomerization of the ribose sugar. To purify the product, a column of silica gel was prepared using 10% MeOH-NH<sub>3</sub> in DCM. The product was dissolved in a minimal amount of MeOH-NH<sub>3</sub> in DCM and loaded onto the column. The column was eluted with a gradient of 10-30% MeOH-NH<sub>3</sub> in DCM. Fractions that appeared to contain the desired product on TLC were combined and concentrated via rotary evaporation to yield 61.3 mg of a light yellow solid. Proton NMR showed <10% of the undesired  $\alpha$ -anomer and >90% of the desired  $\beta$ -anomer. The product was repurified using a pipette column prepared with silica gel using 10% MeOH-NH<sub>3</sub> in DCM. The product was dissolved in a minimal amount of MeOH-NH<sub>3</sub> in DCM and loaded onto the column. The column was eluted with a gradient of 10-20% MeOH-NH<sub>3</sub> in DCM. Fractions that appeared to contain the desired product on TLC were combined and concentrated via rotary evaporation to yield 32.1 mg of a light yellow solid. Proton NMR showed full conversion to the desired product with no starting material remaining and <5% of the undesired  $\alpha$ -anomer.

### 4.3 Results and Discussion

Previous work on the hydrogenation of dC found that rhodium on alumina was an effective catalyst for the reduction of the 5,6-double bond (Figure 4.1).<sup>34-36</sup> However, the authors noted that

overhydrogenation and hydrolysis of the 4-amino group presented difficulties.<sup>35,36</sup> To address these concerns, the hydrogenation reaction was carried out via flow chemistry to control the amount of time the nucleoside spends in contact with the catalyst and hydrogen gas, thereby avoiding overhydrogenation (Figure 4.2). The nucleoside was dissolved in methanol instead of water to prevent hydrolysis of the amino group.

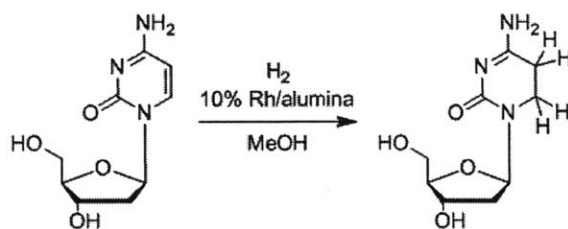


Figure 4.1 Hydrogenation of dC.

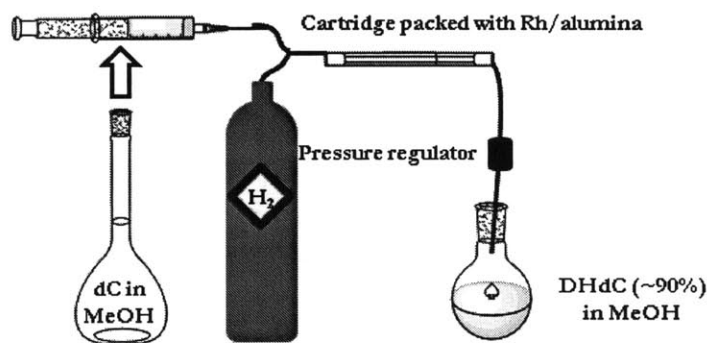
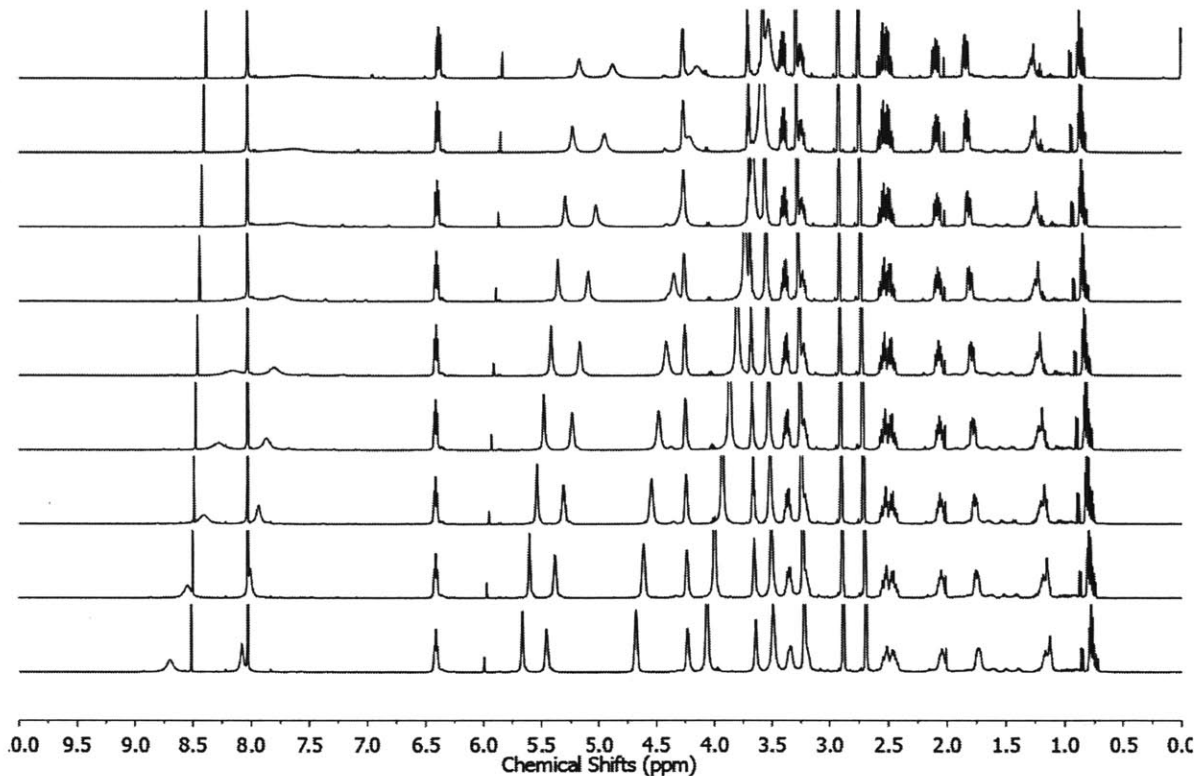


Figure 4.2 Continuous-flow set-up for the hydrogenation of dC.

DHdC was synthesized from dC via hydrogenation with 80-90% conversion in several flow reactions. One challenge presented by the reaction and purification processes was anomerization of the ribose sugar to form the  $\alpha$ -anomer. The  $\alpha$ -anomer of nucleoside derivatives cannot be easily used by RT or DNA polymerases and thus is unlikely to have medicinal relevance. Anomerization appears to be catalyzed by acidic conditions, as anomerization occurs to a great extent when Rh/alumina is used as the catalyst instead of Rh/C. However, slightly acidic conditions typically

increase the rate of hydrogenation reactions, as more protons are available to react with the starting material. As a result, Rh/alumina displays a higher conversion rate than Rh/C (80-90% compared to 20-30%). The increase in conversion rate provided by Rh/alumina was great enough that it was determined to be a better catalyst for the hydrogenation reaction than Rh/C, even though the undesired  $\alpha$ -anomer would have to be separated from the desired  $\beta$ -anomer. To prevent further anomerization during purification on acidic silica gel, MeOH saturated with ammonia (MeOH-NH<sub>3</sub>) was used as the polar solvent in the mobile phase.

To determine whether DHdC displays multiple tautomers that could allow it to base pair with both adenosine and guanosine, 1D, 2D, and VT NMR experiments were carried out on the purified DHdC product (Figure 4.3). At 20°C, exchangeable protons on the base moiety are observed at 7.58 ppm and 4.14 ppm. At -60°C, exchangeable protons on the base moiety are observed at 8.70 ppm, 8.08 ppm, and 4.68 ppm. Protons with a chemical shift of 8.70, 8.08, and 7.58 ppm likely correspond to enolic, amido, or imino protons, while protons with a chemical shift of 4.68 and 4.14 ppm likely correspond to amino protons. The expected chemical shifts for enolic, amido, and imino protons are relatively close, and thus these species cannot be assigned at this stage. However, VT NMR clearly shows three resonances at low temperature, though only two active protons are present on the base moiety. Only the existence of multiple tautomeric species would explain the presence of three exchangeable protons resonances. Furthermore, the chemical shifts observed suggest that both the keto-amino and keto-imino/enol-imino tautomers occur. Based on their chemical structures, the keto-amino tautomer will pair with guanine and the keto-imino/enol-imino tautomers can pair with adenine. Due to the existence of multiple tautomers, DHdC may be able to alternatively base pair and cause adenine to guanine and guanine to adenine mutations, similarly to KP1212.



**Figure 4.3** VT  $^1\text{H}$  NMR spectra of DHdC in  $\text{DMF-d}_7$  from  $20\text{ }^\circ\text{C}$  to  $-60\text{ }^\circ\text{C}$  (top to bottom).

