

**DIRECTED EVOLUTION AND PATHWAY ENGINEERING FOR  
NUCLEOTIDE ANALOGUE BIOSYNTHESIS**

By

Robert Scism

Dissertation

Submitted to the Faculty of the  
Graduate School of Vanderbilt University  
in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry

August, 2010

Nashville, Tennessee

Approved:

Brian Bachmann

Lawrence Marnett

David Wright

Richard Armstrong

Michael Stone

***To my wonderful children, Isabella and Lucas***

*May you pursue life with a thirst for adventure and a heart full of love*

## ACKNOWLEDGEMENTS

First I would like to express my gratitude to Vanderbilt University and the Chemistry Department for allowing me the opportunity to study at this great institution. Throughout the years of classes, research, and teaching assistance at Vanderbilt, I have collected a wealth of memories and experiences that cannot be matched. Completion of this chapter in my life fulfills not only a lifelong dream, but has also prepared me with the confidence and self-reliance to succeed in those adventures which lie ahead.

I would like to thank my advisor, Dr. Brian Bachmann, for his guidance throughout my research, and for expecting only the best from me. Brian's never ending enthusiasm for science, and constant input of new ideas kept the work in our lab exciting. I would also like to thank the members of my Ph.D. committee, Dr. David Wright, Dr. Larry Marnett, Dr. Richard Armstrong, and Dr. Mike Stone, for contributing a wise and experienced perspective, and for keeping me focused on the big picture.

I owe many thanks to Dawn Overstreet and Dr. Wade Calcutt for guidance in the mass spectrometry lab. You both made what seemed like an overwhelming and intimidating task achievable. Also, thanks to Dr. Don Stec for assistance in the NMR lab. Your help with the advanced water peak suppression methods, and general instrument assistance was invaluable.

Members of the Bachmann Lab, past and present, have been my family away from home. I enjoyed our daily news discussions and political debates, and festive pre-group meeting dinners. I especially enjoyed our canoeing adventures on the Harpeth, camping excursions in very hot summer nights, and refreshing waterfall swims at Rock Island. Thank you all for your companionship and camaraderie.

I am especially grateful to my family for a lifetime of love and support. You have all been so wonderful to me. To my sisters; Jamie, thank you for acting as my insightful editor, for all of your career advice, and for your always flattering complements and belief in my abilities. Jennifer, thank you being my second mom, and for always looking out for me, and your constant support. Elizabeth, you are my very special big sis, and have always been a laughing, loving, light in the world. I will always cherish all of those fun dinners and backyard cookouts our families shared together. Mom and Dad, thank you for instilling my strong foundation of learning, and always teaching me to question the world around me. Thank you for teaching me to see a job through to the end, and to have a never ending thirst for knowledge.

A special thanks to my wife, Almary, for always listening to my gripes and recap of the daily grind. Thanks for being there, and thanks for listening, and for laughing with me. Your emotional support and words of encouragement helped me get through the most difficult of times. Finally, I am thankful for my children, Isabella and Lucas, for all of those wonderful hugs, smiles, and laughs. You have given purpose to my life.



## TABLE OF CONTENTS

	Page
DEDICATION .....	ii
ACKNOWLEDGEMENTS .....	iii
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
LIST OF SCHEMES .....	xii
LIST OF ABBREVIATIONS .....	xiii
Chapter	
I. NUCLEOSIDE AND NUCLEOTIDE PHARMACEUTICALS .....	1
Introduction .....	1
Ribavirin .....	7
Chemical Synthesis.....	8
Enzymatic Synthesis .....	10
Nucleosides .....	11
Nucleotides .....	21
Directed Evolution .....	25
Mutation Methods.....	26
Selection.....	33
Screening.....	37
Applications.....	43
Nucleotide Biosynthesis by Directed Evolution .....	43
Dissertation Statement .....	46
References .....	47
II. DIRECTED EVOLUTION AND IN VIVO SCREENING OF HPRT FOR RMP NUCLEOTIDE BIOSYNTHESIS	
Introduction .....	64
Phosphoribosyl Transferases .....	64
Ribavirin and Ribavirin Monophosphate (RMP) .....	69
Inosine Monophosphate Dehydrogenase (IMPDH).....	71

Synopsis of Proposed Screening Method for HPRT Evolution.....	74
Results.....	75
Conclusions.....	85
Experimental Methods.....	86
References.....	92
III. APPLICATION OF AN IMPROVED BIOCATALYST FOR NUCLEOTIDE ANALOGUE BIOSYNTHESIS.....	97
Introduction.....	97
Results.....	98
Discussion.....	103
Activity assay.....	103
Requirements of a Nucleobase.....	104
Experimental Methods.....	106
References.....	112
IV. PATHWAY ENGINEERING FOR CASCADE SYNTHESIS OF NUCLEOTIDE ANALOGUES IN A CROSS-LINKED ENZYME AGGREGATE.....	114
Introduction.....	114
Ribokinase.....	115
Phosphoribosyl Pyrophosphate Synthetase (PPS).....	118
Immobilization techniques.....	119
Results.....	123
Discussion.....	130
ATP Regeneration.....	130
Nucleotide Purification.....	130
CLEA Stability Rationale.....	132
Glutaraldehyde cross-linking mechanism.....	133
Conclusion.....	135
Experimental Methods.....	135
References.....	142
V. DISSERTATION SUMMARY AND FUTURE DIRECTION.....	149
Synopsis.....	149
Significance.....	150
Future directions.....	152
Directed evolution.....	152
CLEA Improvements and Applications.....	155
References.....	158

## Appendix

A. PERCENT CONVERSION NMR FROM CHAPTER III .....	161
B. PURIFIED PRODUCT NMR FROM CHAPTER III.....	172
C. PURIFIED PRODUCT NMR FROM CHAPTER IV .....	187

## LIST OF TABLES

Table		Page
I-1.	Collection of enzymatic syntheses of nucleosides using PNP. Abbreviations: PNP purine nucleoside phosphorylase, UP uridine phosphorylase, TP thymidine phosphorylase, ADA adenosine deaminase, PPM phosphopentomutase .....	13
I-2.	Whole cell catalyzed nucleoside syntheses .....	18
II-1.	Affinity for various purine nucleobases by HPRT and XGPRT from <i>E. coli</i> . Adapted from original data by Guddat, L. W., <i>et al.</i> .....	68
II-2.	Correlation of mutation rate with manganese chloride concentration .....	77
II-3.	Kinetic parameters of hypoxanthine and TCA turnover by wt HPRT and 8B3PRT .....	82
III-1.	Percent conversion of select nucleobase analogues. Adapted from Scism, R. A. <i>et al.</i> .....	102
III-2.	Concentration of nucleobase stock solutions and final concentrations in wells. Adapted from Scism, R. A. <i>et al.</i> .....	108
III-3.	Percent conversion values and base analogue concentrations used in reactions, adapted from Scism, R. A. <i>et al.</i> .....	111
IV-1.	Nucleotide analogues prepared with the 5-enzyme CLEA, purified by anion exchange. This table is copyright of Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission. Originally published by Scism, R. A. <i>et al.</i> .....	129
IV-2.	Bacterial strains and plasmids used. Adapted from Scism, R. A. <i>et al.</i> .....	136
IV-3.	Primers used in cloning of nucleotide biosynthetic genes. Adapted from Scism, R. A. <i>et al.</i> .....	137

## LIST OF FIGURES

Figure	Page
I-1. Nucleoside analogues for HSV treatment.....	2
I-2. Nucleoside analogue and nucleotide analogue reverse transcriptase inhibitors used in HIV treatment. Each of these drugs has modifications in the ribosyl portion of the nucleoside .....	4
I-3. Hepatitis B reverse transcriptase inhibitors .....	5
I-4. Chemotherapeutic nucleoside, nucleotide, and nucleobase analogues .....	6
I-5. Ribavirin tautomers - analogues of guanosine and adenosine.....	7
I-6. Numbering system for purine nucleosides.....	9
I-7. Nucleobase elaboration for synthesis of wyosine and 5-azacytidine. Reagents: (i) CNBr; (ii) NaOEt, EtOH; (iii) BrCH <sub>2</sub> COCH <sub>3</sub> . Adapted from Blackburn, <i>et al</i> .....	9
I-8. Original method of ribavirin synthesis as reported by Witkowski .....	10
I-9. Selected examples of NdRT catalyzed 2'-deoxyribonucleoside analogues .....	19
I-10. Base analogues incorporated into monophosphate nucleotides catalyzed by liver acetone powders .....	22
I-11. DNA Shuffling; adapted from Valetti, <i>et al</i> .....	29
I-12. Staggered extension protocol (StEP); adapted from Zhao, <i>et al</i> .....	30
I-13. Incremental Truncation (ITCHY); adapted from Valetti, <i>et al</i> .....	31
I-14. Sequence homology-independent recombination (SHIPREC); adapted from Zhao, <i>et al</i> .....	32
II-1. Comparison of structural similarities in purine PRTs, adapted from Craig, S. P., <i>et al</i> . From left to right; APRT from <i>Leishmani donovani</i> , HPRT from <i>Homo sapiens</i> , and XPRT from <i>E. coli</i> . Images were rendered from crystal structures obtained by Keough, D. T., Phillips, C. L., and Vos, S. using UCSF Chimera software .....	66
II-2. Open and closed forms of HPRT from <i>Trypanosoma cruzi</i> . Adaptations rendered from original crystal structures by Focia, <i>et al</i> , using UCSF Chimera software .....	67

II-3.	Ribavirin tautomers as analogues of guanosine and adenosine .....	69
II-4.	Ternary complex of <i>H. sapiens</i> IMPDH type II, as adapted from original crystal structures obtained by Colby, T. D., <i>et al</i> , using UCSF Chimera software .....	72
II-5.	Overlap crystal structures of IMPDH of <i>T. foetus</i> bound with IMP and with RMP, adapted from original crystal structures by Prorise, G. L., <i>et al</i> , using UCSF Chimera software. Ligand positions are indiscernable due to identical positioning .....	74
II-6.	TCA toxicity to <i>E. coli</i> containing wt HPRT in pET28a plasmid.....	76
II-7.	Guanosine complementation in <i>E. coli</i> with wt HPRT .....	77
II-8.	Hypoxanthine consumption assay for percent activity of a library.....	78
II-9.	Growth rate comparison of 8B3PRT mutant and wt HPRT in 115 $\mu$ M TCA.....	80
II-10.	Complementation effect of guanosine on 8B3 mutant in comparison with wt HPRT and 8B3 without guanosine .....	80
II-11.	HPLC and MS of biosynthesized RMP: Images on left illustrate the appearance of the RMP peak over time. Mass spectrum of the preparatory fraction of the RMP peak is to the right .....	81
II-12.	Kinetic measurements of hypoxanthine and TCA turnover by wt HPRT and 8B3PRT .....	82
II-13.	Intracellular comparison of RMP in wt and 8b3. The second compound to elute is the isobaric UMP naturally present in the cell.....	83
II-14.	Theoretical model of 8B3PRT with bound RMP, generated with UCSF Chimera software, based on crystal structure of wild type HPRT with bound IMP, as determined by Guddat, L. W. <i>et al</i> . Mutated residues V157A and Y170H are highlighted.....	84
II-15.	Active site point mutation locations shown with native amino acid residues and theoretical bound RMP with rotated amide bond, created with UCSF Chimera software. Adapted from original crystal structure of HPRT from <i>E. coli</i> , as determined by Guddat, L. W. <i>et al</i> .....	85
III-1.	Turnover of nucleobase analogues from the substrate library with relatively lower turnover, plotted on a smaller scale for visibility. Adapted from Scism, R. A. <i>et al</i> .....	100
III-2.	Turnover of nucleobase analogues from substrate library: Highly active substrates listed in order of activity with 8B3PRT. Adapted from Scism, R. A. <i>et al</i> .....	101

III-3.	Key active site residues of HPRT, adapted from crystal structure of wt HPRT from <i>E. coli</i> , as obtained by Guddat <i>et al.</i> . Modeled with UCSF Chimera .....	105
IV-1.	Ribokinase with bound ADP and ribose. Adaptation of crystal structure obtained by Sigrell, J. A. <i>et al.</i> , rendered by UCSF Chimera software.....	116
IV-2.	SDS-PAGE gel of purified enzymes and CLEA; lane 1, marker; lane 2, 1.1 µg RK; lane 3, 0.9 µg PPS; lane 4, 1.8 µg 8B3; lane 5, 1.5 µg PK; lane 6, 1.3 µg AK; lane 7, 690 µg CLEA aggregate. This image copyright of Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission. Originally published by Scism, R. A. <i>et al.</i> .....	124
IV-3.	CLEA preparation - PK (pyruvate kinase), AK (adenylate kinase), RK (ribokinase), PPS (phosphoribosyl pyrophosphate synthetase), 8B3 (8B3PRT evolved biocatalyst), BSA (bovine serum albumin). This image copyright of Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission. Originally published by Scism, R. A. <i>et al.</i> .....	125
IV-4.	Reaction progress by HPLC using ion pairing reagent and reverse phase chromatography. Adapted from Scism, R. A. <i>et al.</i> .....	125
IV-5.	Effect of glutaraldehyde concentration on activity with purified enzymes. Shown in dark brown is activity of a CLEA particle created with non-purified cell lysate. Adapted from Scism, R. A. <i>et al.</i> .....	126
IV-6.	Stability of CLEA vs. unbound enzymes at 37°C. Adapted from Scism, R. A. <i>et al.</i> .....	127
IV-7.	CLEA residual activity after catalysis and subsequent washing. CLEAs prepared with and without added bovine serum albumin. Adapted from Scism, R. A. <i>et al.</i> .....	128
IV-8.	Glutaraldehyde polymers .....	134
IV-9.	Protein adducts with glutaraldehyde polymers including Michael addition and anti-Markownikoff products.....	134
V-1.	Nucleotide analogue inhibitors of inosine monophosphate dehydrogenase .....	154

## LIST OF SCHEMES

Scheme	Page
I-1. General scheme of the one-step nucleoside 2-deoxyribosyltransferase (top) and the two-step nucleoside phosphorylase reaction with ribose-1-phosphate intermediate (bottom). B = purine or pyrimidine nucleobase .....	12
I-2. General method of non-natural nucleotide biosynthesis using salvage enzymes, adapted and combined from Scott, Williamson, and Hennig. Ribokinase (RK), phosphoribosylpyrophosphate synthetase (PPS), phosphoribosyl transferase (PRT), nucleotide monophosphate (NMP), nucleotide diphosphate (NDP), nucleotide triphosphate (NTP), phosphoenol pyruvate (PEP), creatine phosphate (CP), pyrophosphate (PPi).....	23
I-3. dATP recycling enzymes: adenylate kinase (AK), pyruvate kinase (PK), enolase and phosphoglycerate mutase (PGM). Adapted from Da Costa, C. P., <i>et al</i> .....	24
I-4. Biosynthesis of 5-fluorocytidine-5'-triphosphate from 5-fluorocytidine nucleoside using uridine kinase (UK), nucleoside monophosphate kinase (NMK), and pyruvate kinase (PK). Adapted from Hennig, M., <i>et al</i> .....	25
II-1. Proposed mechanism of IMPDH catalyzed conversion of IMP to XMP, adapted from Gan, L. <i>et al</i> .....	73
II-2. <i>In vivo</i> screening method for RMP production. TCA added to growth media passively transports into the cell where it condenses with endogenous PRPP via HPRT catalysis. Resulting RMP inhibits IMPDH effectively inhibiting GMP production and cell growth. Cells grown with guanine or guanosine supplementation can survive the inhibition.....	75
IV-1. Three-step pathway from ribose to nucleotide utilizing ribokinase (RK), phosphoribosyl pyrophosphate synthetase (PPS), and 8B3PRT, with ATP regeneration enzymes adenylate kinase (AK) and pyruvate kinase (PK). Adapted from Scism, R. A. <i>et al</i> .....	115
IV-2. Cross-linking by Schiff base formation, condensed adaptation from Farris, S. <i>et al</i> .....	133
V-1. ATP recycling schemes. Left; original method used for nucleotide pathway using adenylate kinase (AK) and pyruvate kinase (PK). Middle; polyphosphate kinase (PPK) and polyphosphate:AMP phosphotransferase (PAP). Right; Adenylate kinase (AK) and PAP coupled regeneration. Adapted in part from Zhao, <i>et al</i> .....	157



## LIST OF ABBREVIATIONS

2F-ADP	2-fluoroadenosine-5'-diphosphate
2F-AMP	2-fluoroadenosine-5'-monophosphate
2F-ATP	2-fluoroadenosine-5'-triphosphate
8-azaGMP	8-azaguanine monophosphate
ADA	adenosine deaminase
ADP	adenosine diphosphate
AIDS	acquired immune deficiency syndrome
AK	adenylate kinase
AMP	adenosine monophosphate
amu	atomic mass units
APRT	adenine phosphoribosyl transferase
ATP	adenosine triphosphate
AZT	3'-azido-3'-deoxythymidine
BSA	bovine serum albumin
CLEA	cross-linked enzyme aggregate
CTP	cytidine triphosphate
DEAE	diethylaminoethyl
DHR	degenerate homoduplex recombination
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dATP	deoxyadenosine triphosphate
EDTA	ethylenediaminetetraacetic acid