As a comparison, 5,6-dihydro-2'-deoxyuridine (DHdU) was synthesized from 2'-deoxyuridine (dU) via a similar continuous-flow hydrogenation reaction (Figure 4.4). This reaction showed full conversion to DHdU using Rh/C as the catalyst, with no starting material apparent on TLC or proton NMR. However, VT NMR experiments showed a single tautomer of DHdU exists at low temperatures (Figure 4.5). The lack of tautomerization is unsurprising, given the chemical stability of the imide moiety in the uridine base. The low reactivity of DHdU is further supported by the literature, which reports no problems with overhydrogenation.<sup>37</sup> Ionization of the imide proton, the alternative mechanism of lethal mutagenesis observed with 5-F-dU, 5-Br-dU, and 5-I-dU, is unlikely to occur for DHdU. The  $\text{p}K_a$  of DHdU is expected to be greater than that of dU, as the

non-aromatic ring would provide less resonance stabilization to the conjugate base. Thus, it is unlikely that DHdU will cause lethal mutagenesis through tautomerization or ionization.

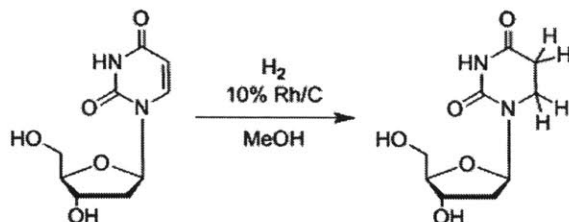


Figure 4.4 Hydrogenation of dU.

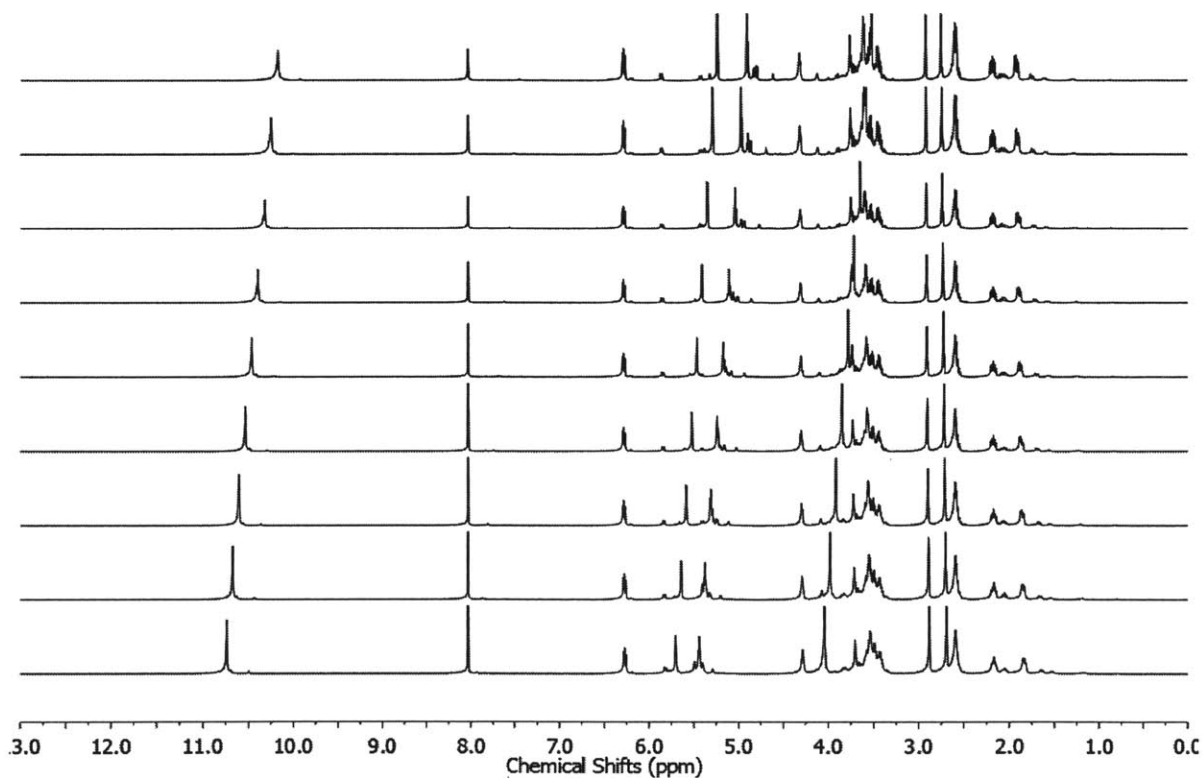
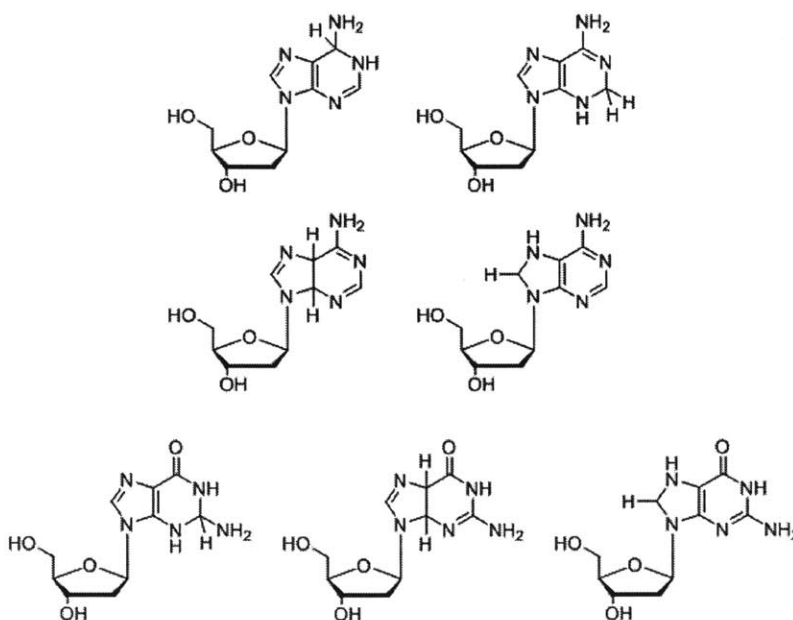


Figure 4.5 VT <sup>1</sup>H NMR spectra of DHdU in DMF-d<sub>7</sub> from 20 °C to -60 °C (top to bottom).

Attempts were made to synthesize hydrogenated purine nucleosides using the same flow hydrogenation methodology. As bicyclic rings, purines could be hydrogenated in multiple positions: the C<sup>2</sup>-N<sup>3</sup> double bond, the C<sup>4</sup>-C<sup>5</sup> double bond, and the N<sup>7</sup>-C<sup>8</sup> double bond, as well as the N<sup>1</sup>-C<sup>6</sup>

double bond found in adenine but not in guanine (Figure 4.6). A solution of 2'-deoxyadenosine (dA) in MeOH was prepared and run through the flow reactor under the same conditions as dC. However, no reaction occurred under these conditions. The low reactivity of dA compared to dC and dU is unsurprising, as more energy would be required to overcome the stability of a bicyclic aromatic ring system. If hydrogenation were to occur on a purine derivative, it would most likely reduce the N<sup>7</sup>-C<sup>8</sup> double bond, as this would break the aromaticity of the smaller imidazole ring while leaving the larger pyrimidine ring intact. In order to hydrogenate the purine nucleosides, an alternative, highly active catalyst, such as palladium on carbon, Raney nickel, or Wilkinson's catalyst, could be utilized. Though increasing the temperature would not be advisable, as it would likely cause decomposition of the starting material, the pressure of hydrogen gas in the flow reactor could be increased. In addition, the flow rate could be decreased to give the starting material more time to interact with the catalyst and undergo hydrogenation. Though hydrogenation of dA and dG will likely be a challenging endeavor, their hydrogenated analogs are worth exploring, as they may show mutagenic tautomerization similar to KP1212 and DHdC.



**Figure 4.6** Possible hydrogenated purine nucleosides, DHdA and DHdG.

## Chapter 5

### Conclusions and Future Directions

Despite major advances in the treatment of HIV/AIDS with existing therapeutics, the high mutation rate of the virus makes drug resistance a constant problem. Lethal mutagenesis of HIV provides an alternative mechanism to traditional therapeutic strategies that uses the virus's high mutation rate against it. KP1212 is a nucleoside analog under study for the treatment of HIV. The molecule increases the rate of mutation, specifically adenine to guanine and guanine to adenine transitions in HIV with low cellular and mitochondrial toxicity. It is hypothesized that KP1212 exists in multiple tautomeric forms, which alternatively base-pair and induce mutations once the nucleoside is incorporated into the HIV genome by RT.

VT NMR spectroscopy was used in conjunction with 1D and 2D NMR experiments to demonstrate the existence of multiple tautomers of KP1212 in DMF at -50 °C. Unlike dC, where the keto tautomers are the dominant species, the majority of KP1212 is found as enol tautomers. The presence of multiple tautomers for KP1212 supports the hypothesis behind its mutagenic properties. VT NMR also showed multiple tautomers for 5-OH-dC, another mutagenic nucleoside thought to increase adenine to guanine and guanine to adenine transitions through tautomerization. Other mutagenic nucleosides, such as 8-oxo-dG, ribavirin, 5-X-dU, and decitabine, did not display evidence of multiple tautomers via VT NMR. However, in all of these cases, other mechanisms have been proposed to explain their mutagenicity, suggesting that the results of VT NMR spectroscopy are consistent with *in vitro* and *in vivo* experiments.