EICAR	1- $\beta$ -D-ribofuranosylimidazole-4-carboxamide
epPCR	error prone polymerase chain reaction
ESI	electrospray ionization
FACS	fluorescence-activated cell sorting
FPLC	fast protein liquid chromatography
GMP	guanosine monophosphate
GSSM	gene site saturation mutagenesis
GTP	guanosine triphosphate
HBV	hepatitis B virus
HGPRT	hypoxanthine-guanine phosphoribosyl transferase
HIV	human immunodeficiency virus
HPLC	high pressure liquid chromatography
HPRT	hypoxanthine phosphoribosyl transferase
HRMS	high resolution mass spectrometry
HSV	herpes simplex virus
hx	hypoxanthine
IC <sub>50</sub>	half maximal inhibitory concentration
IMP	inosine monophosphate
IMPDH	inosine monophosphate dehydrogenase
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
ITCHY	include incremental truncation of hybrid enzymes
IVC	<i>in vitro</i> compartmentalization
kb	kilobase
kDa	kiloDalton

LB	Luria-Bertani
LCMS	liquid chromatography mass spectrometry
mg	milligram
μg	microgram
MHz	megahertz
mL	milliliter
μL	microliter
mM	millimolar
μM	micromolar
MP	monophosphate
MWCO	molecular weight cut-off
NAD <sup>+</sup>	nicotinamide adenine dinucleotide (oxidized)
NADH	nicotinamide adenine dinucleotide (reduced)
NDP	nucleotide diphosphate
NdRT	nucleoside 2'-deoxyribosyltransferase
ng	nanogram
nm	nanometer
nM	nanomolar
NMK	nucleoside monophosphate kinase
NMP	nucleotide monophosphate
NMR	nuclear magnetic resonance
NP	nucleoside phosphorylase
NTP	nucleotide triphosphate
OD	optical density

OPRT	orotidine phosphoribosyl transferase
PAP	polyphosphate:AMP phosphotransferase
PCR	polymerase chain reaction
PDA	photodiode array
PEP	phosphoenol pyruvate
Pi	inorganic phosphate
PK	pyruvate kinase
PNP	purine nucleoside phosphorylase
PPi	pyrophosphate
PPK	polyphosphate kinase
PPM	phosphopentomutase
ppm	parts per million
PPS	phosphoribosyl pyrophosphate synthetase
PRPP	phosphoribosyl pyrophosphate
PRT	phosphoribosyl transferase
psi	pounds per square inch
PVA	polyvinyl alcohol
r-1-p	ribose-1-phosphate
r-5-p	ribose-5-phosphate
RACHITT	random chimeragenesis on transient templates
rcf	relative centrifugal force
RDP	ribavirin diphosphate
RK	ribokinase
RMP	ribavirin monophosphate

RNA	ribonucleic acid
rpm	rotations per minute
RTP	ribavirin triphosphate
SHIPREC	sequence homology-independent recombination
SEM	scanning electron micrograph
StEP	Staggered extension protocol
TCA	1,2,4-triazole-3-carboxamide
TK	thymidine kinase
TP	thymidine phosphorylase
TRIS	tris(hydroxymethyl)aminomethane
UK	uridine kinase
UMP	uridine monophosphate
UP	uridine phosphorylase
UPRT	uracil phosphoribosyl transferase
UTP	uridine triphosphate
UV	ultra violet
V	volts
wt	wild type
XGPRT	xanthine-guanine phosphoribosyl transferase
XMP	xanthine monophosphate

## CHAPTER I

### NUCLEOSIDE AND NUCLEOTIDE PHARMACEUTICALS

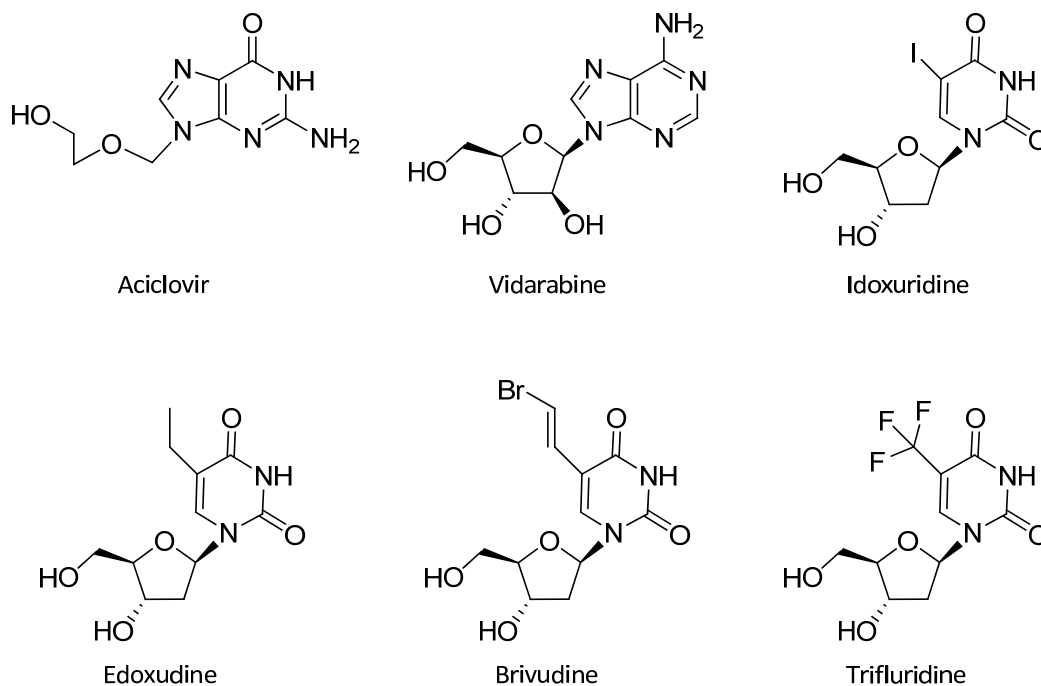
#### Introduction

Nucleoside and nucleotide analogues are vital elements of the frontline defense against viral disease. Immunization programs provide invaluable protection against a significant portion of viral diseases, to the point of near eradication. Measles virus has been eliminated from North America through rigorous vaccination campaigns. Poliovirus has been eliminated from both North and South Americas, and smallpox has been eliminated worldwide.<sup>1</sup> However, it is increasingly apparent that vaccines cannot be manufactured for all viral diseases, and even when a vaccine is available, complete coverage is not possible. In these cases, successful treatment depends on the efficacy of antiviral drugs, and nucleoside and nucleotide analogues are often the most effective treatment available.

Many prominent viral diseases are treated with nucleoside analogues including human immunodeficiency virus (HIV), hepatitis B virus (HBV) and hepatitis C virus, herpes simplex virus (HSV), varicella zoster virus, Epstein Barr virus, cytomegalovirus, and human papilloma virus. To narrow the scope of this discussion, examples of nucleoside and nucleotide analogue treatment for HIV, HBV, and HSV are covered here.

The center for disease control reports one out of five adolescents and adults are infected with HSV. The total number of people who were living with HSV-2 infection worldwide in 2003 was estimated to be 536 million.<sup>2</sup> As there is no cure for HSV, nor any vaccine providing complete immunity, prevention and outbreak management by nucleoside analogues is the only option. Figure I-1 shows several nucleosides for HSV treatment, including analogues of

thymidine, guanosine, and adenosine. These molecules are in fact prodrugs, becoming activated after phosphorylation of the 5' ribosyl hydroxyl upon entering the body. This transformation is catalyzed by thymidine kinase, nucleoside diphosphate kinase, or other similar cellular kinases, resulting in a mixture of mono, di-, and tri-phosphate nucleotides. For non-retroviral diseases, such as HSV, the antiviral activity is a result of the inhibition of DNA or RNA polymerases in infected cells by the phosphorylated nucleotide. Increased activity of viral kinases for these analogues compared to host cell kinases makes treatment with these compounds possible.



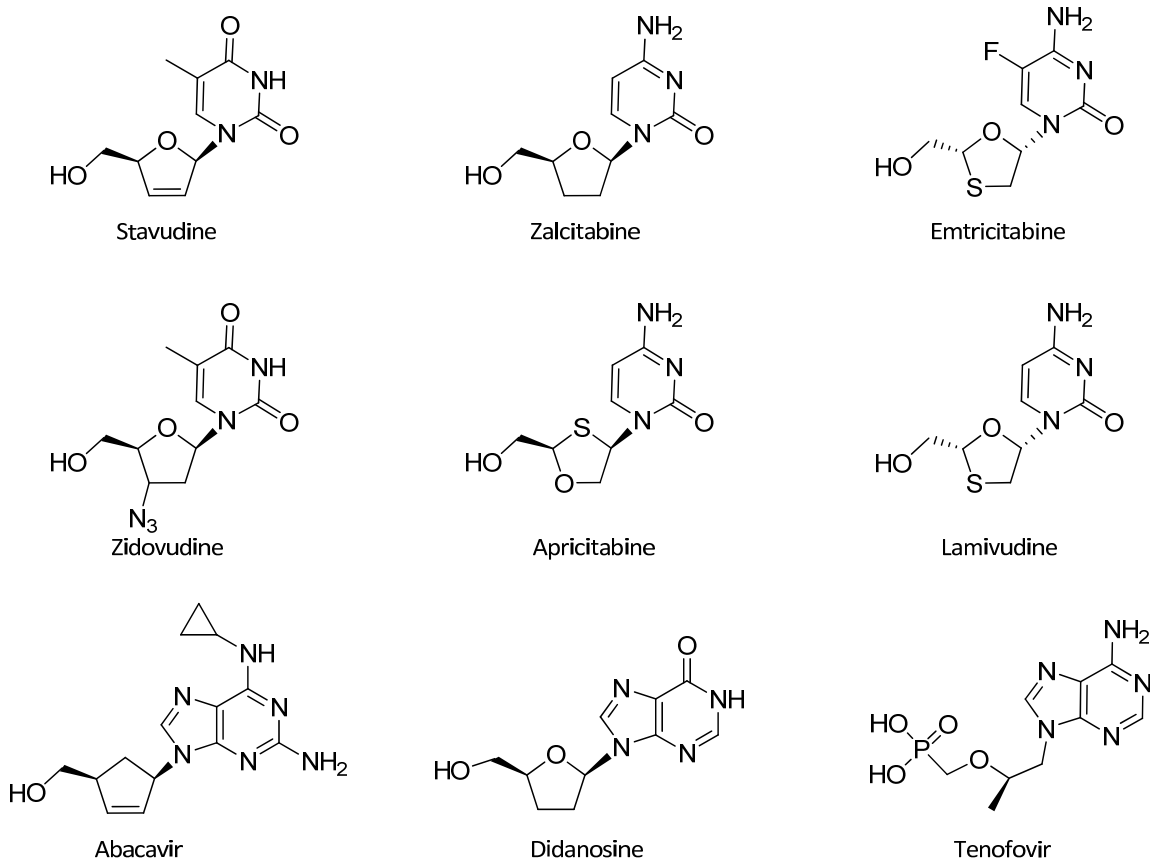
**Figure I-1: Nucleoside analogues for HSV treatment.**

HIV is a significantly more dangerous and deadly disease. It infects cells of the host immune system, resulting in the progressive deterioration of the body's ability to fight off infection and disease.<sup>2</sup> Worldwide, 33 million people live with HIV/AIDS, and an estimated 2.7 million people were newly infected with the virus in 2007. HIV is ultimately fatal, killing an

estimated 2 million people each year.<sup>2</sup> Nucleoside analogues are an important component in many of the numerous combination drug therapies currently used in treatment. Nucleoside analogues used against HIV include thymidine analogues Zidovudine and Stavudine, cytidine analogues Zalcitabine, Emtricitabine, Apricitabine, and Lamivudine, and purine analogues Abacavir and Didanosine. (Figure 1-2) Tenofovir, an analogue of adenosine monophosphate, is one of the less common nucleotide analogues that do not require intracellular phosphorylation.

Analogues for HIV treatment have a different inhibition mechanism than those prescribed for HSV. As with HSV, each HIV nucleoside is a prodrug, phosphorylated by kinases to mono-, di-, and tri-phosphates. Some are ribonucleotide reductase inhibitors in their diphosphate forms. All, as triphosphates, are reverse transcriptase inhibitors (unique to retroviruses) and chain terminating substrates in DNA synthesis.

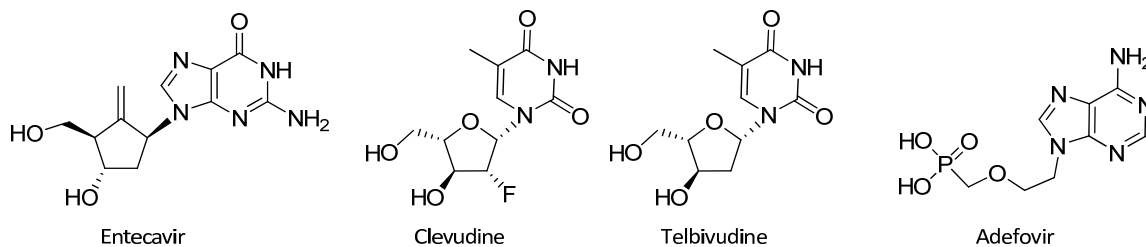




**Figure I-2: Nucleoside analogue and nucleotide analogue reverse transcriptase inhibitors used in HIV treatment. Each of these drugs has modifications in the ribosyl portion of the nucleoside.**

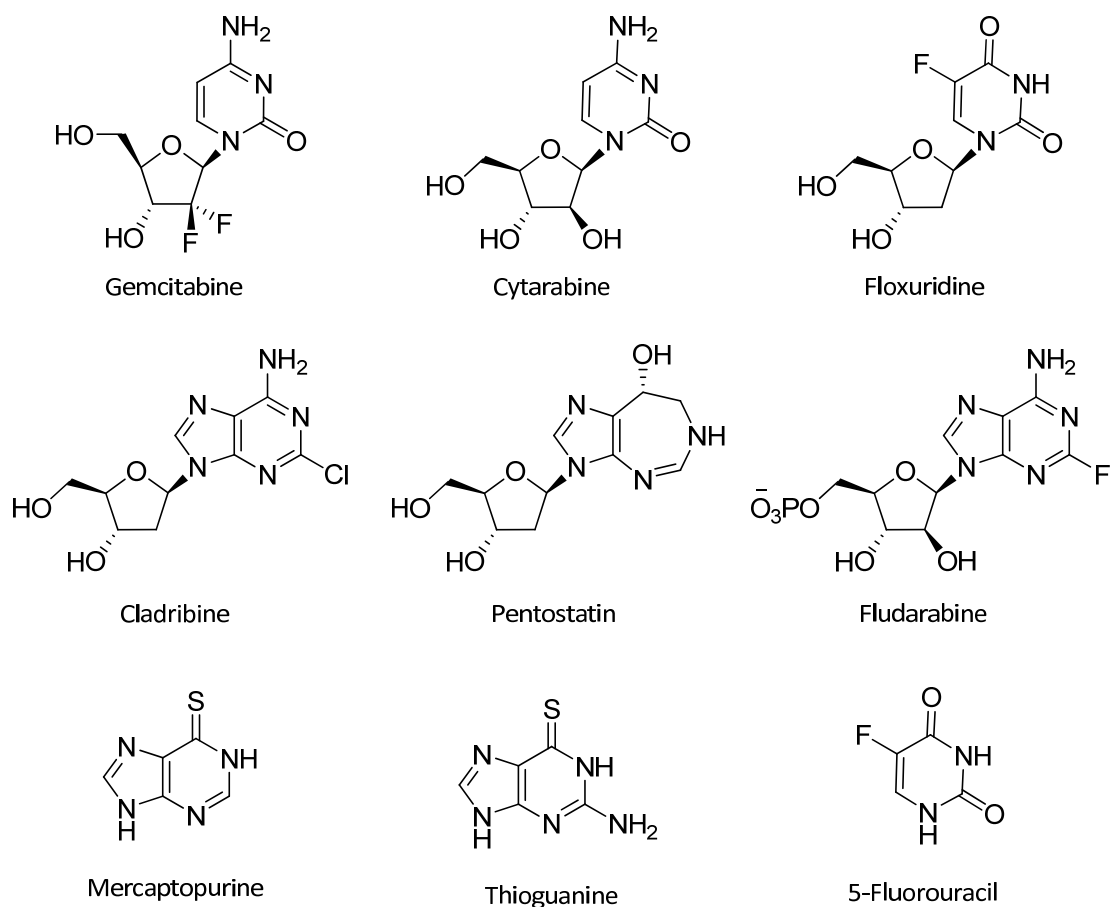
The hepatitis B virus is 50 to 100 times more infectious than HIV. About 2 billion people worldwide have been infected with the virus and about 350 million live with chronic infection. Fortunately, an effective vaccine against hepatitis B has been available since 1982, yet an estimated 600,000 people die each year due to the acute and chronic consequences of the virus.<sup>2</sup> Nucleoside analogues for HBV are also reverse transcriptase inhibitors upon phosphorylation, but not all are necessarily chain terminating. The examples given here exhibit modifications to the sugar moiety with retention of original nucleobase. These include the previously mentioned Lamivudine, as well as Entecavir, Clevudine, and Telbivudine. (Figure I-3)

Analogues of phosphorylated nucleotides include Tenofovir, previously mentioned as an HIV treatment, and Adefovir.