NMR studies on KP1212 and decitabine suggested that hydrogenation of the 5,6-double bond in dC nucleoside analogs could create nucleosides that display multiple tautomers. Such tautomeric nucleoside analogs have the potential to cause viral decay acceleration and thus could be studied as lead compounds for the treatment of viral diseases. In order to study DHdC, we

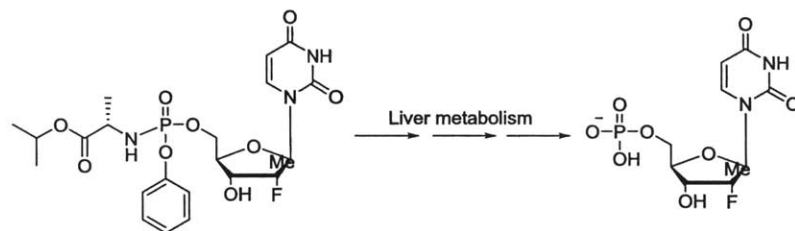
developed a protocol for the selective hydrogenation of dC. Hydrogenation of dC to DHdC was successfully carried out using a continuous-flow reaction set-up that will likely translate well into efficient large-scale preparation. VT NMR studies suggest that DHdC adopts multiple tautomers, which could allow it to alternatively base pair with adenine and guanine. These preliminary NMR experiments indicate that DHdC is a promising antiviral agent and further studies should be done to measure its pharmacologic properties *in vitro* and *in vivo*.

Given the potential mutagenic properties of DHdC, a full library of hydrogenated nucleoside analogs should be prepared. Synthesis of 5-substituted DHdC derivatives such as 5-OH-DHdC and 5-X-DHdC should be possible using the methodology for DHdC. DHdU was synthesized via a similar continuous-flow hydrogenation of dU. However, this modified nucleoside showed no evidence of tautomerization in VT NMR experiments. The hydrogenated purine derivatives, DHdA and DHdG, have not yet been synthesized, but these products should be pursued as possible novel viral mutagens.

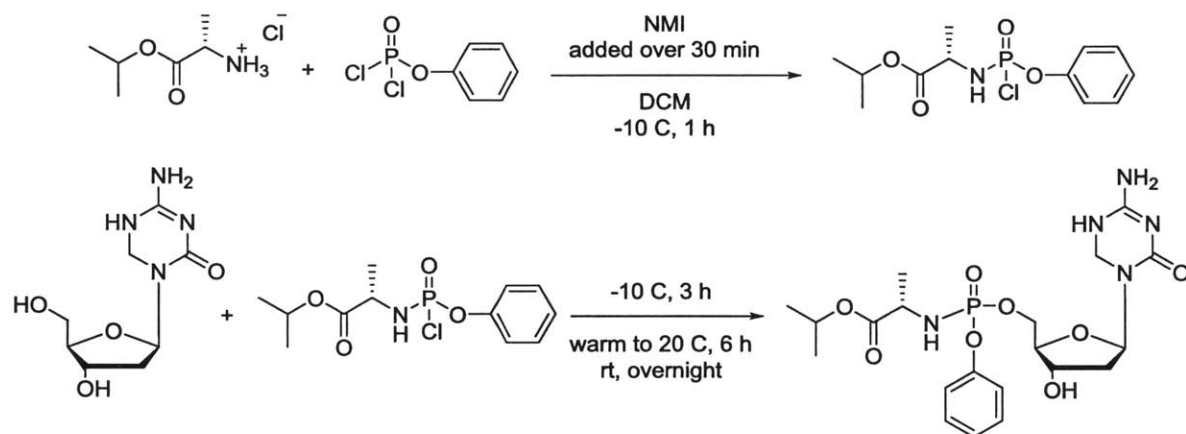
In addition to using hydrogenation of nucleosides to find new lead compounds, work on the next generation of KP1212-derived molecules is ongoing. One limitation of KP1212 as a drug is that, in order to be incorporated into DNA, the nucleoside analog must be phosphorylated into the corresponding nucleotide triphosphate. However, the first phosphorylation step is extremely slow. Sofosbuvir, a nucleoside analog recently approved for the treatment of Hepatitis C, found a way to bypass this step. The nucleoside is converted into a protected phosphoramidate prodrug, which is hydrolyzed in the liver to release the active monophosphate derivative at the desired site of action.<sup>38</sup> To increase the efficacy of KP1212, the KP1212 nucleoside should be converted to the protected phosphoramidate, based on the synthetic route used to prepare sofosbuvir. The phosphoramidate of KP1212 can then be studied in *in vitro* and *in vivo* assays to see how its pharmacokinetic properties compare to those of the unphosphorylated nucleoside analog KP1212. This modification may



increase the bioavailability of KP1212 in its active triphosphate form and improve its efficacy against HIV.



**Figure 5.1** Conversion of the phosphoramidate prodrug sofosbuvir to the monophosphate uridine derivative.



**Figure 5.2** Proposed synthesis of the phosphoramidate prodrug derivative of KP1212, based on the synthesis of sofosbuvir.

## References

1. World Health Organization (WHO). *Core epidemiological slides: HIV/AIDS estimates*; WHO HIV Department. Online, 2013.
2. World Health Organization (WHO). *Global update on HIV treatment 2013: results, impact and opportunities*; WHO HIV Department. Online, 2013.
3. World Health Organization (WHO). *Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection*; WHO. Online, 2013.
4. Mackman, R. The Discovery of GS-9131, an Amidate Prodrug of a Novel Nucleoside Phosphonate HIV Reverse Transcriptase Inhibitor. In *Accounts in Drug Discovery, Case Studies in Medicinal Chemistry*; Barrish, J. C., Carter, P. H., Cheng, P. T. W., Zahler, R., Eds.; RSC Publishing: Cambridge, UK, 2011; pp 215-237.
5. Eigen, M. Error catastrophe and antiviral strategy. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13374-13376.
6. Smith, R. A.; Loeb, L. A.; Preston, B. D. Lethal mutagenesis of HIV. *Virus Res.* **2005**, *107*, 215-228.
7. Ortega-Prieto, A. M. et al. Extinction of hepatitis C virus by ribavirin in hepatoma cells involves lethal mutagenesis. *PLoS ONE* **2013**, *8*, e71039.
8. Dietz, J. et al. Deep sequencing reveals mutagenic effects of ribavirin during monotherapy of hepatitis C virus genotype 1-infected patients. *J. Virol.* **2013**, *87*, 6172-6181.
9. Moreno, H.; Grande-Perez, A.; Domingo, E.; Martin, V. Arenaviruses and lethal mutagenesis. Prospects for new ribavirin-based interventions. *Viruses* **2012**, *4*, 2786-2805.
10. Crotty, S. et al. The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. *Nat. Med.* **2000**, *6*, 1375-1379.
11. Furuta, Y. et al. Favipiravir (T-705), a novel viral RNA polymerase inhibitor. *Antiviral Res.* **2013**, *100*, 446-454.
12. Baranovich, T. et al. T-705 (favipiravir) induces lethal mutagenesis in influenza A H1N1 viruses in vitro. *J. Virol.* **2013**, *87*, 3741-3751.
13. Loeb, L. A. et al. Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1492-1497.
14. Harris, K. S. et al. KP-1212/1461, a nucleoside designed for the treatment of HIV by viral mutagenesis. *Antiviral Res.* **2005**, *67*, 1-9.
15. Murakami, E.; Basavapathruni, A.; Bradley, W. D.; Anderson, K. S. Mechanism of action of a novel viral mutagenic covert nucleotide: molecular interactions with HIV-1 reverse transcriptase and host cell DNA polymerases. *Antiviral Res.* **2005**, *67*, 10-17.
16. Mullins, J. I. et al. Mutation of HIV-1 genomes in a clinical population treated with the mutagenic nucleoside KP1461. *PLoS ONE* **2011**, *6*, e15135.
17. Seckarova, P. et al. Direct determination of tautomerism in purine derivatives by low-temperature NMR spectroscopy. *Tetrahedron Letters* **2004**, *45*, 6259-6263.
18. Singh, V. et al. Direct observation of multiple tautomers of oxythiamine and their recognition by the thiamine pyrophosphate riboswitch. *ACS Chem. Biol.* **2014**, *9*, 227-236.
19. Watson, J. D.; Crick, F. H. Genetical implications of the structure of deoxyribonucleic acid. *Nature* **1953**, *171*, 964-967.
20. Wang, W.; Hellinga, H. W.; Beese, L. S. Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17644-17648.
21. Negishi, K.; Bessho, T.; Hayatsu, H. Nucleoside and nucleobase analog mutagens. *Mutation Research* **1994**, *318*, 227-238.

22. Suen, W.; Spiro, T. G.; Sowers, L. C.; Fresco, J. R. Identification by UV resonance Raman spectroscopy of an imino tautomer of 5-hydroxy-2'-deoxycytidine, a powerful base analog transition mutagen with a much higher unfavored tautomer frequency than that of the natural residue 2'-deoxycytidine. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4500-4505.
23. Kupfer, P. A.; Leumann, C. J. Synthesis, base pairing properties and trans-lesion synthesis by reverse transcriptases of oligoribonucleotides containing the oxidatively damaged base 5-hydroxycytidine. *Nucleic Acids Res.* **2011**, *39*, 9422-9432.
24. Roszkowski, K. Oxidative DNA damage – the possible use of biomarkers as additional prognostic factors in oncology. *Front. Biosci.* **2014**, *19*, 808-817.
25. Le Page, F.; Sarasin, A.; Gentil, A.; Guy, A.; Cadet, J. Repair and mutagenic potency of 8-oxoG:A and 8-oxoG:C base pairs in mammalian cells. *Nucleic Acids Res.* **1998**, *26*, 1276-1281.
26. Taggart, D. J.; Fredrickson, S. W.; Gadkari, V. V.; Suo, Z. Mutagenic potential of 8-oxo-7,8-dihydro-2'-deoxyguanosine bypass catalyzed by human Y-family DNA polymerases. *Chem. Res. Toxicol.* **2014**, [ePub ahead of print].
27. Bonnac, L. F.; Mansky, L. M.; Patterson, S. E. Structure-activity relationships and design of viral mutagens and application to lethal mutagenesis. *J. Med. Chem.* **2013**, *56*, 9403-9414.
28. Wempen, I.; Duschinsky, R.; Kaplan, L.; Fox, J. J. Thiation of nucleosides. IV The synthesis of 5-fluoro-2'-deoxycytidine and related compounds. *J. Am. Chem. Soc.* **1961**, *83*, 4755-4766.
29. Tohyama, K. Utility of DNA methyltransferase inhibitors for the treatment of myelodysplastic syndromes. *Curr. Pharm. Des.* **2012**, *18*, 3190-3197.
30. Tsujioka, T.; Yokoi, A.; Uesugi, M.; Kishimoto, M.; Tochigi, A.; Suemori, S.; Tohyama, K. Effects of DNA methyltransferase inhibitors (DNMTIs) on MDS-derived cell lines. *Exp. Hematol.* **2013**, *41*, 189-197.
31. Jackson-Grusby, L.; Laird, P. W.; Magge, S. N.; Moeller, B. J.; Jaenisch, R. Mutagenicity of 5-aza-2'-deoxycytidine is mediated by the mammalian DNA methyltransferase. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4681-4685.
32. Dapp, M. J.; Clouser, C. L.; Patterson, S.; Mansky, L. M. 5-Azacytidine can induce lethal mutagenesis in human immunodeficiency virus type 1. *J. Virol.* **2009**, *83*, 11950-11958.
33. Rogstad, D. K.; Herring, J. L.; Theruvathu, J. A.; Burdzy, A.; Perry, C. C.; Neidigh, J. W.; Sowers, L. C. Chemical decomposition of 5-aza-2'-deoxycytidine (decitabine): kinetic analysis and identification of products by NMR, HPLC, and mass spectrometry. *Chem. Res. Toxicol.* **2009**, *22*, 1194-1204.
34. Cohn, W. E.; Doherty, D. G. The catalytic hydrogenation of pyrimidine nucleosides and nucleotides and the isolation of their ribose and respective ribose phosphates. *J. Am. Chem. Soc.* **1956**, *78*, 2863-2868.
35. Green, M.; Cohen, S. S. Studies on the biosynthesis of bacterial and viral pyrimidines III. Derivatives of dihydrocytosine. *J. Biol. Chem.* **1957**, *228*, 601-609.
36. Hanze, A. R. Nucleic acids. IV. The catalytic reduction of pyrimidine nucleosides (human liver deaminase inhibitors). *J. Am. Chem. Soc.* **1967**, *89*, 6720-6725.
37. Green, M.; Cohen, S. S. Studies on the biosynthesis of bacterial and viral pyrimidines II. Dihydrouracil and dihydrothymine nucleosides. *J. Biol. Chem.* **1957**, *225*, 397-407.
38. Sofia, M. J.; Bao, D.; Chang, W.; Du, J.; Nagarathnam, D.; Rachakonda, S.; Reddy, P. G.; Ross, B. S.; Wang, P.; Zhang, H.-R.; Bansal, S.; Espiritu, C.; Keilman, M.; Lam, A. M.; Micolochick Steuer, H. M.; Niu, C.; Otto, M. J.; Furman, P. A. Discovery of a  $\beta$ -d-2'-deoxy-2'- $\alpha$ -fluoro-2'- $\beta$ -C-methyluridine nucleotide prodrug (PSI-7977) for the treatment of Hepatitis C virus. *J. Med. Chem.* **2010**, *53*, 7202-7218.

## Appendix

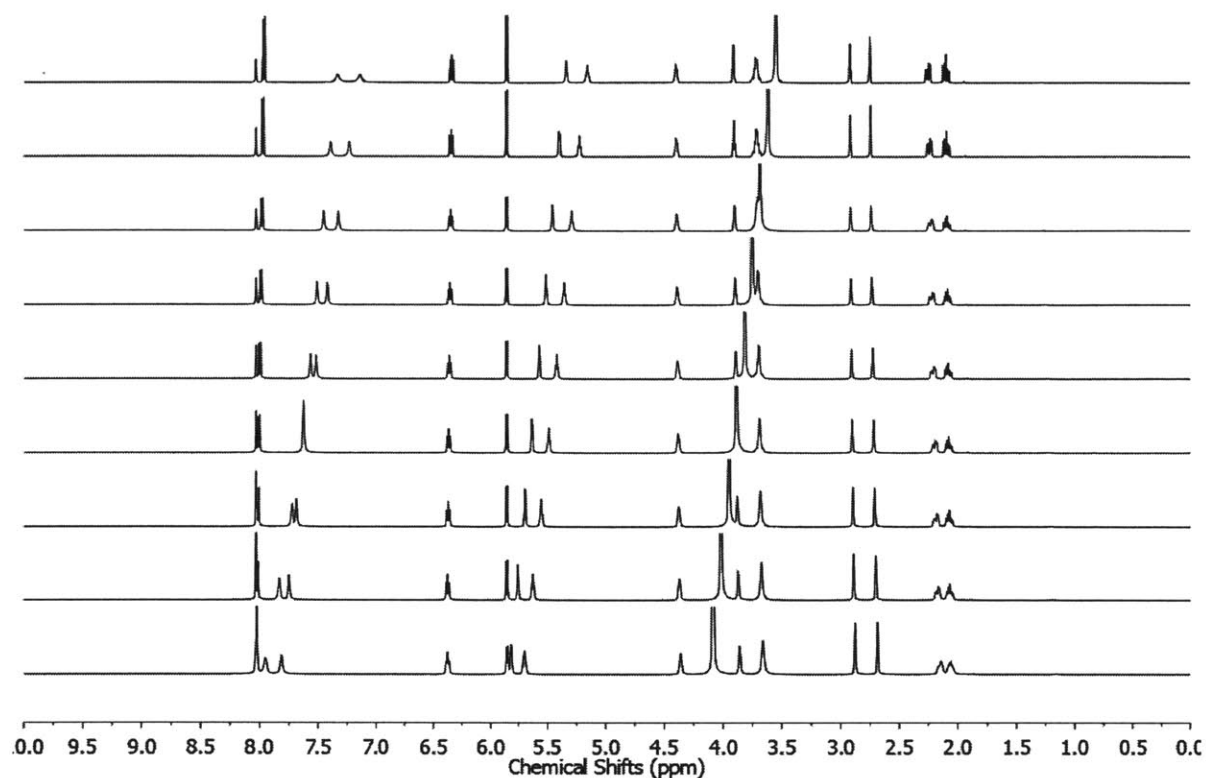
This appendix contains spectral data for all nucleosides studied. All spectra were acquired on a Varian Inova-500 NMR spectrometer.  $^1\text{H}$  NMR spectra were referenced to DMF- $d_6$  (aldehyde proton) at 8.03 ppm.  $^{13}\text{C}$  NMR spectra were referenced to DMF- $d_7$  (carbonyl carbon) at 163.15 ppm.

**Table A.1**  $^1\text{H}$  chemical shift assignments of KP1212 in DMF- $d_7$  at 20 °C.

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
6 a	4.57	Quartet, J = 11.5, 2.0
6 b	4.57	Quartet, J = 11.5, 2.0
1'	6.27	Triplet, J = 7.3
2' a	2.15	Multiplet
2' b	1.87	Multiplet
3'	4.26	Singlet (br)
3'-OH	5.20	Singlet (br)
4'	3.71	Multiplet
5' a	3.57	Singlet (br)
5' b	3.57	Singlet (br)
5'-OH	4.91	Singlet (br)

**Table A.2**  $^{13}\text{C}$  chemical shift assignments of KP1212 in DMF- $d_7$  at 20 °C.

$^{13}\text{C}$ Position	$^{13}\text{C}$ Chemical Shift ( $\delta$ )	$^{13}\text{C}$ Chemical Shift ( $\delta$ ) in $\text{D}_2\text{O}$
2	N/A	161.84 (from HMBC)
4	N/A	160.91 (from HMBC)
6	N/A	52.30 (from HSQC)
1'	83.20 (from HSQC)	85.11
2'	37.06	36.48
3'	72.41	72.41 (reference)
4'	87.39	86.31
5'	63.63	63.11



**Figure A.1** VT  $^1\text{H}$  NMR spectra of 2'-deoxycytidine in  $\text{DMF-d}_7$  from 20 °C to -60 °C (top to bottom).