**Figure I-3: Hepatitis B reverse transcriptase inhibitors.**

Nucleoside and nucleotide analogues are not only antiviral pharmaceuticals, but also included in the chemotherapeutic regimen for many cancers including non-Hodgkin's lymphoma, leukemia, and cancers of the breast and colon. (Figure I-4)



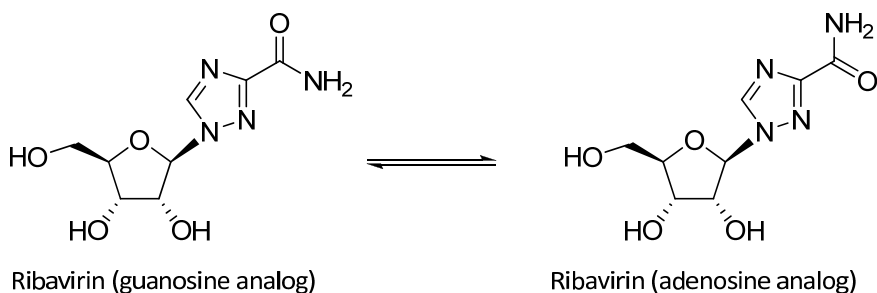
**Figure I-4: Chemotherapeutic nucleoside, nucleotide, and nucleobase analogues.**

These chemotherapeutic agents are toxic to the human body, but because they target enzymes involved in DNA and RNA replication, their toxic effects are higher in the rapidly dividing cells of cancerous tumors. Nucleobases mercaptopurine and thioguanine are similar in their pharmacological mechanisms. Both compete with hypoxanthine and guanine for the enzyme hypoxanthine-guanine phosphoribosyltransferase (HGPRT) which converts the nucleobases to their monophosphate forms. The resulting intracellular nucleotide inhibits several enzymes including inosine monophosphate dehydrogenase (IMPDH), which converts IMP to xanthosine monophosphate (XMP) and inhibits the conversion of IMP to adenosine monophosphate (AMP) via adenylosuccinate. In addition, these metabolites inhibit *de novo*

purine biosynthesis by pseudo-feedback inhibition of glutamine-5-phosphoribosyl-pyrophosphate amidotransferase, the first enzyme unique to the *de novo* pathway for purine ribonucleotides.

Like the antiviral pharmaceuticals, these chemotherapeutic nucleosides are also prodrugs, phosphorylated by various cellular kinases. The diphosphate forms of Gemcitabine and Fludarabine inhibit ribonucleotide reductase. Targets of the other chemotherapeutic nucleosides include adenosine deaminase (Pentostatin), DNA polymerase (Fludarabine), and thymidylate synthase (5-Fluorouracil).

## Ribavirin



**Figure I-5: Ribavirin tautomers - analogues of guanosine and adenosine.**

A unique nucleoside analogue worthy of mention is Ribavirin, first synthesized in 1970 by Joseph T. Witkowski.<sup>3</sup> This prodrug exhibits broad-spectrum antiviral activity, and is currently approved for treatment of hepatitis C viral infections. Remarkably, it has also shown antiviral activity against influenza, HBV, polio, rabies, measles, smallpox, and RSV. Ribavirin can mimic both guanine and adenosine nucleosides depending on the orientation of the amide moiety. (Figure I-5) The major metabolite found intracellularly is the tri-phosphate form, originating from the same kinases that convert other nucleoside analogues. No single mechanism can

account for its antiviral activity, though it has been shown to inhibit IMPDH in its monophosphate form.

IMPDH inhibition prevents *de novo* synthesis of AMP and GMP, but this alone cannot account for the sum of its antiviral properties. It has also been suggested that the incorporation of the triphosphate into growing RNA by polymerases could cause an accumulation of mutations resulting in an "error catastrophe" lethal to viruses. The engineered biosynthesis of this intriguing analogue and its IMPDH inhibitory properties will be the focus of much discussion later in this dissertation.

### Chemical Synthesis

Nucleoside analogue synthesis is achieved through the formation of a glycosidic bond between the sugar and a pre-formed nucleobase, or by step-wise construction of a heterocyclic base on to a pre-formed sugar via substituent at the anomeric position. Addition of preformed bases to sugar moieties is well established, with publications dating back to a 1901. The Koenigs-Knorr reaction utilized silver salts of purines to catalyze the nucleophilic displacement of a halogen substituent from the 1' position of a protected sugar. Other methods were later developed including addition of silylated bases to protected sugars, and  $S_N1$  transglycosylation methods. The greatest disadvantage of these methods, in addition to the multiple protecting group manipulations, is the lack of regio- and stereospecificity. Both silyl base condensation synthesis, and the favored transglycosylation of nucleosides, result in a mixture of products connected at either the 7 or 9 position of the purine base (Figure I-6) and often a mixture of both  $\alpha$  and  $\beta$  anomers. Some control can be achieved through neighboring group participation using acyloxy substituents at the 2' position (Baker's 1, 2-*trans* rule). Where 2' neighboring group participation is not possible, chromatographic separations are necessary.<sup>4</sup>

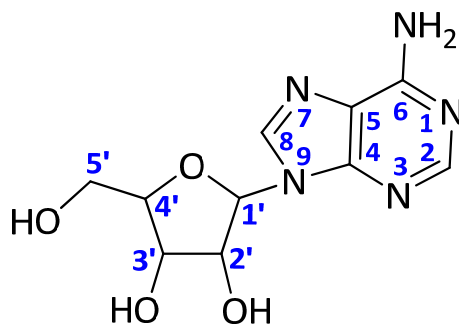


Figure I-6: Numbering system for purine nucleosides

Elaboration of the base moiety is most useful in analogues with highly modified sugars linked to normal or modified nucleobases. Examples of this method are less common and highly specific to each product. As an example, (Figure I-7) an anomeric isocyanate group at 1' of a protected ribose allowed the construction of 5-azacytidine or the fluorescent nucleoside wyosine depending of the synthetic route chosen.<sup>4</sup>

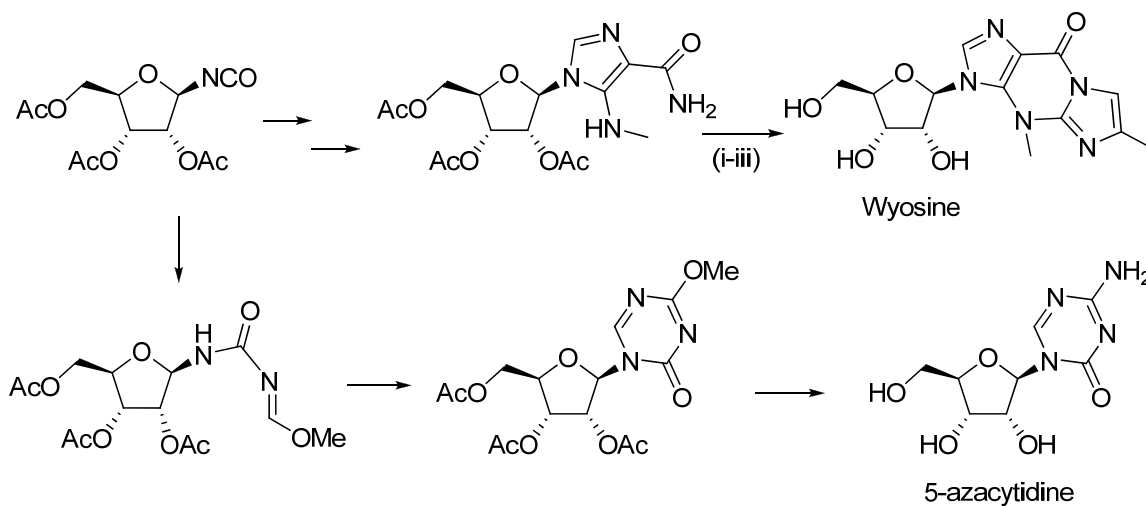
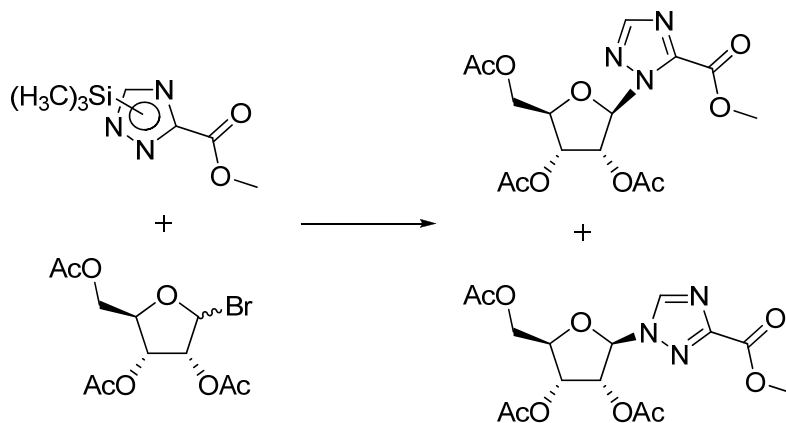


Figure I-7: Nucleobase elaboration for synthesis of wyosine and 5-azacytidine. Reagents: (i) CNBr; (ii) NaOEt, EtOH; (iii) BrCH<sub>2</sub>COCH<sub>3</sub>. Adapted from Blackburn, *et al.*<sup>4</sup>

Both approaches for nucleoside synthesis have been demonstrated in the preparation of ribavirin. The original report featured addition of a silyl triazole base with a 1'-brominated ribose as well as an alternative fusion procedure using triazole base and acyl protected ribose.<sup>3</sup> (Figure I-8) A second article was later published providing ribavirin by construction of the base upon a hydrazine substituted ribose.<sup>5</sup>



**Figure I-8: Original method of ribavirin synthesis as reported by Witkowski.<sup>3</sup>**

Chemically synthesized nucleotide analogues are obtained through phosphorylation of the corresponding nucleoside analogues. This further reduces yield by requiring additional protections, separations, and purifications. In relation to the previous example of ribavirin, phosphorylation of the nucleoside proceeds by addition of trimethyl phosphate and phosphorous oxychloride, giving a final yield of 32%.<sup>6</sup>

### Enzymatic Synthesis

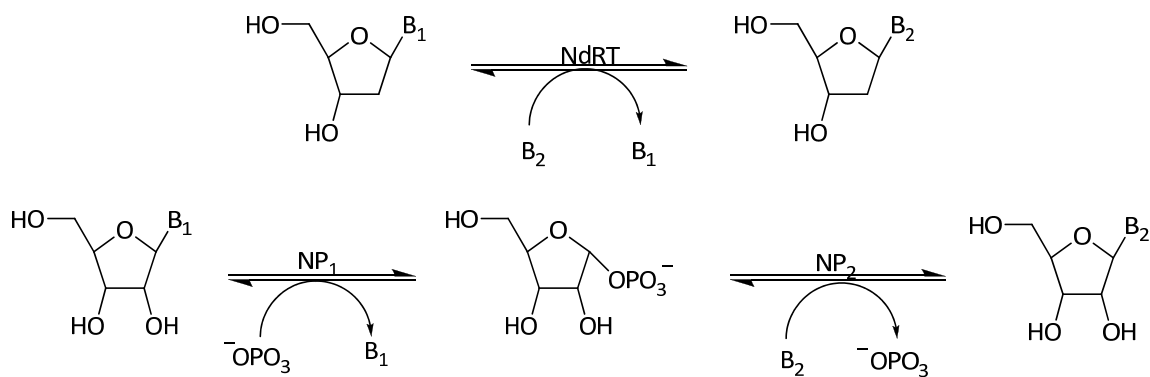
The previously mentioned disadvantages of chemical synthesis of nucleoside and nucleotide analogues include lack of control of regio and stereochemistry, protective group manipulations, low yield, and specialized synthetic methods for each nucleotide product.

Enzymatic syntheses have provided invaluable solutions to these limitations, and present additional environmental benefits due to reduction in generated waste. Chemical synthetic methods consume massive amounts of solvents and hazardous chemicals. The resulting cost can be prohibitive to industrial production, especially when considering the monetary and environmental cost of toxic waste disposal. Enzymatic synthesis occurs in aqueous solutions, often at ambient temperature and pressure, and greatly reduces or eliminates the amount of toxic waste. A review of biocatalytic methods for nucleoside analogues is presented next, followed by a review of biosynthetic methods for nucleotide analogue synthesis.

## **Nucleosides**

Enzymatic synthesis of nucleoside analogues typically uses one of two common methods employing either nucleoside phosphorylase (NP) or nucleoside 2'-deoxyribosyltransferase (NdRT). The net reaction of both methods is essentially the transfer of a free purine or pyrimidine base to the ribosyl portion of an existing nucleoside. (Scheme I-1) Regardless of the enzymatic method used, the ribosyl-accepting base moiety can be naturally or non-naturally occurring and either a purine or pyrimidine. NP acceptance of the ribosyl portion is generally non-specific, and biosynthetic reactions can utilize glycosyl donors of ribose, deoxyribose, or arabinose. NdRTs however, are limited to 2'-deoxyribose glycosyl donors or derivatives thereof. Following is a review of some examples of each method.





**Scheme I-1: General scheme of the one-step nucleoside 2-deoxyribosyltransferase (top) and the two-step nucleoside phosphorylase reaction with ribose-1-phosphate intermediate (bottom). B = purine or pyrimidine nucleobase.**

### ***Biocatalysis by Nucleoside Exchange***

Nucleoside phosphorylases (NP) can be further divided into purine nucleoside phosphorylase (PNP) and the pyrimidine phosphorylases, thymidine phosphorylase (TP) or uridine phosphorylase (UP). Typically a pyrimidine phosphorylase and PNP are used together in a one-pot reaction using a pyrimidine nucleoside as the glycosyl donor and a non-natural purine base as the glycosyl acceptor. Equilibrium often favors the purine nucleoside, providing decent yields with this method. (Table I-1) Alternatively, a single NP enzyme type can be used if the reaction is performed in two-steps, with isolation of the ribose-1-phosphate intermediate product of the first reaction. The net reaction is a pyrimidine to pyrimidine or purine to purine analogue exchange. Without isolation of the ribose-1-phosphate intermediate, reverse exchange of the desired ribonucleoside occurs, resulting in a mixture of nucleoside products.<sup>7</sup> The problematic reversal and unfavorable equilibrium has been addressed in some unique and interesting ways, including removal of substrates for reverse reactions by chemical or enzymatic scavengers, or *in situ* generation of ribose-1-phosphate, as described below.

**Table I-1: Collection of enzymatic syntheses of nucleosides using PNP. Abbreviations: PNP purine nucleoside phosphorylase, UP uridine phosphorylase, TP thymidine phosphorylase, ADA adenosine deaminase, PPM phosphopentomutase.**

glycosyl source	enzymes	product	yield	Reference
uridine	UP, PNP	9-( $\beta$ -D-ribofuranosyl)-allopurinol	4%	<sup>8</sup>
uridine	UP, PNP	9-( $\beta$ -D-ribofuranosyl)-oxoallopurinol	<1%	<sup>8</sup>
thymidine	TP, PNP	N,N-6-dimethylamino-9-(2'-deoxyribofuranosyl)purine	81%	<sup>9</sup>
uridine	UP, PNP	2-amino-6-chloro-9-ribofuranosylpurine	76%	<sup>9</sup>
7-methylguanosine	PNP	3-deazaadenosine	53%	<sup>7</sup>
7-methylguanosine	PNP	Ribavirin (1-( $\beta$ -D-ribofuranosyl)-1,2,4,-triazole-3-carboxamide)	44%	<sup>7</sup>
uracil arabinoside	NP	adenine arabinoside	40%	<sup>10</sup>
thymidine	TP, PNP	M <sub>1</sub> G and M <sub>1</sub> G derivatives	55%-100%	<sup>11</sup>
adenosine, 2'-deoxyadenosine	TP, PNP, ADA	ribosyl and 2'-deoxyribosyl fluorouridine analogues	36%-73%	<sup>12</sup>
2',5'-dideoxythymidine	TP	2',5'-dideoxy-5-fluorouridine	59%	<sup>13</sup>
2',3'-dideoxythymidine	TP, PNP	6-alkoxypurine 2',3'-dideoxynucleosides	11%-69%	<sup>14</sup>
2',5'-dideoxythymidine	TP, PNP	2',5'-dideoxy-6-thiopurine nucleosides	35%-48%	<sup>15</sup>
5'-deoxyadenosine	ADA, PNP	5'-deoxy-6-thiopurine nucleosides	41%-45%	<sup>15</sup>
ribose-5'-phosphate	PPM, PNP	purine and pyrimidine nucleosides	7%-100%	<sup>16</sup>

### *One-pot Reactions with Coupled Phosphorylases*

In one of the earliest examples of PNP catalyzed nucleoside synthesis, ribonucleosides of allopurinol and oxoallopurinol, originally detected in human urine, were recreated on a preparative scale using NPs from guinea pig intestines.<sup>8</sup> The guinea pig protein fractions were partially purified, and a single fraction contained both pyrimidine and purine phosphorylase activities. Although the yield was quite low (<1% for oxoallopurine, 4% allopurine) the 1 L batch processes from six guinea pigs provided 10 mg to 60 mg of material. While this method is impractical by today's standards, it was a seminal step in nucleoside biocatalysis.

Fourteen years later the first example of biosynthesis using purified enzymes from *E. coli* was described. Three separate enzymes were purified from 550 liter cultures: uridine phosphorylase, thymidine phosphorylase, and purine nucleoside phosphorylase.<sup>9</sup> These enzymes in their naturally occurring abundance yielded 2000 to 24,000 units (1U = 1  $\mu$ mol product/min) of pure protein, allowing the synthesis of N,N-6-dimethylamino-9-(2'-deoxyribofuranosyl) purine and 2-amino-6-chloro-9-ribofuranosylpurine from thymidine and uridine, respectively. This is one of the first examples of the classical two-step, one-pot procedure using the combination of pyrimidine and purine phosphorylases.

Burns, *et al* elaborated upon the same enzyme preparation procedure from *E. coli* by immobilizing the purified TP and PNP in cellulose for facilitation of catalyst removal.<sup>14</sup> To demonstrate the synthetic utility and flexibility of this system for multiple nucleoside products, twenty-one 6-alkoxypurine 2',3'-dideoxynucleosides were created from 3'-deoxythymidine, with yields ranging from 11% to 69%.