**Table A.3**  $^1\text{H}$  chemical shift assignments of 2'-deoxycytidine in  $\text{DMF-d}_7$  at 20 °C.

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
4-NH <sub>2</sub> a	7.34	Singlet (br)
4-NH <sub>2</sub> a	7.14	Singlet (br)
5	5.86	Doublet, J = 7.5
6	7.96	Doublet, J = 7.5
1'	6.33	Triplet, J = 6.0
2' a	2.25	Multiplet
2' b	2.10	Multiplet
3'	4.40	Multiplet
3'-OH	5.34	Doublet, J = 4.0
4'	3.91	Quartet, J = 3.5
5' a	3.72	Multiplet
5' b	3.71	Multiplet
5'-OH	5.16	Triplet, J = 5.3

**Table A.4**  $^{13}\text{C}$  chemical shift assignments of 2'-deoxycytidine in DMF- $d_7$  at 20 °C.

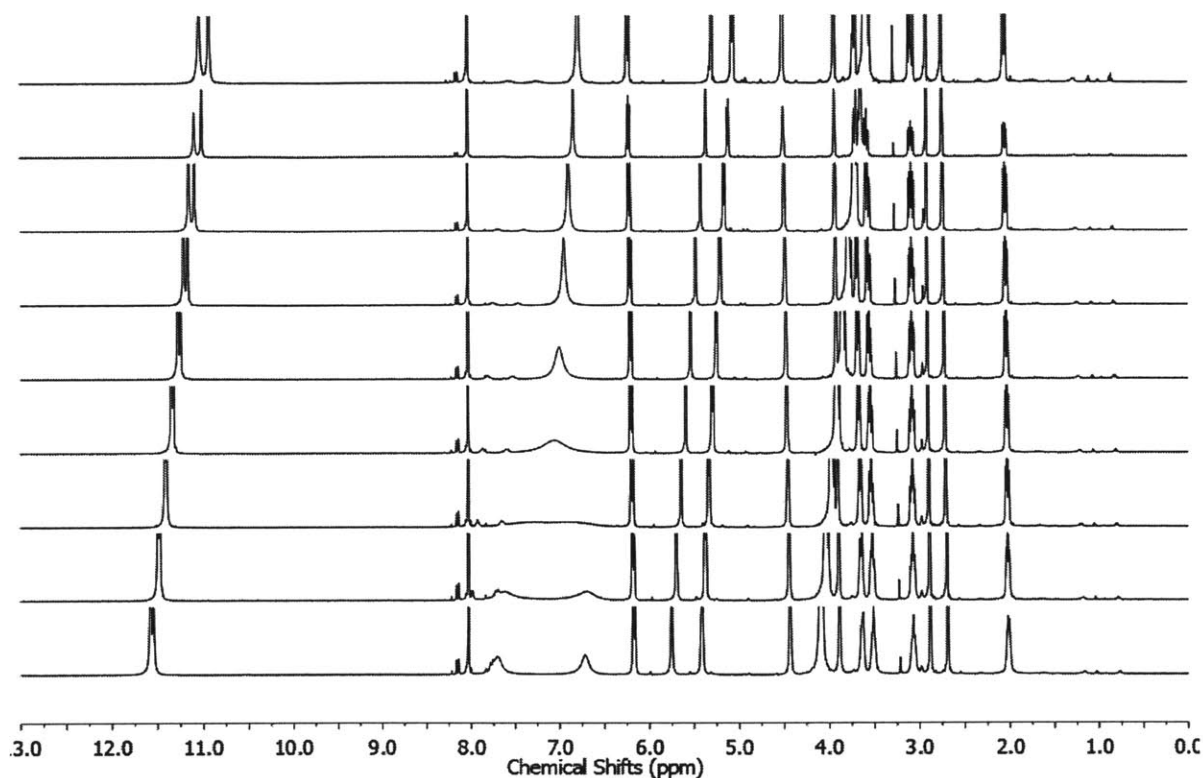
$^{13}\text{C}$ Position	$^{13}\text{C}$ Chemical Shift ( $\delta$ )
2	156.64
4	167.25
5	94.84
6	142.32
1'	86.68
2'	41.89
3'	72.16
4'	88.94
5'	63.06

**Table A.5**  $^1\text{H}$  chemical shift assignments of 5-OH-dC in DMF- $d_7$  at 20 °C.

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
4-NH <sub>2</sub>	8.52	Singlet (br)
4-NH	10.40	Singlet (br)
5-OH	9.57	Singlet (br)
6	7.94	Singlet
1'	6.33	Triplet, J = 7.0
2' a	2.30	Multiplet
2' b	2.20	Multiplet
3'	4.44	Multiplet
3'-OH	5.41	Singlet (br)
4'	3.97	Quartet, J = 3.0
5' a	3.75	Multiplet
5' b	3.75	Multiplet
5'-OH	5.26	Singlet (br)

**Table A.6**  $^{13}\text{C}$  chemical shift assignments of 5-OH-dC in DMF- $d_7$  at 20 °C.

$^{13}\text{C}$ Position	$^{13}\text{C}$ Chemical Shift ( $\delta$ )
2	151.99
4	159.69
5	129.03
6	124.04
1'	87.02
2'	41.65
3'	72.33
4'	89.61
5'	62.92



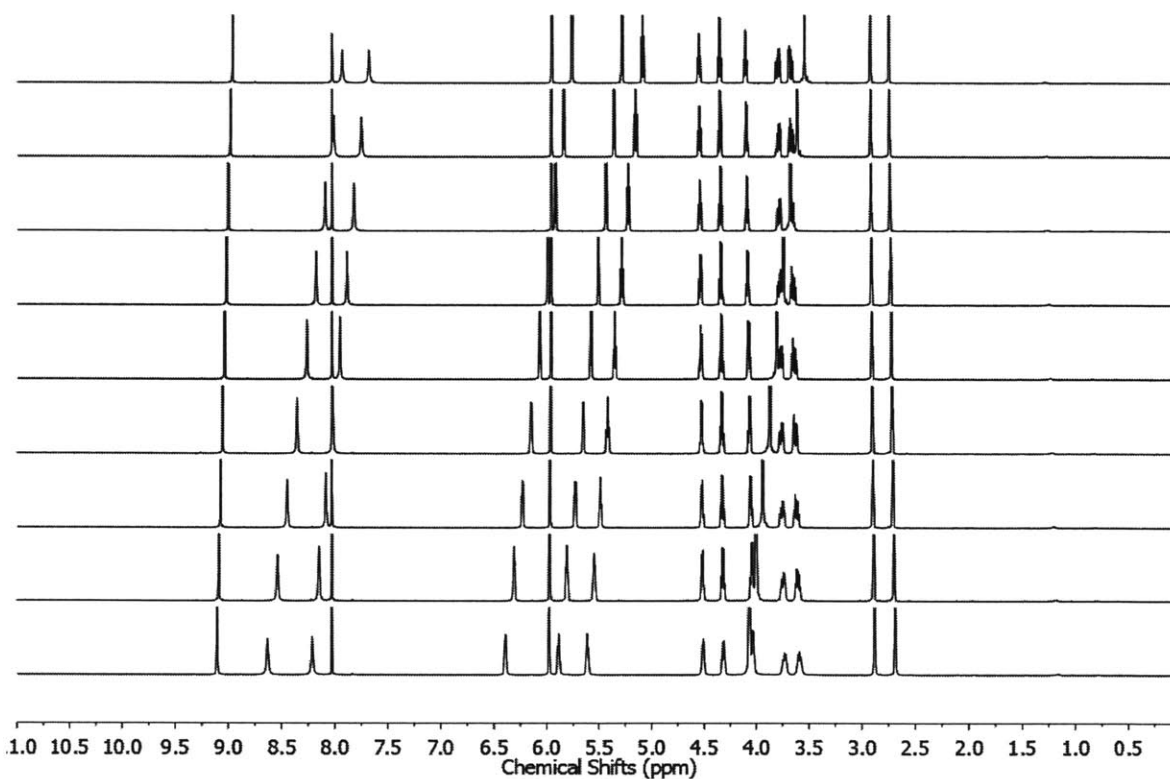
**Figure A.2** VT  $^1\text{H}$  NMR spectra of 8-oxo-dG in  $\text{DMF-d}_7$ , from 20 °C to -60 °C (top to bottom).