Yet another excellent example of the base substrate flexibility of PNP is demonstrated with the synthesis of eight nucleosides from thymidine, and utilizing TP and PNP enzymes.<sup>11</sup> Corresponding 2'-deoxynucleoside analogues were made from the bases *O*<sup>6</sup>-ethylguanine, *O*<sup>6</sup>-benzylguanine, *N*<sup>6</sup>-methyladenine, *N*<sup>6</sup>-( $\beta$ -hydroxyethyl)adenine, 6-chloropurine, M<sub>1</sub>G, 7-methyl-M<sub>1</sub>G, and 8-methylguanine in 55% to 100% yield. While this example demonstrates the incredibly relaxed specificity of PNP for the base moiety, it also introduces a previously unpublished difficulty of PNP catalysis; N7 linkage of the purine base to ribose. The desired N9 product thereby required chromatographic separation from the N7 diastereomer (9:1 ratio before separation).

### *Single Enzyme Phosphorylase Reactions*

The preceding examples established that a one-pot combination of PNP and pyrimidine phosphorylases is an effective method for the biosynthesis of a wide range of non-natural nucleosides. Alternatively, a single enzyme type can also be used alone (PNP only or a pyrimidine phosphorylase only), when the equilibrium lies in favor of the desired product. The reaction scheme is identical to that illustrated in Scheme I-1, with the difference of NP1 and NP2 are the same protein. Examples of both TP and PNP in solitary reactions are reviewed here.

In a typical example, 2',5'-dideoxy-5-fluorouridine was created, in reasonable yield, from 5'-deoxythymidine using a single TP biocatalyst.<sup>13</sup> The authors were fortunate to find equilibrium in favor of the fluorinated product, to the extent of 69% conversion for the fluoro analogue, as determined by <sup>19</sup>F NMR. A similar example was later reported using adenosine deaminase to generate 5'-deoxyinosine *in situ* from 5'-deoxyadenosine. With addition of the corresponding base analogues to PNP, 5'-deoxyinosine served as the glycosyl donor for 5'-deoxy-6-thioguanosine and 5'-deoxy-6-mercaptopurine riboside.<sup>15</sup> However, yields above 48% could not be attained, possibly due to reaction reversal.

A rather unique example is described with the discovery of a novel, single enzyme capable of both pyrimidine and purine exchange.<sup>10</sup> This previously unreported enzyme acted as a general NP with both base and sugar substrate non-specificity. The NP accepted 2'-deoxyribose, arabinose and ribose nucleosides. As an example synthesis, adenine arabinoside was made from uracil arabinoside. Not surprisingly, due to the non-specificity, the adenine arabinoside product partially reverted to adenine in the presence of uracil, and yields higher than 45% could not be obtained. With control over the reverse reaction, this nonspecific catalyst could be useful as a synthetic tool.

### *Preventing Reaction Reversal*

As mentioned previously, different strategies to eliminate reaction reversal have emerged. In a direct approach, a two pot method can be employed in which the intermediate ribose-1-phosphate (r-1-p) is purified after the first step (catalyzed by TP or PNP) and then used as starting material for the glycosylation step. Ribavirin (1-( $\beta$ -D-ribofuranosyl)-1,2,4,-triazole-3-carboxamide) was made in this way using inosine, 1,2,4,-triazole-3-carboxamide, and partially purified PNP.<sup>17</sup> Disadvantages of this method include the time and cost of double purifications, and instability of the intermediate.

PNP catalyzed ribavirin synthesis was improved upon by eliminating the intermediate purification step. In this case, ribavirin was biosynthesized in a one-pot process using only PNP.<sup>7</sup> Ribose-1-phosphate did not need to be isolated due to the key glycosyl donor, 7-methylguanosine. In this method, PNP creates r-1-p, and the departing 7-methylguanine base precipitates, thus preventing the reaction reversal.

In another example, inosine and 2'-deoxyinosine served as ribosyl donors, and the hypoxanthine by-product was removed by xanthine oxidase in the production of fluorinated analogues.<sup>12</sup> Usually, equilibrium favors purine products, but in this case equilibrium was forced toward pyrimidine analogue product from a purine ribosyl source. Although inosine and 2'-deoxyinosine are commercially available today, these were generated *in situ* from adenosine and 2'-deoxyadenosine, respectively, via adenosine deaminase. In the first transferase step, hypoxanthine was released from inosine, and the nucleobase was rendered nonreactive by transformation to urate by xanthine oxidase. The ribose 1-phosphate intermediate subsequently reacted with fluorinated base analogues via UP to yield fluorinated nucleosides. Nucleosides 5-fluorouridine, 5-chlorouridine, 5-trifluoromethyluridine, 2'-deoxyuridine, 2'-deoxy-5-fluorouridine, 2'-deoxy-5-chlorouridine, and 2'-deoxy-5-trifluoromethyluridine were produced in

this way. This method of control of reaction direction is very useful for instances with inosine as a glycosyl donor, and yield increased from 29% to 74% when xanthine oxidase was included.

Another way to bypass the reversal problem completely is to use phosphopentomutase (PPM) to go directly to r-1-p from ribose-5-phosphate (r-5-p).<sup>16</sup> In this recent example, chemically synthesized r-5-p was used in combination with PPM to make r-1-p, 2'-deoxy r-1-p, and arabinose-1-phosphate. Commercially available TP and PNP were then used in separate reactions with PPM and r-5-p (or equivalent) to make eleven different analogues from various combinations of sugars and bases. This method eliminates the problem of competing bases with NP reactions.

#### *Nucleoside Biosynthesis by Whole Cells*

All of the methods described thus far utilized purified or partially purified enzymes. A significant amount of work has also focused on the whole cell catalysis of nucleosides. Although the catalytic protein is not over expressed or purified in these examples, it can be deduced from the added substrates and resulting products that NPs are the catalytic agents. Examples of nucleoside analogue syntheses by whole cells are summarized below in Table I-2.

Table I-2: Whole cell catalyzed nucleoside syntheses.

glycosyl donor	glycosyl acceptor	product	organism	reference
inosine	TCA	ribavirin	<i>Erwinia carotovora</i> , <i>Bacillus brevis</i>	18
uridine, orotidine	TCA	ribavirin	<i>E. carotovora</i>	19
guanosine	TCA	ribavirin	<i>Brevibacterium acetylicum</i>	20
uridine, cytidine	TCA	ribavirin	<i>Enterobacter aerogenes</i>	21
uracil arabinoside	adenine	adenine arabinoside	<i>E. aerogenes</i>	22
uracil arabinoside	2-chlorohypoxanthine	9- $\beta$ -D-arabinofuranosyl-2-chlorohypoxanthine	<i>E. aerogenes</i>	23
2'-amino-2'-deoxyuridine	hypoxanthine	2'-amino-2'-deoxyinosine	<i>Erwinia herbicola</i>	24
uracil arabinoside	2,6-substituted purines	2,6-substituted purine arabinosides	<i>E. aerogenes</i>	25
5-trifluoromethyluracil	none	5-trifluoromethyl-2'-deoxyuridine	<i>Escherichia coli</i>	26
uridine, inosine	2,6-diaminopurine, thymine	2,6-diaminopurine riboside, ribosylthymine	<i>E. herbicola</i>	27
ribose-1-phosphate	guanine 7-N-oxide	guanosine 7-N-oxide	<i>Bacillus subtilis</i>	28
2',3'-dideoxyuridine	adenine, hypoxanthine	2',3'-dideoxyadenosine, 2',3'-dideoxyinosine	<i>E. coli</i>	29
2',3'-dideoxyuridine	2,6-substituted purines	2',3'-dideoxy-2,6-substituted purines	<i>E. coli</i>	30
2'-deoxyadenosine	pyrimidine and purine analogues	pyrimidine and purine nucleoside analogues	<i>Bacillus stearothermophilus</i>	31

### ***Nucleoside Biosynthesis by Nucleoside 2'-deoxyribosyltransferase***

Nucleoside phosphorylase methods are the most commonly used biocatalysts for nucleoside analogue biosynthesis, but not without limitation. Occasionally addition of a purine base at N7 has been noted,<sup>11</sup> but the most commonly encountered problem is unwanted reaction reversal. While viable solutions to these problems have been described, the ability to

use a biocatalyst without these drawbacks is of obvious utility. For these reasons, nucleoside 2'-deoxyribosyltransferase (NdRT) often is used for biosynthesis after attempts with PNP have failed. A few examples comparing the two enzymatic methods follow.

NdRT catalyzed the formation of a total of 19 new analogues from the antiviral agent 3'-azido-3'-deoxythymidine (AZT) when combined with several 2-amino-6-substituted purine analogues.<sup>32</sup> (Figure I-9) Additional compounds were prepared by chemical or enzymatic modification of the resultant 6-substituted 2-amino purine analogues for a total of over 40 products in 13% to 96% yield. Some of these new compounds exhibited activity against HIV. The unique substrate specificity of NdRT allowed product where previous attempts using PNP had proven unsuccessful.

Another example found NdRT to be useful in the deoxyribosylation of 5,6,7,9-tetrahydro-7-acetoxy-9-oxoimidazo[1,2-*a*]purine (Figure I-9) after attempts with PNP were ineffective.<sup>33</sup> Deoxycytidine provided the deoxyribose for this and seven other analogues, including both purine and pyrimidine analogues (14% to >95%). PNP gave a mixture of isomers connected at N7 and N9 positions in a ratio of 2:3 in the preparation of 2'-deoxyguanosine. Use of NdRT over PNP was of additional importance due to its increased regioselectivity.

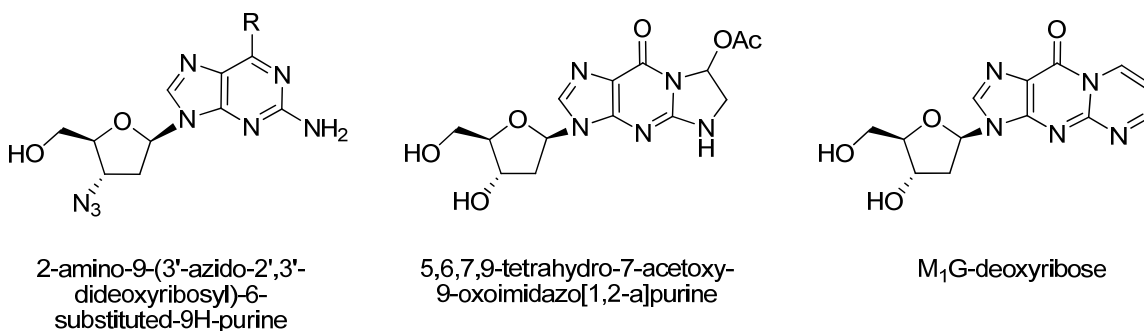


Figure I-9: Selected examples of NdRT catalyzed 2'-deoxyribonucleoside analogues.<sup>33-35</sup>



Occasional N7 coupling by PNP versus the preferred N9 was also noted by the Marnett laboratory in the preparation of M<sub>1</sub>G-deoxyribose<sup>34</sup> (Figure I-9). A combination of PNP and TP were used, with thymidine as the source of deoxyribose, giving only 2% to 5% yield product. NdRT provided only the N9 isomer using 2'-deoxycytidine and the modified base as substrates, in 10% to 15% yield. Additionally, purification was more laborious in the PNP procedure, as both reverse and normal phase separations were required to separate thymidine and product nucleoside, whereas NdRT product purification involved only simple column chromatography.

An example further demonstrating the utility of NdRT is provided in the synthesis of 2-amino-6-chloropurine and 2,6-diaminopurine.<sup>35</sup> While the end goal of this biosynthesis was 2'-deoxyguanosine, the aminopurine analogues were purified at a yield around 40%, a slightly higher yield with NdRT than PNP. Adenosine deaminase (ADA) was then used to convert the amino purines to guanosine. Interestingly, addition of ADA to the NdRT reaction for a one-pot synthesis resulted in 2'-deoxyguanosine reverting to guanine when thymine was present. To avoid this, a one-pot synthesis was performed in sequential steps by adjusting pH, essentially denaturing NdRT with NaOH before neutralization and addition of ADA.

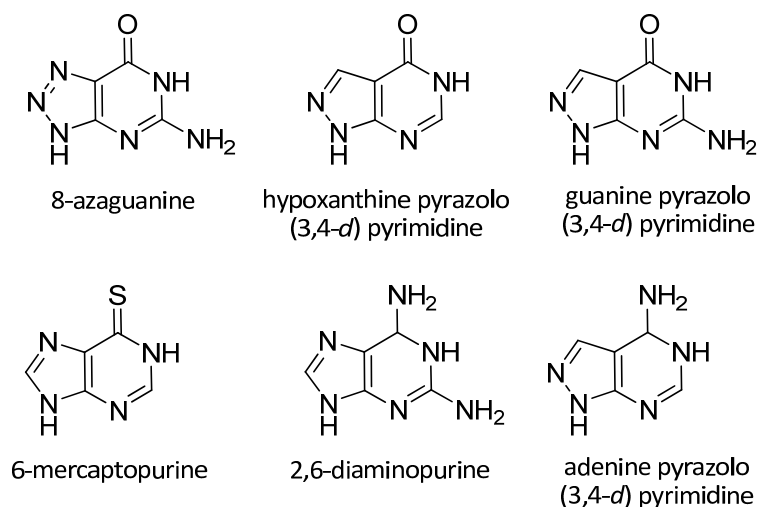
The methods reviewed here highlight the utility of the NP and NdRT enzymes due to the broad specificity for both sugar and base substrates. PNP is more versatile for glycosyl donor acceptance, and accepts a wide variety of nucleobases. NdRT has an equally useful, but slightly shifted specificity for nucleobases. NdRT is stricter in accepting only deoxy sugars, but has the advantage of more regiospecific base addition. Together, this group of enzymes, each with its own window of flexibility, allows multiple opportunities for non-natural nucleoside biocatalysis.

## Nucleotides

Enzyme catalyzed synthesis of phosphorylated nucleotides is much more limited in scope than for nucleosides. The majority of published synthetic methods are for isotopically labeled nucleotides, the biosynthesis of which does not require any promiscuity of the biocatalysts. Only a handful of examples exist pertaining to the biosynthesis of non-natural nucleotide analogues, all of which employ salvage enzymes.

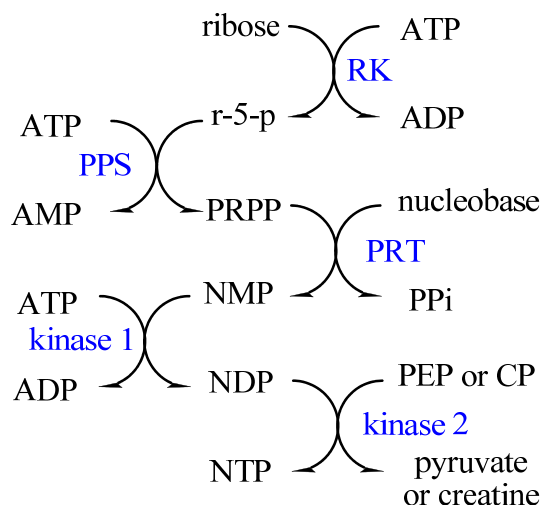
The purine and pyrimidine salvage enzymes used for non-natural nucleotide biosynthesis include hypoxanthine phosphoribosyl transferase (HPRT), xanthine/guanine phosphoribosyl transferase (XGPRT), adenine phosphoribosyl transferase (APRT), or uracil phosphoribosyl transferase (UPRT). Each of these enzymes catalyze nucleotide formation by condensation of the corresponding purine or pyrimidine base to phosphoribosylpyrophosphate (PRPP).

An early example of enzymatic nucleotide synthesis used liver acetone powders of yeast, porcine, and bovine origin.<sup>36</sup> Incubation of 8-azaguanine with PRPP and dialyzed extract of bovine liver acetone powder resulted in 8-azaguanosine monophosphate. An unexpected side product, 8-azaxanthine monophosphate was also created due to a guanase enzyme (guanine deaminase) co-collected in the liver preparation. Subsequent syntheses using porcine liver powder, apparently devoid of the guanase enzyme, then provided pure 8-azaGMP. Each liver preparation of yeast, pig, and beef showed different specificities, and each was tested for turnover with a collection of base analogues. In this manner, five additional nucleotide analogues were prepared. (Figure I-10)



**Figure I-10: Base analogues incorporated into monophosphate nucleotides catalyzed by liver acetone powders<sup>36</sup>**

Not including the work described in this dissertation, Scott, Williamson, and Hennig, from the Scripps Research Institute, have published the only biosyntheses of nonnatural nucleotide analogues using purified enzymes.<sup>37-39</sup> The end goal of the syntheses is for NMR investigation of RNA structures via chemical shifts due to incorporation of the analogues. RNA synthesis by RNA polymerase requires a triphosphorylated nucleotide, which is provided by a two-step phosphorylation of the monophosphate, ultimately originating from ribose. The following describes these methods in greater detail.



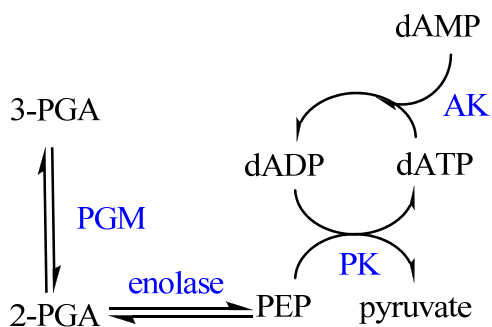
**Scheme I-2: General method of non-natural nucleotide biosynthesis using salvage enzymes, adapted and combined from Scott, Williamson, and Hennig.<sup>37-39</sup> Ribokinase (RK), phosphoribosylpyrophosphate synthetase (PPS), phosphoribosyl transferase (PRT), nucleotide monophosphate (NMP), nucleotide diphosphate (NDP), nucleotide triphosphate (NTP), phosphoenol pyruvate (PEP), creatine phosphate (CP), pyrophosphate (PPi).**

In each of the following mmol scale reactions, a general theme for nucleotide biosynthesis is followed (Scheme I-2). Ribokinase converts ribose to ribose-5'-phosphate (r-5-p) using a phosphate from ATP. An additional molecule of ATP is then needed for phosphoribosyl pyrophosphate synthetase (PPS) to convert r-5-p to phosphoribosyl pyrophosphate (PRPP). Depending on the desired product, the appropriate nucleobase is then added with the corresponding salvage enzyme, replacing the diphosphate of PRPP, allowing the nucleotide monophosphate (NMP). This NMP could then be collected if desired, but for the application of RNA incorporation, phosphorylation to the nucleotide diphosphate and finally triphosphate is necessary. The kinase used for the first step is substrate specific, and depends on the NMP. The final phosphorylation is more relaxed, and either pyruvate kinase or creatine phosphokinase is used.

In this way, 470 mg of 2-fluoroadenosine-5'-triphosphate (2F-ATP) was made from 2-fluoroadenine and ribose in 90% yield.<sup>39</sup> The 40 mL biosynthetic reaction took 4 days at 37°C,

using commercially available enzymes. The salvage enzyme used, APRT, condensed 2-fluoroadenine with PRPP. Essential to purification of this product, dATP was used as the phosphate source instead of ATP. This allowed selective isolation of product by boronate affinity resin, which binds only to *cis* diols. Adenylate kinase transformed 2F-AMP to 2F-ADP, and creatine phosphokinase added the final phosphate, giving 2F-ATP. The entire process required flexibility from all of the enzymes involved to accept non-natural substrates.

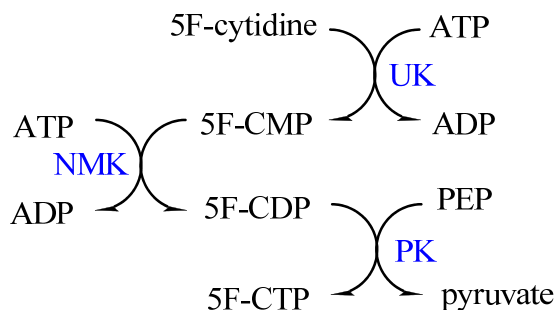
Similarly, biosynthesis of 8-azaguanine-5'-triphosphate from 8-azaguanine used dATP and boronate affinity resin to aid in purification.<sup>37</sup> A bacterial XGPRT was cloned, along with ribokinase and phosphoribosyl pyrophosphate synthetase. All other enzymes were purchased. In this case, product was obtained in 60% yield over a period of 11 days at 37°C, and used guanylate kinase and pyruvate kinase for penultimate and final phosphorylations, respectively. This scheme was also improved from the previous method by addition of a dATP recycling system (Scheme I-3), including *in situ* generation of PEP.<sup>37</sup> This lowers the amount of starting dATP needed to catalytic amounts.



**Scheme I-3: dATP recycling enzymes: adenylate kinase (AK), pyruvate kinase (PK), enolase and phosphoglycerate mutase (PGM). Adapted from Da Costa, C. P., *et al.*<sup>37</sup>**

In the final example, UPRT catalyzed biosynthesis of 5-fluorouridine triphosphate from 5-fluorouracil in 80% yield.<sup>38</sup> Di- and tri-phosphate nucleotides were provided by the

combination of nucleoside monophosphate kinase and pyruvate kinase, respectively. The biosynthesis of 5-fluorocytidine triphosphate was also described in this publication, but a different route was used since a cytosine PRT does not exist. In nature, CTP is made from UTP by CTP synthase, but 5F-UTP is not a substrate. Thus, 5F-CMP was made through phosphorylation of the nucleoside by uridine kinase (Scheme I-4). Nucleoside monophosphate kinase and pyruvate kinase then allowed 5F-CTP. In both methods, products were not purified, but biocatalyst proteins were precipitated over 12 hours, and supernatant used directly in RNA synthesis.



**Scheme I-4: Biosynthesis of 5-fluorocytidine-5'-triphosphate from 5-fluorocytidine nucleoside using uridine kinase (UK), nucleoside monophosphate kinase (NMK), and pyruvate kinase (PK). Adapted from Hennig, M., *et al.*<sup>38</sup>**

## Directed Evolution

The biocatalysis of many diverse non-natural nucleoside and nucleotides have been outlined above. These methods however are limited by the substrate flexibility of the natural enzymes, specific to their natural substrates. Directed evolution can be used to increase the utility of these enzymes by relaxing substrate specificity or increasing stability in normally inhibitory environments.

Directed evolution, also known as molecular evolution or *in vitro* evolution, is an approach to alter enzyme properties through random genetic mutations, followed by a screen

or selection of translated proteins exhibiting the desired phenotype. Genes of selected enzymes are then subjected to subsequent rounds of mutation and selection or screening. Directed evolution has been used to enhance enzyme solubility, increase stability in extremes of temperature and pH, increase activity in heterologous hosts, increase stability and activity in organic solvents, add enantioselectivity, and to alter substrate specificity. An introduction into methods of creating genetic diversity and methods for selections and screening follow, with an emphasis on directed evolution for biosynthesis.

## **Mutation Methods**

### ***Random Mutagenesis***

The first step in a directed evolution experiment is the introduction of genetic diversity. Mutational methods include error prone polymerase chain reaction (epPCR), chemical mutagenesis, UV irradiation, and mutating strains. The most common and successful is epPCR, in which an error prone, thermostable polymerase lacking proofreading ability (*Taq*), is used to amplify a gene of interest, incorporating random mutations along the entire length of the gene. The error rate of the polymerase alone is not high enough for directed evolution purposes, so a higher mutation rate is achieved through addition of  $Mn^{2+}$ , and by manipulation of the number of thermocycles and template and dNTP concentrations. Ideally, a mutation rate resulting in 1 to 2 amino acid changes per protein is desired, though in some cases, higher mutation rates have been beneficial. Although the position of the mutation in the gene is random, the change in nucleotide and ultimately the resulting change in amino acid are not. A single point mutation in a gene can only provide 5 to 6 other amino acids due to the nature of the genetic code. Expanding the amino acid possibilities through the accumulation of two mutations in a single

codon is unlikely. Additionally, the polymerase tends to incur transitions (purine to purine or pyrimidine to pyrimidine) over transversions (purine to pyrimidine or pyrimidine to purine). Thus, only a fraction of all the possible amino acid substitutions is accessible by point-mutation methods.

Mutazyme (Stratagene) is a polymerase engineered to have a higher than normal mutation rate, with less bias than the typical *Taq* polymerase. It has also been combined with traditional epPCR in a combined library. Separate libraries – one created with *Taq* and the other with Mutazyme - were combined by DNA shuffling. The purpose was to provide an end library with less bias.<sup>40</sup> Although the mutational bias was carefully analyzed in this report, no link of bias to increased or decreased enzyme performance was made, since subsequent screening or selection was not performed.

After amplification of a gene through epPCR, the pcr product must then be digested and ligated into a suitable vector, causing a decrease in yield. To avoid the ligation step necessary with epPCR, mutating strains or whole cell mutation by UV irradiation can be used. These techniques, however, can have unpredictable and undesirable results due to host genome mutation. Alternatively, whole plasmid epPCR is possible as a mutation method (Bringer strategy) thanks to various commercially available kits for whole plasmid PCR.<sup>41</sup> Although the host genome is not affected, similar drawbacks exist with off target mutations of ribosomal binding sites or antibiotic resistance markers.

Cassette mutagenic methods, using synthetic oligonucleotides, bypass the bias of epPCR and provide diversity in specific locals, unlike epPCR, whole cell, or whole plasmid methods. Various methods of cassette mutagenesis methods are used, often for the purpose of active site ligand binding analysis. Reetz introduced combinatorial active-site saturation test, or CASTing for widened substrate specificity of a lipase. This method is essentially saturation mutagenesis



of adjacent residues, focused upon the active site.<sup>42</sup> This method reduces library size required by focusing diversity, but unexpected beneficial mutations distal from the active site can be overlooked and stop codons are occasionally introduced. Similarly, gene site saturation mutagenesis (GSSM) uses degenerate primers to produce all 19 amino acid substitutions at every position of the gene.<sup>43</sup> This was used to produce every possible single amino acid variant of a 330-amino acid nitrilase, providing improved enantioselectivity.

### ***Recombination Methods***

In epPCR libraries, higher mutation rates provide sequences more diverse from the parent gene, providing greater opportunity for improvement, but also increasing the percent inactive enzymes, since most mutations are deleterious if not neutral. With high mutation rates, both good and bad mutations are combined. However, *in vitro* formation of recombinant genes overcomes this obstacle by recombining homologous sequences, in parallel with sexual recombination in nature. This allows collection of combined beneficial mutations, and elimination of those, which are deleterious. The original DNA shuffling method by Stemmer recombines homologous DNA sequences; either naturally occurring homologues of a gene family, or select mutants of a single gene, generated by epPCR.<sup>44</sup> (Figure I-11) The pool of genes is digested by DNase I, and fragments are separated by agarose gel electrophoresis. The fragments are then reassembled into full length genes by a primer free PCR. During this stochastic process, the fragments prime each other on the homologous regions, resulting in recombination when a fragment derived from one parental gene primes on a fragment from a second parental gene. Finally, the full length genes are amplified by standard PCR with flanking primers, resulting in a library of hybrid genes with multiple cross-overs. This method could also be used for combining two or more improved properties, evolved separately in different

experiments. An example of this concept is provided with the activity of subtilisin for a non-natural substrate sequence, and simultaneous stability toward hydrogen peroxide.<sup>45</sup>

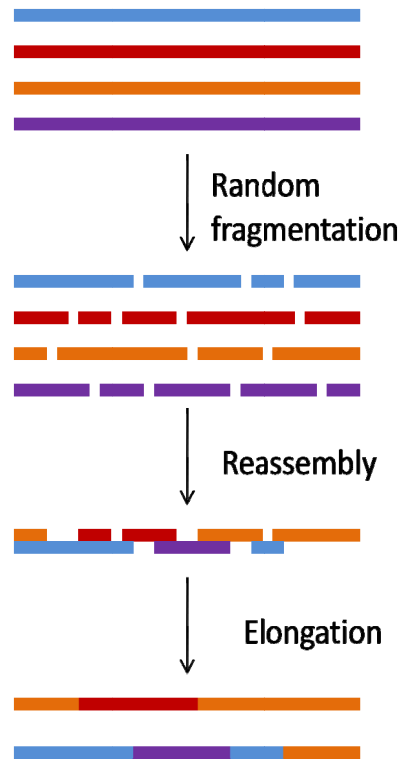


Figure I-11: DNA Shuffling; adapted from Valetti, *et al.*<sup>46</sup>

Many other methods for recombination have been reported over the last 15 years, the majority of them modifications of Stemmer's seminal process. Random chimeragenesis on transient templates (RACHITT) is similar to DNA shuffling, but is intended to produce chimeras with a higher number of crossovers.<sup>47</sup> As in DNA shuffling, a pool of homologous genes is digested into fragments. However, the RACHITT technique requires a single strand of one of the parental homologues to be left undigested as a "transient template" to which complementary fragments of the other parents anneal. After reassembly on the full-length strand of the remaining parent, any unhybridized, or mismatched sections of the fragments are trimmed, extended along the template, and then ligated to generate full-length genes. Finally, the

template strand is destroyed leaving only the ligated gene fragments, which are then converted to double-stranded DNA. The advantage of this assembly process is that it creates a greater number of crossovers, at approximately 14 per gene, and amplification of parent sequences is avoided. The disadvantages are that sequences similar to the scaffold will be incorporated more frequently into the chimeras, and the additional steps required; the generation of single strands, the removal of mismatched DNA and removal of the template strand.

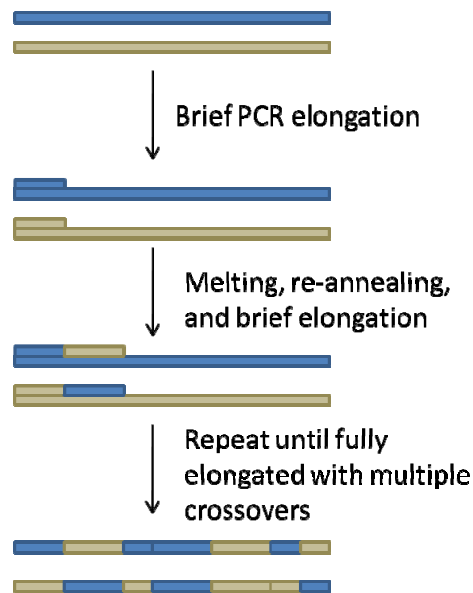


Figure I-12: Staggered extension protocol (StEP); adapted from Zhao, *et al.*<sup>48</sup>

Staggered extension protocol (StEP) is also similar to DNA shuffling, but generates chimeras without digesting the parental DNA.<sup>48</sup> In StEP, the DNA homologous templates are mixed together with one or more primers, and subjected to very short annealing and extension steps. (Figure I-12) The short cycles result in only partial elongation of products, which are then melted from the template, and randomly annealed to another parent strand in the next cycle, resulting in a crossover. This process is repeated until full-length genes are obtained. Optimization of timing the annealing and extension cycles can be challenging, and determining

whether recombination has occurred is not always straightforward. Full length products may simply be amplified template rather than chimeras. Prior sequence analysis and comparison of parent restriction sites to product can facilitate determination of the degree of recombination. Despite the difficulties, once a specific StEP method has been validated, the crossover procedure can be easier than with DNA shuffling.<sup>49</sup>

The above shuffling methods are useful for recombining only highly homologous sequences. Methods for recombination of nonhomologous sequences include incremental truncation of hybrid enzymes (ITCHY) and sequence homology-independent recombination (SHIPREC).<sup>50, 51</sup> In the ITCHY method, parental DNA is mixed and progressively truncated via carefully timed exonuclease III digestion on the 3' end of one strand, and S1 nuclease digestion on the 5' end. (Figure I-13) Blunt end ligation results in single crossover chimeras that may or may not be analogous to their position in the template gene. Thus, out-of-frame and nonsense products will arise along with new, potentially useful hybrid sequences.

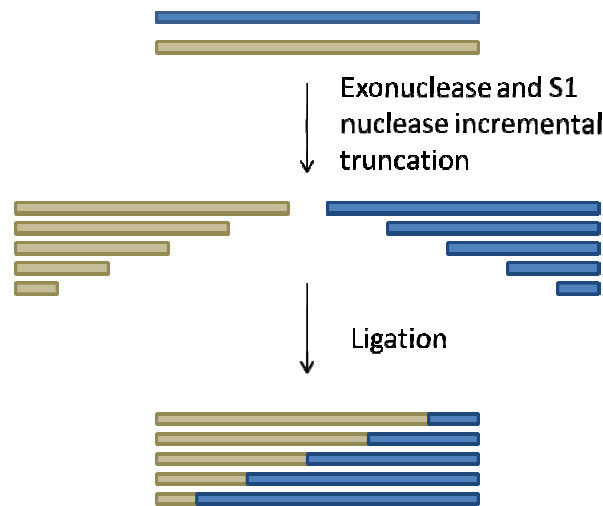


Figure I-13: Incremental Truncation (ITCHY); adapted from Valetti, *et al.*<sup>46</sup>

SHIPREC combines two genes of interest connected through a linker sequence containing multiple restriction sites. (Figure I-14) This heterodimer is then digested by DNase I in the presence of  $Mn^{2+}$  to provide double-stranded fragments of random lengths. Fragments are treated with S1 nuclease to produce blunt ends, and gel filtered for correct size. Blunt end ligation produces circularized DNA, which is then digested and linearized, removing the linker sequence. This creates a library of single-crossover hybrids where the gene that was at the 5' position in the dimer will now be at the 3' position, donating the C terminus of the new protein.

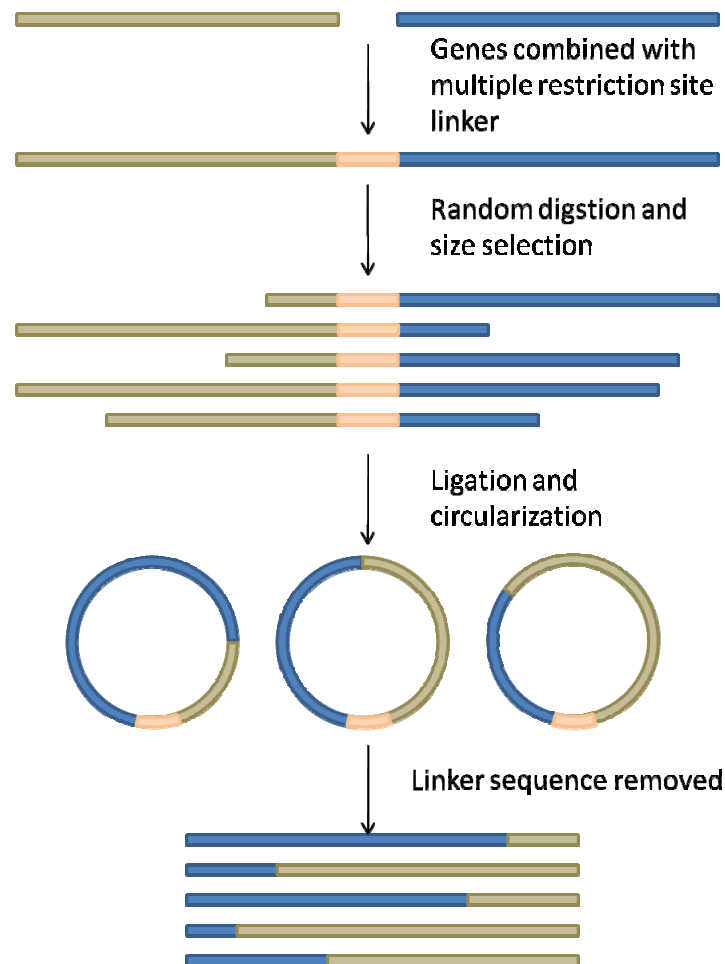


Figure I-14: Sequence homology-independent recombination (SHIPREC); adapted from Zhao, *et al.*<sup>52</sup>

Greater control over the rate and type of crossovers occurring in DNA shuffling is achieved in the three following similar methods; degenerate homoduplex recombination (DHR), assembly of designed oligonucleotides (ADO), and synthetic shuffling.<sup>53, 54</sup> In each case, an assembly of synthetic, overlapping, and partially degenerate oligonucleotides anneal to homologous oligos. The synthetic oligos essentially take the place of the digested oligos of traditional DNA shuffling, but comprise all possible recombinations of alleles. The result is a library of cross-over chimeras in which all shuffled variants are equally likely, regardless of homology to parental DNA, with minimized hybridization of parental genes.