**Table A.7**  $^1\text{H}$  chemical shift assignments of 8-oxo-dG in  $\text{DMF-d}_7$  at 20 °C.

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
1-NH	11.01	Singlet (br)
2-NH <sub>2</sub> a	6.77	Singlet (br)
2-NH <sub>2</sub> b	6.77	Singlet (br)
7-NH	10.91	Singlet (br)
1'	6.23	Quartet, $J = 6.5, 2.5$
2' a	3.09	Multiplet
2' b	2.06	Multiplet
3'	4.51	Multiplet
3'-OH	5.27	Doublet, $J = 3.5$
4'	3.94	Multiplet
5' a	3.70	Multiplet
5' b	3.59	Multiplet
5'-OH	5.04	Doublet of doublets, $J = 5.0, 4.0$

**Table A.8**  $^{13}\text{C}$  chemical shift assignments of 8-oxo-dG in DMF- $d_7$  at 20 °C.

$^{13}\text{C}$ Position	$^{13}\text{C}$ Chemical Shift ( $\delta$ )
2	152.59
4	148.53
5	100.17
6	155.03
8	153.31
1'	83.25
2'	37.64
3'	73.52
4'	89.43
5'	64.30



**Figure A.3** VT  $^1\text{H}$  NMR spectra of ribavirin in DMF- $d_7$  from 20 °C to -60 °C (top to bottom).

**Table A.9**  $^1\text{H}$  chemical shift assignments of ribavirin in DMF- $d_7$  at 20 °C.

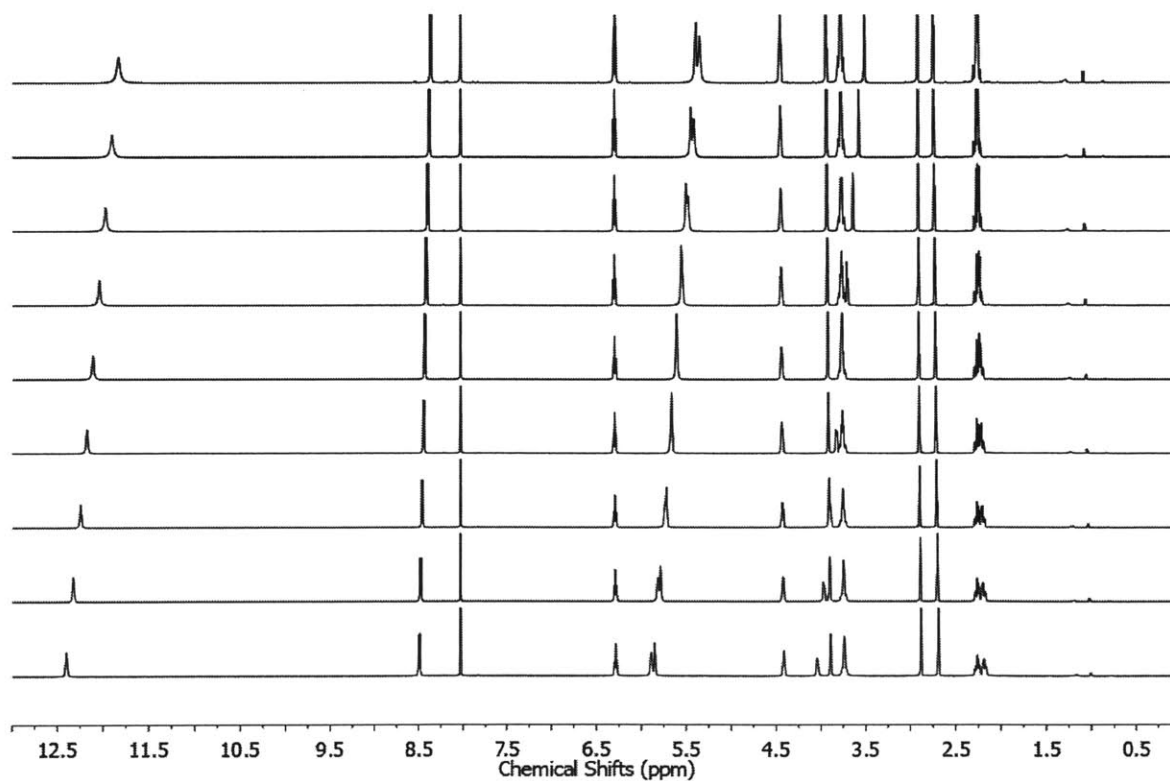
$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
5	8.96	Singlet
6-NH <sub>2</sub> a	7.91	Singlet (br)
6-NH <sub>2</sub> b	7.66	Singlet (br)



1'	5.95	Doublet, J = 3.5
2' a	4.55	Quartet, J = 5.0, 4.0
2'-OH	5.74	Doublet, J = 5.5
3'	4.35	Quartet, J = 5.0
3'-OH	5.25	Doublet, J = 5.5
4'	4.10	Quartet, J = 4.5
5' a	3.78	Multiplet
5' b	3.69	Multiplet
5'-OH	5.06	Triplet, J = 5.5

**Table A.10**  $^{13}\text{C}$  chemical shift assignments of ribavirin in  $\text{DMF-d}_7$  at 20 °C.

$^{13}\text{C}$ Position	$^{13}\text{C}$ Chemical Shift ( $\delta$ )
3	158.93
5	146.00
6	161.78
1'	93.75
2'	76.46
3'	71.76
4'	87.30
5'	62.96



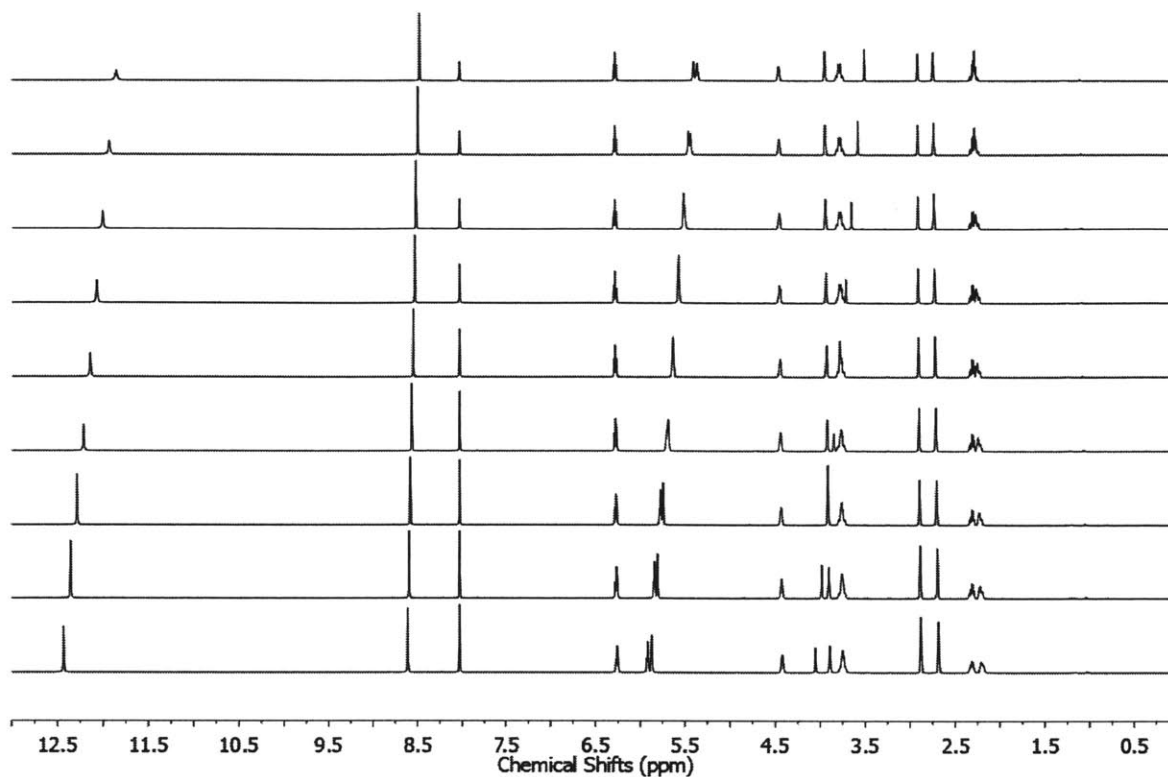
**Figure A.4** VT  $^1\text{H}$  NMR spectra of 5-F-dU in  $\text{DMF-d}_7$  from 20 °C to -60 °C (top to bottom).

**Table A.11**  $^1\text{H}$  chemical shift assignments of 5-F-dU in DMF- $d_7$  at 20 °C.

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
3-NH	11.83	Singlet (br)
6	8.36	Doublet, J = 7.5
1'	6.30	Triplet of doublets, J = 6.5, 2.0
2' a	2.26	Multiplet
2' b	2.26	Multiplet
3'	4.46	Singlet (br)
3'-OH	5.40	Singlet (br)
4'	3.94	Quartet, J = 3.0
5' a	3.79	Multiplet
5' b	3.77	Multiplet
5'-OH	5.35	Singlet (br)

**Table A.12**  $^{13}\text{C}$  chemical shift assignments of 5-F-dU in DMF- $d_7$  at 20 °C.

$^{13}\text{C}$ Position	$^{13}\text{C}$ Chemical Shift ( $\delta$ )	$^{13}\text{C}$ - $^{19}\text{F}$ Coupling Constant (Hz)
2	150.44	N/A
4	158.32	J = 26.3
5	141.56	J = 228.8
6	125.89	J = 33.8
1'	86.21	N/A
2'	41.42	N/A
3'	72.08	N/A
4'	89.34	N/A
5'	62.85	N/A



**Figure A.5** VT  $^1\text{H}$  NMR spectra of 5-Cl-dU in DMF- $d_7$ , from 20 °C to -60 °C (top to bottom).

**Table A.13**  $^1\text{H}$  chemical shift assignments of 5-Cl-dU in DMF- $d_7$ , at 20 °C.

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
3-NH	11.86	Singlet (br)
6	8.48	Singlet
1'	6.29	Triplet, $J = 7.0$
2' a	2.32	Multiplet
2' b	2.29	Multiplet
3'	4.47	Multiplet
3'-OH	5.40	Doublet, $J = 4.0$
4'		Quartet, $J = 3.0$
5' a	3.80	Multiplet
5' b	3.78	Multiplet
5'-OH	5.37	Triplet, $J = 4.5$

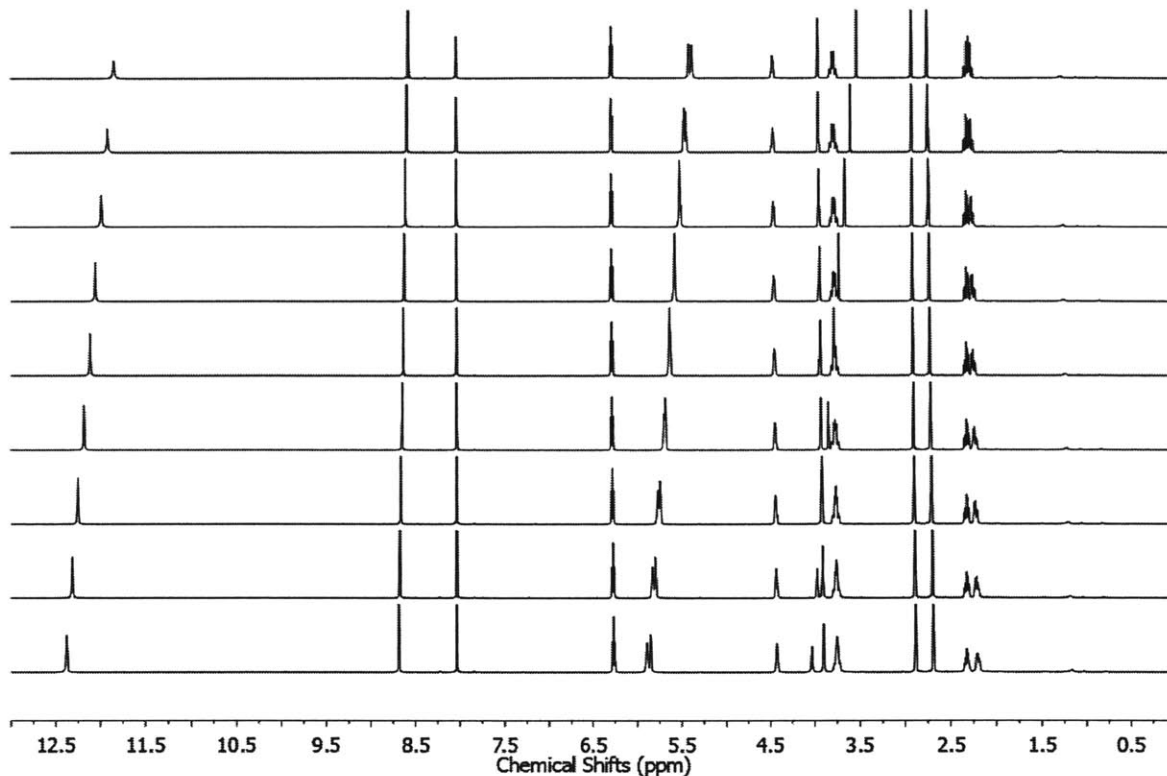


Figure A.6 VT  $^1\text{H}$  NMR spectra of 5-Br-dU in  $\text{DMF-d}_7$  from 20  $^\circ\text{C}$  to -60  $^\circ\text{C}$  (top to bottom).

Table A.14  $^1\text{H}$  chemical shift assignments of 5-Br-dU in  $\text{DMF-d}_7$  at 20  $^\circ\text{C}$ .

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
3-NH	11.83	Singlet (br)
6	8.56	Singlet
1'	6.28	Triplet, $J = 7.0$
2' a	2.32	Multiplet
2' b	2.28	Multiplet
3'	4.46	Multiplet
3'-OH	5.40	Doublet, $J = 4.0$
4'	3.96	Quartet, $J = 3.0$
5' a	3.80	Multiplet
5' b	3.78	Multiplet
5'-OH	5.36	Triplet, $J = 5.0$

Table A.15  $^{13}\text{C}$  chemical shift assignments of 5-Br-dU in  $\text{DMF-d}_7$  at 20  $^\circ\text{C}$ .

$^{13}\text{C}$ Position	$^{13}\text{C}$ Chemical Shift ( $\delta$ )
--------------------------	---

2	151.22
4	160.36
5	96.87
6	141.65
1'	86.49
2'	41.77
3'	71.93
4'	89.42
5'	62.66

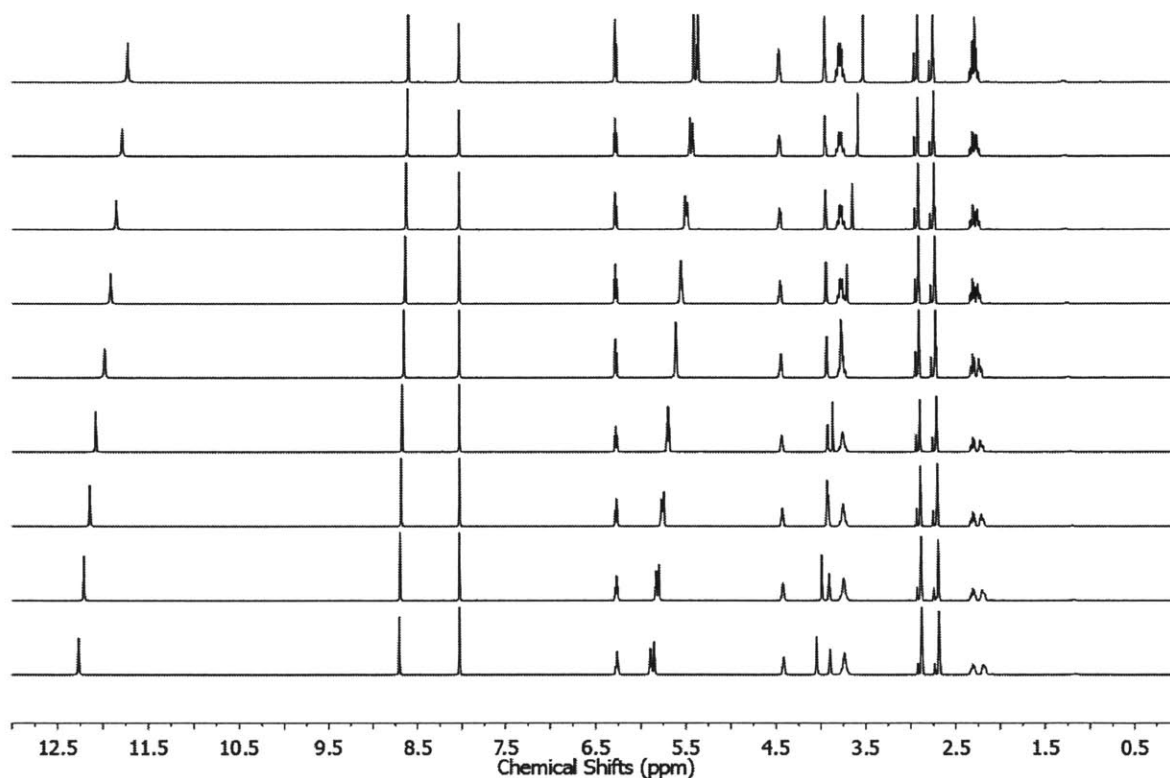


Figure A.7 VT  $^1\text{H}$  NMR spectra of 5-I-dU in  $\text{DMF-d}_7$  from 20 °C to -60 °C (top to bottom).

Table A.16  $^1\text{H}$  chemical shift assignments of 5-I-dU in  $\text{DMF-d}_7$  at 20 °C.

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
3-NH	11.73	Singlet (br)
6	8.59	Singlet
1'	6.28	Triplet, $J = 7.0$
2' a	2.31	Multiplet
2' b	2.26	Multiplet
3'	4.46	Multiplet

3'-OH	5.41	Doublet, J = 4.5
4'	3.95	Quartet, J = 3.0
5' a	3.80	Multiplet
5' b	3.76	Multiplet
5'-OH	5.36	Triplet, J = 5.0