## **Selection**

To find improved enzymes, genotype must be linked to phenotype. Accordingly, desired activity can be linked to host cell survival (selection) or cells can be lysed and expressed proteins analyzed in microtiter format (screening). Selection methods have an advantage of being able to filter through larger libraries;  $10^9$ - $10^{10}$  for selections, limited only by ligation and transformation efficiency, and around  $10^6$  for screens, limited by cost and time. Not all enzymes can be linked to cell survival, and bacterial hosts often find ways to compensate and create metabolic solutions to survive in unfavorable environments. A unique screen is required for every directed evolution experiment, as screening methods are often highly specific to the function of the enzyme being assayed. Much information covering screening and selection methods in directed evolution have been published including several review articles<sup>55-58</sup> and a book describing various methods and protocols.<sup>59</sup>

## **Positive Selection**

Selection methods are typically positive selections for growth (variants exhibiting desired activity have an advantage for survival) and relatively few negative selections can be found. Some methods reviewed here include selection for thermostability, complementation of engineered deficiencies in the host (auxotrophs), and resistance to environmental toxins. Selection for thermostable enzymes involves varying degrees of complexity, most including the cloning of a library of genes of a non-thermostable enzyme into a thermostable host. In the first example, a gene encoding resistance to the antibiotic kanamycin (kanamycin nucleotidyl-transferase) was cloned from *Staphylococcus aureus* and added to the thermostable *Thermus thermophilus*.<sup>60</sup> Diversity was introduced into the gene by DNA shuffling and survivors of a  $3.2 \times 10^5$  size library, grown in the presence of kanamycin, were selected. Thermostability of the enzyme increased by 20 degrees over wild type. Thermostability of a ribonuclease HI was increased using an *E. coli* strain MIC3001 as a host, which shows an RNase H dependent, temperature-sensitive growth phenotype.<sup>61</sup> Only those colonies carrying a thermostable RNase H were capable of DNA replication and survival. Finally, thermostability of 3-isopropylmalate dehydrogenase was increased after growth at elevated temperatures in a stable host. The *leuB* gene of the thermophile was replaced with a temperature-sensitive chimeric *leuB* gene, resulting in a final thermostable mutant.<sup>62-64</sup>

The latter two examples in thermostability utilized selection of auxotrophic hosts with engineered deficiencies complemented by expression of the mutated gene. The following examples also use auxotrophy for enhancing traits other than thermostability. An aspartate aminotransferase (necessary in conversion of amino acids aspartate and glutamate) was evolved into a  $\beta$ -branched aminotransferase via an engineered deficiency in the endogenous  $\beta$ -branched aminotransferase.<sup>65</sup> The gene was mutated by random mutagenesis, and the best mutant

accepted valine over the natural substrate, aspartate, in a  $2.1 \times 10^6$  improvement. Later experiments isolated a mutant aspartate aminotransferase with increased binding of substrate 2-oxovaline.<sup>65, 66</sup> Another group evolving the same enzyme, and using similar techniques, changed the aspartate aminotransferase into a tyrosine aminotransferase by auxotrophic growth in absence of tyrosine and phenylalanine.<sup>67</sup> The final example of these selection methods relied on the ability of *E. coli* to secrete active trypsin into the periplasm and on the inability of *E. coli* to metabolize *p*-naphthylamide derivatives of amino acids.<sup>68</sup> *E. coli* X90 cells auxotrophic for arginine expressed trypsin variants, which cleaved an arginyl naphthylamide bond. Thus, when X90 cells are grown on a medium containing arginine  $\beta$ -naphthylamide as the only source of arginine, they survived only with expression of active, Arg-specific trypsin.

The most common method for selection is insertion of a gene library, which conveys resistance to an environmental toxin by degradation. A review of catabolic enzymes and pathways has been published, including several examples of whole genome shuffling by chemical mutagenesis and growth selection in the presence of toxic substrates.<sup>69</sup> One of the early examples of growth selection of mutants can be found in the original DNA shuffling method by W.P.C. Stemmer.<sup>44</sup> A mix of  $\beta$ -lactamase genes were shuffled and selected for antibiotic degradation. Improved mutants showed increased activity toward the antibiotic cefotaxime by 32,000 fold over wild type. Other antibiotic examples include resistance to ampicillin and kanamycin.<sup>60, 70</sup> Hydroxybiphenyl monooxygenase and biphenyl dioxygenase with increasing ability to degrade polychlorinated biphenyls were obtained by shuffling critical segments of *bphA* genes from *Burkholderia* sp. strain LB400, *Comamonas testosteroni* B-356, and *Rhodococcus globerulus* P6.<sup>71</sup> These variants exhibited high activity toward dichlorobiphenyls, unlike the parental enzymes. Human *O*-6-methylguanine-DNA methyltransferase mutants were selected for the ability to provide alkyltransferase-deficient



*Escherichia coli* with resistance to the methylating agent N-methyl-N'-nitro-N-nitrosoguanidine<sup>72</sup>  
Resistance to 5-fluorouridine in human thymidylate synthases was selected by Encell and Landis.<sup>73, 74</sup>

Glutathione S-transferases in mammalian cells catalyze the detoxification of electrophilic environmental carcinogens and alkylating drugs.<sup>75</sup> Random mutagenesis of three regions comprising the glutathione S-transferase electrophile-binding site, followed by selection of *Escherichia coli* expressing the enzyme library with mechlorethamine, yielded mutant enzymes that showed significant improvement in catalytic efficiency for mechlorethamine conjugation.

Not just single enzymes, but entire pathways have been evolved and selected for using growth selection. Cramer, *et al* evolved a three enzyme pathway encoding arsenate resistance in *E. coli*.<sup>76</sup> All genes, derived from *S. aureus*, were cloned into one plasmid and selected for increased resistance to arsenate. This is the first example of using DNA shuffling for a pathway.

### **Negative Selection**

All of the examples given above relied on positive selections for growth and survival. Less common is the negative growth selection in a library, in which the desired catalytic process results in cell death or growth inhibition. To implement this method, cells must be sorted and replicated into stock and test microtiter plates to facilitate backtracking and resurrection of the desired stock cultures. Examples of negative growth selections follow.

Herpes simplex virus type 1 thymidine kinase (TK) was transformed into TK deficient *E. coli* and selected for ability to phosphorylate nucleoside analogues.<sup>77</sup> Colonies with TK activity were grown in microtiter plates and replicated into plates containing nucleoside analogues

ganciclovir and /or acyclovir. Clones able to phosphorylate these analogues were unable to grow and thus selected as hits.

The same concept was used in the negative selection of HSV TK mutants with increased phosphorylation of zidovudine (AZT).<sup>78</sup> Genes of HSV types 1 and 2 TK were combined by DNA shuffling and cloned into *E. coli*. Phosphorylation of AZT resulted in lethality at 32 fold lower amounts than HSV-1 parent, and 16,000 times lower than HSV-2 parent. More discussion on this and the previously mentioned TK experiment can be found in the later section in this chapter, titled "Nucleotide biosynthesis by directed evolution".

### ***Positive and Negative Selections***

Incorporation of unnatural amino acids was achieved using both positive and negative selections for evolution of a tyrosyl tRNA synthetase.<sup>79, 80</sup> Positive selection was based on an amber stop codon placed in the middle of an antibiotic resistance gene. The engineered tRNA synthetase, which recognizes the stop codon and incorporates unnatural amino acids, allowed complete peptide product, and antibiotic resistance. Negative selection against incorporation of the natural tyrosine amino acid was achieved by placement of the amber stop codon in a toxic gene for barnase. Incorporation of the amino acid allows translation of the complete length lethal protein.

### **Screening**

Many reviews covering screening methods in directed evolution have been published, as mentioned in the selections section.<sup>55-58, 81</sup> Each screen for every directed evolution experiment is essentially unique. Screens require the isolation of individual colonies, typically followed by cell lysis and activity assay. There are however a handful of applications in which the product

itself presents as a spectroscopic identifier, either as a change in the visible wavelength, or as a change in fluorescence.

### **Visual Response Assays**

Variants with colorful products can be identified by visual inspection in solid phase agar plates, or in the wells of a microtiter plate. This has been particularly useful in the evolution of carotenoid production. Wang et al increased production of astaxanthin by heterologous expression of geranylgeranyl diphosphate synthase in *E. coli*.<sup>82</sup> Colonies with higher color intensity were selected by visual inspection. Using a similar screen, lycopene production by co-expression of isoprenoid pathway genes was increased by encoding phytoene desaturase from *Rhodobacter sphaeroides*, in *E. coli*.<sup>83</sup>

Visual screening was also used in directed evolution of novel isoprenoids by several groups. A C-30 carotenoid synthase was evolved to function as a C-40 synthase, and used to synthesize previously unknown C-45 and C-50 carotenoids.<sup>84</sup> Genes for three different carotenoid desaturases were shuffled with a carotenoid hydratase, a cyclase, and a hydroxylase and visually screened in *E. coli*, resulting in four novel hydroxycarotenoids.<sup>85</sup> Many new C35 and C40 carotenoids were evolved by evolution of enzymes from different organisms by Lee.<sup>86</sup>

The most cited use of color screening of carotenoids is that of Claudia Schmidt-Dannert.<sup>87,88</sup> Genes encoding geranylgeranyl diphosphate synthase and phytoene synthase were transformed into *E. coli*. A library of DNA shuffled desaturase mutants of two homologous phytoene desaturase genes were co-transformed into the same *E. coli*. Several novel carotenoid compounds were isolated. DNA shuffling of two homologous lycopene cyclase genes further extended the evolved tetrahydrolycopene pathway and created many different colored variants.

Solid phase screens with a visually detectable response include the screen used for the directed evolution of  $\beta$ -galactosidase, giving a color response on agar when combined with X-gal.<sup>89</sup> A *Bacillus stearothermophilus* lactate dehydrogenase was screened in a similar manner,<sup>90</sup> as well as a  $\beta$ -glucuronidase evolution using X-gluc.<sup>91</sup>

### ***Fluorescent Response Assays***

Just as cells, expressing visually detectable compounds, can be screened on solid phase agar plates or microtiter plates, evolution experiments in which a fluorescent response is elicited can be screened in the same format with fluorescent spectrometry. Fluorescent response can be a direct result of the protein itself, as with green or red fluorescent proteins, through a fluorogenic product of the enzyme, or by linking the response of the enzyme with a secondary, fluorescent reaction.

Several examples using directed evolution for variations of fluorescent proteins can be found. These include evolution of green<sup>92-94</sup>, red<sup>95-97</sup>, red/green<sup>98, 99</sup> orange, and yellow<sup>98, 100</sup> derivatives of green fluorescent protein. These have been evolutionary targets since early in the history of directed evolution, and later used as reporters for evolution of other proteins for either activity<sup>101-103</sup> or solubility.<sup>104-106</sup>

To screen evolved P450 enzymes for activity in aromatic hydroxylation of naphthalene, a secondary reaction using a coexpressed horseradish peroxidase (HRP) resulted in intracellular fluorescent activity.<sup>107</sup> The H<sub>2</sub>O<sub>2</sub> byproduct of P450 reacted with hydroxylated product, converting it into a fluorescent polymer, which remained inside the cell. The random mutant library was screened by fluorescence digital imaging on an agar plates. Libraries of other hydroxylation catalysts have been screened using this method.<sup>108</sup> Enzyme activities of epoxide hydrolases, acylases, phosphatases, or lipases are easily detected by fluorescence.<sup>109</sup> A

periodate coupled fluorogenic assay utilized fluorescent product to detect activity via microtiter format.

Fluorescent response also has the advantage of allowing a higher throughput screening method — fluorescence-activated cell sorting (FACS). FACS is a type of flow cytometry in which a mixture of cells is sorted, one cell at a time, based upon the specific light scattering and fluorescent characteristics of each cell. The introduction of this method provided a great advancement for screening directed evolution libraries.

An evolution experiment of 2'-deoxynucleoside kinase from *Drosophila melanogaster* gave a fluorescent product trapped inside a whole cell. The nucleoside analogue kinase library was created with GeneMorph II kit (Stratagene, La Jolla, CA) with 2-4 mutations per gene, and transformed into *E. coli*.<sup>110</sup> Cell cultures were incubated with fluorogenic nucleoside analogues, which became phosphorylated by mutant kinases. Washing of cells removed unreacted nucleoside, while charged nucleotide product remained trapped in the cell. FACS screening of  $5 \times 10^5$  cells per round isolated a mutant with a 30-fold preference for analogue over natural substrate. Other examples of directed evolution in which the fluorescent product is retained in the cell have been used to screen mutants of glutathione transferase,  $\beta$ -galactosidase, and glycosyltransferase.<sup>111-114</sup>

FACS methods utilizing *in vitro* compartmentalization (IVC) allow the entrapment of fluorescent products while simultaneously maintaining a linkage between the product, enzyme, and the enzyme encoding gene. IVC uses a water-in-oil emulsion droplet, with transcription, translation and encoded enzyme reactions all occurring in the inner, aqueous portion, essentially acting as an artificial cell. In the original publication, the emulsion was used to evolve a DNA methyltransferase, which methylated its own DNA, thus protecting it from subsequent restriction enzyme digestion.<sup>115</sup> A later and more complex example of IVC for directed evolution

of a phosphotriesterase was reported.<sup>116</sup> Genotype and phenotype were linked by attaching DNA to a streptavidin coated bead, and substrate bound to biotin. Beads were then selected using antibodies with affinity tags and fluorescent labels, finally sorted by FACS. Literature provides more discussion and reviews of FACS of IVC.<sup>117-122</sup>

Most screening methods, however do not lend themselves to direct whole cell imaging by change in visible color or by fluorescent response, thereby requiring lysing and assaying enzymatic activity of individually isolated colonies in microtiter plates. Screening methods of this type are reviewed in the general screening reviews mentioned above. Briefly, examples include evolution of beta-glucosidase,<sup>123</sup> subtilisin S41,<sup>124</sup> hydantoinase,<sup>125</sup> aldolase,<sup>126</sup> and fungal peroxidase<sup>127</sup> to name only a few.

### ***Screening by Display Techniques***

An increasing amount of screening methods using display technologies have emerged over the years including mRNA display, ribosomal display, cell surface display, and the most common, phage display.

Both ribosome display and mRNA display use *in vitro* transcription to convert the DNA library to mRNA, circumventing cellular transformation. This allows libraries of up to  $10^{14}$  to be constructed, bypassing any issues with cell toxicity of substrates or products. In mRNA display, an mRNA molecule is tagged with a chain terminating puromycin nucleotide analogue, separated from the gene with spacer sequences. During *in vitro* translation, the ribosome stalls at the puromycin, and creates a covalent linkage between mRNA and peptide.<sup>128-130</sup> Ribosomal display is similar in the induced ribosomal stalling, but unique in that it is initiated by omission of the stop codon.<sup>130-133</sup> The ribosome remains attached to the newly synthesized protein and mRNA. The resulting complexes are then screened for ability to bind to desired target

molecules. As such, these methods are most used for evolving binding interactions rather than catalytic activities.<sup>134</sup>

Cell surface display libraries present proteins at the surface of cells via fusion to a membrane associated protein.<sup>135-137</sup> Cells of *E. coli*, yeast, and mammalian species have been used. These techniques are amenable to FACS screening or as with other display techniques, increased binding affinity for substrate.<sup>138-140</sup> Several reviews are available on this method.<sup>141-143</sup>

Phage display is a very popular screening technique, used primarily to select for binding interactions.<sup>144-151</sup> The method has also been used to screen for binding to transition state analogues,<sup>151</sup> and in screening for stability.<sup>152</sup> In phage display screening, the DNA of a gene library is fused to a viral coat protein gene, and combined in a phagemid (a vector containing both plasmid and phage origins of replication). The phagemid DNA then becomes packaged in the phage capsid that displays the protein. The phage are then subjected to chromatographic separation or panning to select for desired binding activity. The phage is then recovered and infects *E. coli*. The phagemid is replicated as a plasmid and this DNA can be recovered for analysis and later rounds of directed evolution.<sup>153</sup>

There is no universal screening method for searching the sequence landscape for improved activity of all enzymatic reactions. Some lower throughput methods are however closer to universal than the previously mentioned highly specific screens. These methods include an analytical HPLC screen for alkaline phosphatase,<sup>154</sup> several mass spectrometry screens involving proteases<sup>155, 156</sup> and enantioselective enzymes,<sup>157</sup> one of which claims to be able to screen  $10^4$  samples per day.<sup>158</sup> Enantioselectivity has also been assayed by capillary electrophoresis (up to 7000 samples per day),<sup>159</sup> circular dichroism,<sup>160</sup> pH (enantioselectivity of hydrolytic reactions),<sup>161</sup> and infrared thermal imaging.<sup>162</sup> Infrared spectroscopy has also been used to screen an acylation reaction.<sup>163</sup>