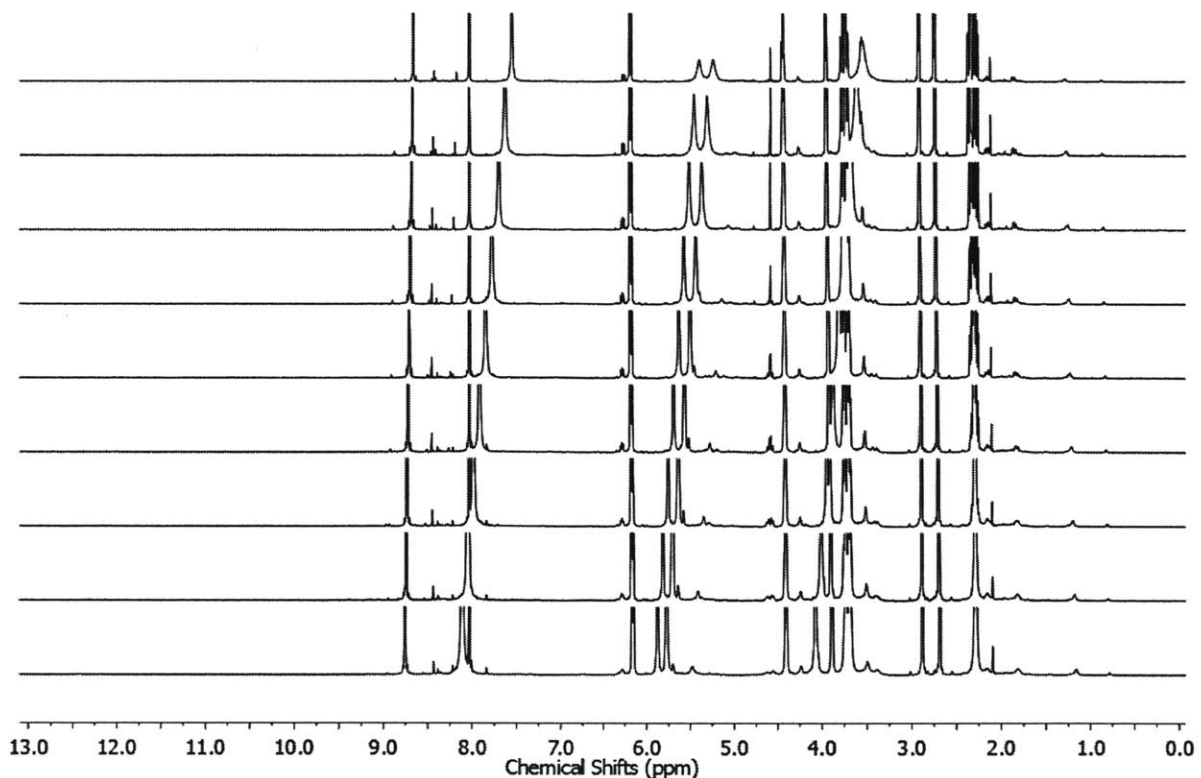


Figure A.8 VT  $^1\text{H}$  NMR spectra of 5-aza-dC in  $\text{DMF-d}_7$ , from 20 °C to -60 °C (top to bottom).

Table A.17  $^1\text{H}$  chemical shift assignments of 5-aza-dC in  $\text{DMF-d}_7$ , at 20 °C.

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
4-NH <sub>2</sub> a	7.52	Singlet (br)
4-NH <sub>2</sub> b	7.52	Singlet (br)
6	8.66	Singlet
1'	6.19	Triplet, J = 6.5
2' a	2.28	Multiplet
2' b	2.35	Multiplet
3'	4.46	Multiplet
3'-OH	5.37	Singlet (br)
4'	3.96	Quartet, J = 3.5
5' a	3.78	Multiplet

5' b	3.73	Multiplet
5'-OH	5.21	Singlet (br)

**Table A.18**  $^{13}\text{C}$  chemical shift assignments of 5-aza-dC in DMF- $d_7$  at 20 °C.

$^{13}\text{C}$ Position	$^{13}\text{C}$ Chemical Shift ( $\delta$ )
2	154.65
4	167.77
6	157.21
1'	86.94
2'	37.15
3'	71.79
4'	89.47
5'	62.66

**Table A.19**  $^1\text{H}$  chemical shift assignments of DHdC in DMF- $d_7$  at 20 °C.

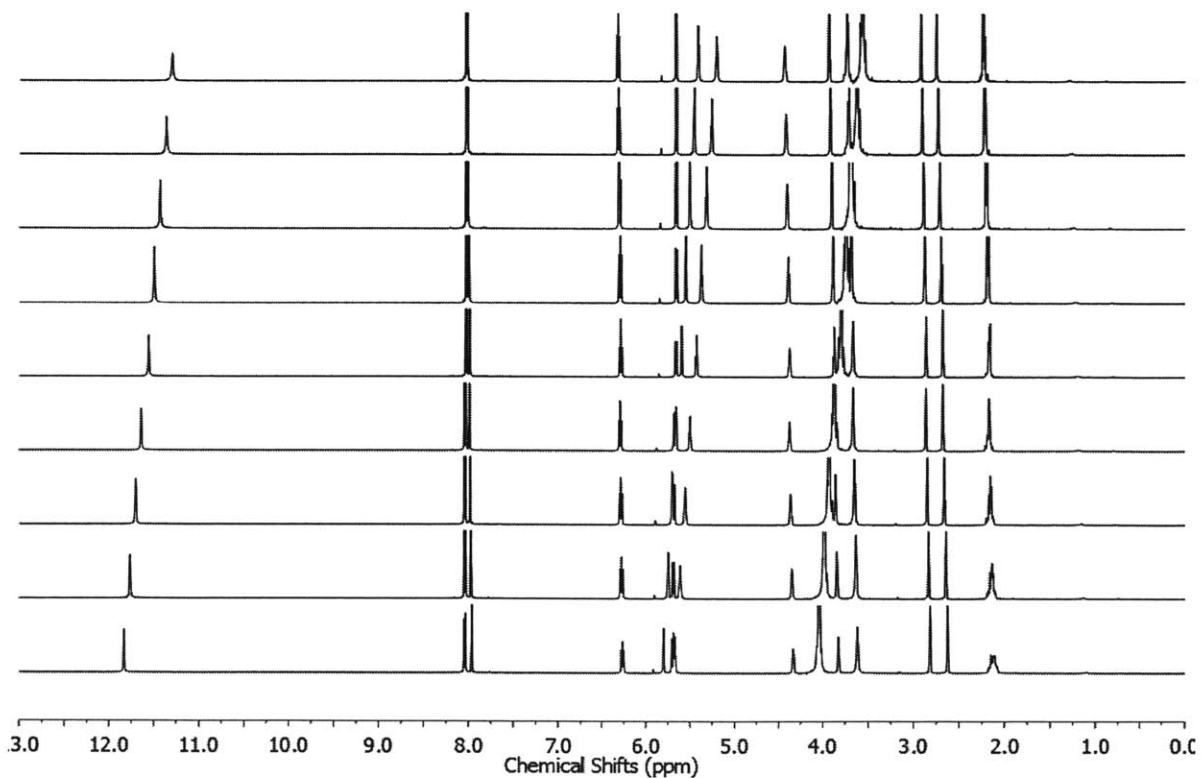
$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
5 a	2.54	Multiplet
5 b	2.51	Multiplet
6 a	3.41	Multiplet
6 b	3.25	Multiplet
1'	6.38	Triplet
2' a	2.09	Multiplet
2' b	1.84	Multiplet
3'	4.27	Multiplet
3'-OH	5.16	Singlet (br)
4'	3.71	Multiplet
5' a	3.58	Multiplet
5' b	3.58	Multiplet
5'-OH	4.86	Singlet (br)

**Table A.20**  $^{13}\text{C}$  chemical shift assignments of DHdC in DMF- $d_7$  at 20 °C.

$^{13}\text{C}$ Position	$^{13}\text{C}$ Chemical Shift ( $\delta$ )
2	N/A
4	N/A
5	N/A
6	36.91
1'	85.25
2'	37.38
3'	72.49
4'	87.19
5'	63.74

**Table A.21**  $^1\text{H}$  chemical shift assignments of DHdU in  $\text{DMF-d}_7$  at  $20\text{ }^\circ\text{C}$ .

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
3-NH	10.16	Singlet (br)
5 a	2.60	Multiplet
5 b	2.60	Multiplet
6 a	3.54	Multiplet
6 b	3.44	Multiplet
1'	6.28	Quartet, $J = 6.5, 2.0$
2' a	2.19	Multiplet
2' b	1.93	Multiplet
3'	4.32	Multiplet
3'-OH	5.26	Doublet, $J = 4.0$
4'	3.77	Multiplet
5' a	3.61	Multiplet
5' b	3.61	Multiplet
5'-OH	4.92	Triplet, $J = 5.5$



**Figure A.9** VT  $^1\text{H}$  NMR spectra of 2'-deoxyuridine in  $\text{DMF-d}_7$  from  $20\text{ }^\circ\text{C}$  to  $-60\text{ }^\circ\text{C}$  (top to bottom).



**Table A.22**  $^1\text{H}$  chemical shift assignments of 2'-deoxyuridine in  $\text{DMF-d}_7$  at 20 °C.

$^1\text{H}$ Position	$^1\text{H}$ Chemical Shift ( $\delta$ )	$^1\text{H}$ Peak Multiplicity and Coupling Constant (Hz)
3-NH	11.28	Singlet (br)
5	5.66	Doublet, J = 8.0
6	8.02	Doublet, J = 8.0
1'	6.32	Triplet, J = 6.5
2' a	2.24	Multiplet
2' b	2.23	Multiplet
3'	4.44	Multiplet
3'-OH	5.37	Doublet, J = 4.0
4'	3.93	Quartet, J = 3.5, 3.0
5' a	3.74	Multiplet
5' b	3.73	Multiplet
5'-OH	5.17	Triplet, J = 5.0

**Table A.23**  $^{13}\text{C}$  chemical shift assignments of 2'-deoxyuridine in  $\text{DMF-d}_7$  at 20 °C.

$^{13}\text{C}$ Position	$^{13}\text{C}$ Chemical Shift ( $\delta$ )
2	151.91
4	164.34
5	102.81
6	141.68
1'	85.78
2'	41.26
3'	72.29
4'	89.21
5'	63.07