## Applications

Applications for directed evolution and the scope of enzyme targets are nearly endless. Most targets fall into one of the four major categories; hydrolases, oxygenases, transferases, and lyases. The majority of reports of successful directed evolution experiments involve catabolizing enzymes participating in hydrolysis. This is partly due to the many possibilities of screening methods for hydrolases, and the high promiscuity of these enzymes. Enzymes in this category that have been modified by directed evolution include lipases and esterases,<sup>164-167</sup> proteases,<sup>48, 124, 168, 169</sup>  $\beta$ -lactamases,<sup>44, 170, 41</sup> acylases,<sup>171-173</sup> glycosidases,<sup>174, 175</sup> and organo-phosphate hydrolases.<sup>176-178</sup>

Oxygenases are equally popular, and many reviews have been written on these enzymes including monooxygenases (including several P450s and non-heme monooxygenases), dioxygenases, and peroxidases.<sup>179, 46, 180-186</sup>

The transferase category encompasses a wider variety of enzyme catalyzed reactions. These include glutathione transferases,<sup>75, 113, 187, 188</sup> and methyltransferases,<sup>189-192</sup> and the increasingly popular evolution of glycosylating enzymes for glycorandomization.<sup>193-195</sup>

Other than the transferases, directed evolution of bond building enzymes is restricted to lyases. These include the carbon-carbon bond forming reactions of aldolases.<sup>196-198</sup> Decarboxylases are also in the lyase category, and include the recently evolved enzyme rubisco.<sup>199</sup>

## Nucleotide Biosynthesis by Directed Evolution

Only a handful of examples of nucleoside or nucleotide analogue biosynthesis using directed evolution have been reported. In the earliest example by Black, *et al*, evolved herpes simplex virus type 1 (HSV-1) thymidine kinase (TK).<sup>77</sup> Wild type HSV-1 TK phosphorylates not



only the natural substrate thymidine, but also nucleoside analogues. It is this relaxed specificity of HSV-1 TK, and the rigidity of human TK, that allows the use of nucleoside analogues in the treatment of herpetic infections. The monophosphate nucleotide analogue product of TK is anabolized further by cellular kinases to nucleoside triphosphates, which ultimately are incorporated into DNA and result in chain elongation termination. This difference in species specificity also provides the basis for potential cancer gene therapy. An HSV gene with a much higher rate of phosphorylation than human host TK would allow less drug administration, and lower cell toxicity upon introduced into a tumor. To this end, HSV-1 TK was mutated by epPCR and selected for improved catalytic activity with both negative and positive complementation methods. The mutant library of genes was transformed into TK deficient *E. coli* and selected for TK activity by addition of 5'-fluorodeoxyuridine, uridine, and thymidine. Active TK enzymes phosphorylate 5'-fluorodeoxyuridine, the product of which inhibits thymidylate synthase of the *de novo* pathway. Uridine is added to inhibit thymidine phosphorylase. Thus, the requirement for dTMP can then be fulfilled only by an active thymidine kinase (positive selection). Colonies with TK activity were grown in microtiter plates and replicated into plates containing nucleoside analogues ganciclovir and/or acyclovir. Clones able to phosphorylate these analogues were unable to grow and thus selected as hits (negative selection).

Stemmer's lab added to this concept by selecting HSV TK mutants with increased phosphorylation of zidovudine (AZT).<sup>78</sup> Genes of HSV types 1 and 2 TK were combined by DNA shuffling and cloned into *E. coli*. The positive selection step for TK activity with thymidine was omitted, and only the negative selection used, as described above. Phosphorylation of AZT resulted in lethality at 32 fold lower amounts than HSV-1 parent, and 16,000 times lower than HSV-2 parent.

A much more recent example by Lutz was another evolution of a kinase for nucleotide analogues.<sup>110</sup> A mutant library of 2'-deoxyribonucleoside kinase from *Drosophila melanogaster* was created with GeneMorph II kit (Stratagene, La Jolla, CA) with 2-4 mutations per gene, and transformed into *E. coli*. Cell cultures were then incubated with fluorogenic nucleoside analogues, which became phosphorylated by mutant kinases. Washing of cells removed unreacted nucleoside, while charged nucleotide product remained trapped in the cell. FACS screening of  $5 \times 10^5$  cells per round isolated a mutant with 30 fold preference for the analogue over the natural substrate.

Also worthy of a brief mention, an orotate phosphoribosyl transferase was evolved to function with its natural orotate substrate, but with a restricted selection of amino acids.<sup>200</sup> Combinatorial and site directed mutagenesis created an OPRT homologue with a total of 73 amino acid substitutions. Mutant genes were placed in an *E. coli* lacking the OPRTase gene, auxotrophic for uracil, and selected by growth. The final variant used only 9 amino acids in 88% of the 188 positions and ostensibly retained catalytic function, although a kinetic comparison to wild type enzyme was not made.

In the only example of evolution of an enzyme other than a kinase, two nucleoside 2'-deoxyribosyltransferases (NdRT) were forced to accept dideoxy and didehydronucleosides by directed evolution.<sup>201</sup> Genes encoding NdRT were cloned from two different *Lactobacillus* species and randomly mutated using both epPCR with increased manganese, as well as the previously described GeneMorph II (Stratagene) error prone polymerase. Transformation into an *E. coli* auxotrophic for uracil allowed for a unique growth selection method. Genes *pyrC*, *codA*, and *ccd* which code for dihydroorotase, cytosine deaminase, and cytidine deaminase were deleted from the *E. coli*. This resulted in only a single biosynthetic route to UMP via the salvage enzyme UPRT, using uracil as substrate. Cells harboring the mutant NdRT genes were then

grown with 2',3'-dideoxyuracil (ddU) and cytosine supplements. Active NdRTs catalyzed the reaction  $ddU + C = ddC + U$ , thereby releasing uracil into the cell, and complementing the auxotroph. Ten clones survived from an estimated  $10^6$  colonies. The best mutant enzyme selected from each *Lactobacillus* species was purified and specific activity tested with various combinations of glycosyl donors and acceptors. Turnover increased by a factor greater than 100 for 2',3'-dideoxycytidine (ddC). Additionally, 2',3'-didehydro-2',3'- dideoxythymidine increased 250 fold despite the fact that this substrate was not selected for. Despite the versatility of NdRT in its utility as a nucleoside analogue biocatalyst, the potential of this enzyme for further evolution remains to be fully explored.

### **Dissertation Statement**

Nucleotide analogues are essential pharmaceuticals with challenges in preparation due to need for control over regio and stereochemistry, and protecting group manipulations. Enzymatic biosynthesis of these analogues provides solutions to these synthetic issues, but suffers from stringent substrate requirements of the enzymes. Directed evolution presents a possible solution to amend these enzymes toward alternative substrate acceptance. We propose a superior method for nucleotide analogue production, achieved through directed evolution of nucleotide synthesizing enzymes, using an *in vivo* screening method. Combination of the evolved biocatalyst into an engineered biosynthetic pathway will provide an efficient system for nucleotide analogue production from ribose.

Chapter II describes using directed evolution as a means to alter whole cell turnover rates and active site specificity of hypoxanthine phosphoribosyl transferase (HPRT) for increased production of the nucleotide analogue ribavirin monophosphate. A novel *in vivo* screening method for negative selection of variant HPRT enzymes is introduced, and the subsequent

validation of the screening method and characterization of the mutant encoded enzymes is described. Chapter III details the utilization of the evolved HPRT as a general nucleotide analogue biocatalyst. Also described is a method for assaying general nucleobase turnover, and end product purification and characterization. Chapter IV illustrates the addition of a series of enzymes which allow a multistep synthesis of nucleotides from ribose. Inclusion of additional enzymes for cofactor recycling adds utility and relieves byproduct inhibition. Finally, stability and reusability of the pathway is increased by self-immobilization in a cross-linked enzyme aggregate.

The sum of this work presents a significant improvement over current methods in nucleotide synthesis by alleviating current impediments due to lack of control over regio and stereochemistry, and by eliminating need for isolation of unstable intermediates, thus providing a green synthetic alternative in the synthesis of nucleotide analogues.

## References

1. Strauss, J.; Strauss, E., *Viruses and Human Disease*. Academic Press: San Diego, 2002.
2. [www.who.int](http://www.who.int)
3. Witkowski, J. T.; Simon, L. N.; Sidwell, R. W.; Robins, R. K., Design, Synthesis, and Broad-Spectrum Antiviral Activity of 1-Beta-D-ribofuranosyl-1,2,4-Triazole-3-Carboxamide and Related Nucleosides. *Journal of Medicinal Chemistry* **1972**, 15, (11), 1150.
4. Blackburn, G. M.; Gait, M. J., *Nucleic Acids in Chemistry and Biology*. Oxford University Press: New York, 1990; p 446.
5. Ito, Y.; Nii, Y.; Kobayashi, S.; Ohno, M., Regioselective Synthesis of Virazole Using Benzyl Cyanofornate as a Synthone. *Tetrahedron Letters* **1979**, (27), 2521-2524.
6. Allen, L. B.; Boswell, K. H.; Khwaja, T. A.; Meyer, R. B.; Sidwell, R. W.; Witkowski, J. T.; Christensen, L. F.; Robins, R. K., Synthesis and Anti-Viral Activity of Some Phosphates of Broad-Spectrum Anti-Viral Nucleoside, 1-Beta-D-Ribofuranosyl-1,2,4-Triazole-3-Carboxamide (Ribavirin). *Journal of Medicinal Chemistry* **1978**, 21, (8), 742-746.

7. Hennen, W. J.; Wong, C. H., A New Method for the Enzymatic-Synthesis of Nucleosides Using Purine Nucleoside Phosphorylase. *Journal of Organic Chemistry* **1989**, 54, (19), 4692-4695.
8. Krenitsky, T. A.; Elion, G. B.; Strelitz, R. A.; Hitching, G. H., Ribonucleosides of Allopurinol and Oxallopurinol - Isolation from Human Urine Enzymatic Synthesis and Characterization. *Journal of Biological Chemistry* **1967**, 242, (11), 2675.
9. Krenitsky, T. A.; Koszalka, G. W.; Tuttle, J. V., Purine Nucleoside Synthesis, an Efficient Method Employing Nucleoside Phosphorylases. *Biochemistry* **1981**, 20, (12), 3615-3621.
10. Ling, F.; Inoue, Y.; Kimura, A., Purification and Characterization of a Novel Nucleoside Phosphorylase from a Klebsiella Sp and Its Use in the Enzymatic Production of Adenine-Arabinoside. *Applied and Environmental Microbiology* **1990**, 56, (12), 3830-3834.
11. Chapeau, M. C.; Marnett, L. J., Enzymatic-Synthesis of Purine Deoxynucleoside Adducts. *Chemical Research in Toxicology* **1991**, 4, (6), 636-638.
12. Hori, N.; Uehara, K.; Mikami, Y., Enzymatic-Synthesis of 5-Methyluridine from Adenosine and Thymine with High-Efficiency. *Bioscience Biotechnology and Biochemistry* **1992**, 56, (4), 580-582.
13. Chae, W. G.; Cauchon, N. S.; Kozlowski, J. F.; Chang, C. J., Facile Synthesis of 2',5'-Dideoxy-5-Fluorouridine by Thymidine Phosphorylase. *Journal of Organic Chemistry* **1992**, 57, (3), 1002-1005.
14. Burns, C. L.; Stclair, M. H.; Frick, L. W.; Spector, T.; Averett, D. R.; English, M. L.; Holmes, T. J.; Krenitsky, T. A.; Koszalka, G. W., Novel 6-Alkoypurine 2',3'-Dideoxynucleosides as Inhibitors of the Cytopathic Effect of the Human-Immunodeficiency-Virus. *Journal of Medicinal Chemistry* **1993**, 36, (3), 378-384.
15. Chae, W. G.; Chan, T. C. K.; Chang, C. J., Facile synthesis of 5'-deoxy- and 2',5'-dideoxy-6-thiopurine nucleosides by nucleoside phosphorylases. *Tetrahedron* **1998**, 54, (30), 8661-8670.
16. Taverna-Porro, M.; Bouvier, L. A.; Pereira, C. A.; Montserrat, J. M.; Iribarren, A. M., Chemoenzymatic preparation of nucleosides from furanoses. *Tetrahedron Letters* **2008**, 49, (16), 2642-2645.
17. Utagawa, T.; Morisawa, H.; Yamanaka, S.; Yamazaki, A.; Hirose, Y., Enzymatic-Synthesis of Nucleoside Antibiotics .6. Enzymatic-Synthesis of Virazole by Purine Nucleoside Phosphorylase of Enterobacter-Aerogenes. *Agricultural and Biological Chemistry* **1986**, 50, (1), 121-126.
18. Shirae, H.; Yokozeki, K.; Kubota, K., Enzymatic Production of Ribavirin. *Agricultural and Biological Chemistry* **1988**, 52, (1), 295-296.

19. Shirae, H.; Yokozeki, K.; Kubota, K., Enzymatic Production of Ribavirin from Orotidine by *Erwinia-Carotovora* Aj-2992. *Agricultural and Biological Chemistry* **1988**, 52, (6), 1499-1504.
20. Shirae, H.; Yokozeki, K.; Uchiyama, M.; Kubota, K., Enzymatic Production of Ribavirin from Purine Nucleosides by *Brevibacterium-Acetylicum* Atcc-954. *Agricultural and Biological Chemistry* **1988**, 52, (7), 1777-1783.
21. Shirae, H.; Yokozeki, K.; Kubota, K., Enzymatic Production of Ribavirin from Pyrimidine Nucleosides by *Enterobacter-Aerogenes* Aj-11125. *Agricultural and Biological Chemistry* **1988**, 52, (5), 1233-1237.
22. Utagawa, T.; Morisawa, H.; Miyoshi, T.; Yoshinaga, F.; Yamazaki, A.; Mitsugi, K., Novel and Simple Method for the Preparation of Adenine-Arabinoside by Bacterial Transglycosylation Reaction. *Febs Letters* **1980**, 109, (2), 261-263.
23. Morisawa, H.; Utagawa, T.; Miyoshi, T.; Yoshinaga, F.; Yamazaki, A.; Mitsugi, K., New Method for the Synthesis of Some 9-Beta-D-Arabinofuranosylpurines by a Combination of Chemical and Enzymatic-Reactions. *Tetrahedron Letters* **1980**, 21, (5), 479-482.
24. Utagawa, T.; Morisawa, H.; Nakamatsu, T.; Yamazaki, A.; Yamanaka, S., Enzymatic-Synthesis of Purine 2'-Amino-2'-Deoxyriboside. *Febs Letters* **1980**, 119, (1), 101-104.
25. Utagawa, T.; Morisawa, H.; Yamanaka, S.; Yamazaki, A.; Yoshinaga, F.; Hirose, Y., Enzymatic-Synthesis of Nucleoside Antibiotics .2. Microbial Synthesis of Purine Arabinosides and Their Biological-Activity. *Agricultural and Biological Chemistry* **1985**, 49, (7), 2167-2171.
26. Heidelberger, C.; Parsons, D. G.; Remy, D. C., Syntheses of 5-trifluoromethyluracil and 5-trifluoromethyl-2'-deoxyuridine. *Journal of Medicinal Chemistry* **1964**, 7, (1), 1-5.
27. Kamimura, A.; Mitsugi, K.; Okumura, S., Bacterial Synthesis of Nucleosides by Pentosyl Transfer-Reaction from Nucleoside to Base. *Agricultural and Biological Chemistry* **1973**, 37, (9), 2063-2072.
28. Nishii, M.; Fujihara, Y.; Inagaki, J.; Ogawa, K.; Nohara, F.; Ubukata, M.; Isono, K., Microbial Preparation of Guanosine 7-N-Oxide. *Agricultural and Biological Chemistry* **1986**, 50, (10), 2697-2699.
29. Shirae, H.; Kobayashi, K.; Shiragami, H.; Irie, Y.; Yasuda, N.; Yokozeki, K., Production of 2',3'-Dideoxyadenosine and 2',3'-Dideoxyinosine from 2',3'-Dideoxyuridine and the Corresponding Purine-Bases by Resting Cells of *Escherichia-Coli* Aj-2595. *Applied and Environmental Microbiology* **1989**, 55, (2), 419-424.
30. Murakami, K.; Shirasaka, T.; Yoshioka, H.; Kojima, E.; Aoki, S.; Ford, H.; Driscoll, J. S.; Kelley, J. A.; Mitsuya, H., *Escherichia-Coli* Mediated Biosynthesis and In vitro Anti-Hiv Activity of Lipophilic 6-Halo-2',3'-Dideoxypurine Nucleosides. *Journal of Medicinal Chemistry* **1991**, 34, (5), 1606-1612.

31. Pal, S.; Nair, V., Biocatalytic synthesis of base-modified 2'-deoxy-beta-D-ribonucleosides with bacterial whole cells. *Biotechnology Letters* **1997**, 19, (4), 349-351.
32. Freeman, G. A.; Shaver, S. R.; Rideout, J. L.; Short, S. A., 2-Amino-9-(3-Azido-2,3-Dideoxy-Beta-D-Erythro-Pentofuranosyl)-6-Substituted-9h-Purines - Synthesis and Anti-Hiv Activity. *Bioorganic & Medicinal Chemistry* **1995**, 3, (4), 447-458.
33. Muller, M.; Hutchinson, L. K.; Guengerich, F. P., Addition of deoxyribose to guanine and modified DNA bases by *Lactobacillus helveticus* trans-N-deoxyribosylase. *Chemical Research in Toxicology* **1996**, 9, (7), 1140-1144.
34. Schnetz-Boutaud, N.; Chapeau, M.-C.; Marnett, L. J., Enzymatic Synthesis of M1G-Deoxyribose. *Current Protocols in Nucleic Acid Chemistry* **2000**, 1.2, 121-128.
35. Okuyama, K.; Shibuya, S.; Hamamoto, T.; Noguchi, T., Enzymatic synthesis of 2'-deoxyguanosine with nucleoside deoxyribosyltransferase-II. *Bioscience Biotechnology and Biochemistry* **2003**, 67, (5), 989-995.
36. Way, J.; Parks, R. E., Enzymatic Synthesis of 5'-phosphate Nucleotides of Purine Analogues. *Journal of Biological Chemistry* **1958**, 231, 467-480.
37. Da Costa, C. P.; Fedor, M. J.; Scott, L. G., 8-Azaguanine reporter of purine ionization states in structured RNAs. *Journal of the American Chemical Society* **2007**, 129, (11), 3426-3432.
38. Hennig, M.; Scott, L. G.; Sperling, E.; Bermel, W.; Williamson, J. R., Synthesis of 5-fluoropyrimidine Nucleotides as sensitive NMR probes of RNA structure. *Journal of the American Chemical Society* **2007**, 129, (48), 14911-14921.
39. Scott, L. G.; Geierstanger, B. H.; Williamson, J. R.; Hennig, M., Enzymatic synthesis and F-19 NMR studies of 2-fluoroadenine-substituted RNA. *Journal of the American Chemical Society* **2004**, 126, (38), 11776-11777.
40. Vanhercke, T.; Ampe, C.; Tirry, L.; Denolf, P., Reducing mutational bias in random protein libraries. *Analytical Biochemistry* **2005**, 339, (1), 9-14.
41. Bichet, A.; Bureik, M.; Lenz, N.; Bernhardt, R., The "bringer" strategy. *Applied Biochemistry and Biotechnology* **2004**, 117, (2), 115-122.
42. Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D. X.; Vogel, A., Expanding the range of substrate acceptance of enzymes: Combinatorial active-site saturation test. *Angewandte Chemie-International Edition* **2005**, 44, (27), 4192-4196.
43. DeSantis, G.; Wong, K.; Farwell, B.; Chatman, K.; Zhu, Z. L.; Tomlinson, G.; Huang, H. J.; Tan, X. Q.; Bibbs, L.; Chen, P.; Kretz, K.; Burk, M. J., Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM). *Journal of the American Chemical Society* **2003**, 125, (38), 11476-11477.

44. Stemmer, W. P. C., Rapid Evolution of a Protein in-Vitro by DNA Shuffling. *Nature* **1994**, 370, (6488), 389-391.
45. Kuchner, O.; Arnold, F. H., Directed evolution of enzyme catalysts. *Trends in Biotechnology* **1997**, 15, (12), 523-530.
46. Valetti, F.; Gilardi, G., Directed evolution of enzymes for product chemistry. *Natural Product Reports* **2004**, 21, (4), 490-511.
47. Coco, W. M.; Levinson, W. E.; Crist, M. J.; Hektor, H. J.; Darzins, A.; Pienkos, P. T.; Squires, C. H.; Monticello, D. J., DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nature Biotechnology* **2001**, 19, (4), 354-359.
48. Zhao, H. M.; Giver, L.; Shao, Z. X.; Affholter, J. A.; Arnold, F. H., Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nature Biotechnology* **1998**, 16, (3), 258-261.
49. Neylon, C., Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Research* **2004**, 32, (4), 1448-1459.
50. Ostermeier, M.; Shim, J. H.; Benkovic, S. J., A combinatorial approach to hybrid enzymes independent of DNA homology. *Nature Biotechnology* **1999**, 17, (12), 1205-1209.
51. Sieber, V.; Martinez, C. A.; Arnold, F. H., Libraries of hybrid proteins from distantly related sequences. *Nature Biotechnology* **2001**, 19, (5), 456-460.
52. Zhao, H. M.; Chockalingam, K.; Chen, Z. L., Directed evolution of enzymes and pathways for industrial biocatalysis. *Current Opinion in Biotechnology* **2002**, 13, (2), 104-110.
53. Coco, W. M.; Encell, L. P.; Levinson, W. E.; Crist, M. J.; Loomis, A. K.; Licato, L. L.; Arensdorf, J. J.; Sica, N.; Pienkos, P. T.; Monticello, D. J., Growth factor engineering by degenerate homoduplex gene family recombination. *Nature Biotechnology* **2002**, 20, (12), 1246-1250.
54. Zha, D. X.; Eipper, A.; Reetz, M. T., Assembly of designed oligonucleotides as an efficient method for gene recombination: A new tool in directed evolution. *ChemBiochem* **2003**, 4, (1), 34-39.
55. Demirjian, D. C.; Shah, P. C.; Moris-Varas, F., Screening for novel enzymes. *Biocatalysis - from Discovery to Application* **1999**, 200, 1-29.
56. Boersma, Y. L.; Droge, M. J.; Quax, W. J., Selection strategies for improved biocatalysts. *Federation of European Biochemical Societies Journal* **2007**, 274, (9), 2181-2195.
57. Wahler, D.; Reymond, J. L., Novel methods for biocatalyst screening. *Current Opinion in Chemical Biology* **2001**, 5, (2), 152-158.



58. Olsen, M.; Iverson, B.; Georgiou, G., High-throughput screening of enzyme libraries. *Current Opinion in Biotechnology* **2000**, 11, (4), 331-337.
59. Arnold, F. H.; Georgiou, G., *Directed Enzyme Evolution*. Humana Press, Inc.: Austin, TX, 2003.
60. Hoseki, J.; Yano, T.; Koyama, Y.; Kuramitsu, S.; Kagamiyama, H., Directed evolution of thermostable kanamycin-resistance gene: A convenient selection marker for *Thermus thermophilus*. *Journal of Biochemistry* **1999**, 126, (5), 951-956.
61. Haruki, M.; Noguchi, E.; Akasako, A.; Oobatake, M.; Itaya, M.; Kanaya, S., A Novel Strategy for Stabilization of *Escherichia-Coli* Ribonuclease Hi Involving a Screen for an Intragenic Suppressor of Carboxyl-Terminal Deletions. *Journal of Biological Chemistry* **1994**, 269, (43), 26904-26911.
62. Akanuma, S.; Yamagishi, A.; Tanaka, N.; Oshima, T., Serial increase in the thermal stability of 3-isopropylmalate dehydrogenase from *Bacillus subtilis* by experimental evolution. *Protein Science* **1998**, 7, (3), 698-705.
63. Kotsuka, T.; Akanuma, S.; Tomuro, M.; Yamagishi, A.; Oshima, T., Further stabilization of 3-isopropylmalate dehydrogenase of an extreme thermophile, *Thermus thermophilus*, by a suppressor mutation method. *Journal of Bacteriology* **1996**, 178, (3), 723-727.
64. Tamakoshi, M.; Yamagishi, A.; Oshima, T., Screening of Stable Proteins in an Extreme Thermophile, *Thermus-Thermophilus*. *Molecular Microbiology* **1995**, 16, (5), 1031-1036.
65. Yano, T.; Oue, S.; Kagamiyama, H., Directed evolution of an aspartate aminotransferase with new substrate specificities. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, 95, (10), 5511-5515.
66. Oue, S.; Okamoto, A.; Yano, T.; Kagamiyama, H., Redesigning the substrate specificity of an enzyme by cumulative effects of the mutations of non-active site residues. *Journal of Biological Chemistry* **1999**, 274, (4), 2344-2349.
67. Rothman, S. C.; Kirsch, J. F., How does an enzyme evolved in vitro compare to naturally occurring homologs possessing the targeted function? Tyrosine aminotransferase from aspartate aminotransferase. *Journal of Molecular Biology* **2003**, 327, (3), 593-608.
68. Evnin, L. B.; Vasquez, J. R.; Craik, C. S., Substrate-Specificity of Trypsin Investigated by Using a Genetic Selection. *Proceedings of the National Academy of Sciences of the United States of America* **1990**, 87, (17), 6659-6663.
69. Parales, R. E.; Ditty, J. L., Laboratory evolution of catabolic enzymes and pathways. *Current Opinion in Biotechnology* **2005**, 16, (3), 315-325.
70. Yano, T.; Kagamiyama, H., Directed evolution of ampicillin-resistant activity from a functionally unrelated DNA fragment: A laboratory model of molecular evolution.

*Proceedings of the National Academy of Sciences of the United States of America* **2001**, 98, (3), 903-907.

71. Barriault, D.; Plante, M. M.; Sylvestre, M., Family shuffling of a targeted bphA region to engineer biphenyl dioxygenase. *Journal of Bacteriology* **2002**, 184, (14), 3794-3800.
72. Christians, F. C.; Loeb, L. A., Novel human DNA alkyltransferases obtained by random substitution and genetic selection in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, 93, (12), 6124-6128.
73. Encell, L. P.; Loeb, L. A., Redesigning the substrate specificity of human O-6-alkylguanine-DNA alkyltransferase. Mutants with enhanced repair of O-4-methylthymine. *Biochemistry* **1999**, 38, (37), 12097-12103.
74. Landis, D. M.; Loeb, L. A., Random sequence mutagenesis and resistance to 5-fluorouridine in human thymidylate synthases. *Journal of Biological Chemistry* **1998**, 273, (40), 25809-25817.
75. Gulick, A. M.; Fahl, W. E., Forced Evolution of Glutathione-S-Transferase to Create a More Efficient Drug Detoxication Enzyme. *Proceedings of the National Academy of Sciences of the United States of America* **1995**, 92, (18), 8140-8144.
76. Cramer, A.; Dawes, G.; Rodriguez, E.; Silver, S.; Stemmer, W. P. C., Molecular evolution of an arsenate detoxification pathway DNA shuffling. *Nature Biotechnology* **1997**, 15, (5), 436-438.
77. Black, M. E.; Newcomb, T. G.; Wilson, H. M. P.; Loeb, L. A., Creation of drug-specific herpes simplex virus type 1 thymidine kinase mutants for gene therapy. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, 93, (8), 3525-3529.
78. Christians, F. C.; Scapozza, L.; Cramer, A.; Folkers, G.; Stemmer, W. P. C., Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling. *Nature Biotechnology* **1999**, 17, (3), 259-264.
79. Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G., Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99, (17), 11020-11024.
80. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G., Expanding the genetic code of Escherichia coli. *Science* **2001**, 292, (5516), 498-500.
81. Wahler, D.; Reymond, J. L., High-throughput screening for biocatalysts. *Current Opinion in Biotechnology* **2001**, 12, (6), 535-544.

82. Wang, C. W.; Oh, M. K.; Liao, J. C., Directed evolution of metabolically engineered *Escherichia coli* for carotenoid production. *Biotechnology Progress* **2000**, 16, (6), 922-926.
83. Wang, C. W.; Liao, J. C., Alteration of product specificity of *Rhodobacter sphaeroides* phytoene desaturase by directed evolution. *Journal of Biological Chemistry* **2001**, 276, (44), 41161-41164.
84. Umeno, D.; Arnold, F. H., Evolution of a pathway to novel long-chain carotenoids. *Journal of Bacteriology* **2004**, 186, (5), 1531-1536.
85. Albrecht, M.; Takaichi, S.; Steiger, S.; Wang, Z. Y.; Sandmann, G., Novel hydroxycarotenoids with improved antioxidative properties produced by gene combination in *Escherichia coli*. *Nature Biotechnology* **2000**, 18, (8), 843-846.
86. Lee, P. C.; Momen, A. Z. R.; Mijts, B. N.; Schmidt-Dannert, C., Biosynthesis of structurally novel carotenoids in *Escherichia coli*. *Chemistry & Biology* **2003**, 10, (5), 453-462.
87. Schmidt-Dannert, C., Engineering novel carotenoids in microorganisms. *Current Opinion in Biotechnology* **2000**, 11, (3), 255-261.
88. Schmidt-Dannert, C.; Umeno, D.; Arnold, F. H., Molecular breeding of carotenoid biosynthetic pathways. *Nature Biotechnology* **2000**, 18, (7), 750-753.
89. Stefan, A.; Radeghieri, A.; Rodriguez, A. G. V. Y.; Hochkoeppler, A., Directed evolution of beta-galactosidase from *Escherichia coli* by mutator strains defective in the 3' → 5' exonuclease activity of DNA polymerase III. *Febs Letters* **2001**, 493, (2-3), 139-143.
90. Allen, S. J.; Holbrook, J. J., Production of an activated form of *Bacillus stearothermophilus* L-2-hydroxyacid dehydrogenase by directed evolution. *Protein Engineering* **2000**, 13, (1), 5-7.
91. Matsumura, I.; Wallingford, J. B.; Surana, N. K.; Vize, P. D.; Ellington, A. D., Directed evolution of the surface chemistry of the reporter enzyme beta-glucuronidase. *Nature Biotechnology* **1999**, 17, (7), 696-701.
92. Cramer, A.; Whitehorn, E. A.; Tate, E.; Stemmer, W. P. C., Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nature Biotechnology* **1996**, 14, (3), 315-319.
93. Ehrig, T.; Okane, D. J.; Prendergast, F. G., Green-Fluorescent Protein Mutants with Altered Fluorescence Excitation-Spectra. *Febs Letters* **1995**, 367, (2), 163-166.
94. Heim, R.; Tsien, R. Y., Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology* **1996**, 6, 178-182.

95. Strack, R. L.; Strongin, D. E.; Bhattacharyya, D.; Tao, W.; Berman, A.; Broxmeyer, H. E.; Keenan, R. J.; Glick, B. S., A noncytotoxic DsRed variant for whole-cell labeling. *Nature Methods* **2008**, 5, (11), 955-957.
96. Strongin, D. E.; Bevis, B.; Khuong, N.; Downing, M. E.; Strack, R. L.; Sundaram, K.; Glick, B. S.; Keenan, R. J., Structural rearrangements near the chromophore influence the maturation speed and brightness of DsRed variants. *Protein Engineering Design & Selection* **2007**, 20, (11), 525-534.
97. Campbell, R. E.; Tour, O.; Palmer, A. E.; Steinbach, P. A.; Baird, G. S.; Zacharias, D. A.; Tsien, R. Y., A monomeric red fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99, (12), 7877-7882.
98. Strack, R. L.; Bhattacharyya, D.; Glick, B. S.; Keenan, R. J., Noncytotoxic orange and red/green derivatives of DsRed-Express2 for whole-cell labeling. *Bmc Biotechnology* **2009**, 9, -.
99. Tersikh, A.; Fradkov, A.; Ermakova, G.; Zaraksky, A.; Tan, P.; Kajava, A. V.; Zhao, X. N.; Lukyanov, S.; Matz, M.; Kim, S.; Weissman, I.; Siebert, P., "Fluorescent timer": Protein that changes color with time. *Science* **2000**, 290, (5496), 1585-1588.
100. Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N.; Palmer, A. E.; Tsien, R. Y., Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnology* **2004**, 22, 1567-1572.
101. Ghadessy, F. J.; Ramsay, N.; Boudsocq, F.; Loakes, D.; Brown, A.; Iwai, S.; Vaisman, A.; Woodgate, R.; Holliger, P., Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution. *Nature Biotechnology* **2004**, 22, (6), 755-759.
102. Santoro, S. W.; Wang, L.; Herberich, B.; King, D. S.; Schultz, P. G., An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nature Biotechnology* **2002**, 20, (10), 1044-1048.
103. Wang, J. D.; Herman, C.; Tipton, K. A.; Gross, C. A.; Weissman, J. S., Directed evolution of substrate-optimized GroEL/S chaperonins. *Cell* **2002**, 111, 1027-1039.
104. Pedelacq, J. D.; Piltch, E.; Liang, E. C.; Berendzen, J.; Kim, C. Y.; Rho, B. S.; Park, M. S.; Terwilliger, T. C.; Waldo, G. S., Engineering soluble proteins for structural genomics. *Nature Biotechnology* **2002**, 20, (9), 927-932.
105. Waldo, G. S., Genetic screens and directed evolution for protein solubility. *Current Opinion in Chemical Biology* **2003**, 7, (1), 33-38.
106. Waldo, G. S.; Standish, B. M.; Berendzen, J.; Terwilliger, T. C., Rapid protein-folding assay using green fluorescent protein. *Nature Biotechnology* **1999**, 17, (7), 691-695.
107. Joo, H.; Lin, Z. L.; Arnold, F. H., Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation. *Nature* **1999**, 399, (6737), 670-673.

108. Joo, H.; Arisawa, A.; Lin, Z. L.; Arnold, F. H., A high-throughput digital imaging screen for the discovery and directed evolution of oxygenases. *Chemistry & Biology* **1999**, *6*, (10), 699-706.
109. Badalassi, F.; Wahler, D.; Klein, G.; Crotti, P.; Reymond, J. L., A versatile periodate-coupled fluorogenic assay for hydrolytic enzymes. *Angewandte Chemie-International Edition* **2000**, *39*, (22), 4067-4070.
110. Liu, L. F.; Li, Y. F.; Liotta, D.; Lutz, S., Directed evolution of an orthogonal nucleoside analog kinase via fluorescence-activated cell sorting. *Nucleic Acids Research* **2009**, *37*, (13), 4472-4481.
111. Aharoni, A.; Thieme, K.; Chiu, C. P. C.; Buchini, S.; Lairson, L. L.; Chen, H. M.; Strynadka, N. C. J.; Wakarchuk, W. W.; Withers, S. G., High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nature Methods* **2006**, *3*, (8), 609-614.
112. Bergquist, P. L.; Reeves, R. A.; Gibbs, M. D., Degenerate oligonucleotide gene shuffling (DOGS) and random drift mutagenesis (RNDM): Two complementary techniques for enzyme evolution. *Biomolecular Engineering* **2005**, *22*, (1-3), 63-72.
113. Griswold, K. E.; Aiyappan, N. S.; Iverson, B. L.; Georgiou, G., The evolution of catalytic efficiency and substrate promiscuity in human theta class 1-1 glutathione transferase. *Journal of Molecular Biology* **2006**, *364*, (3), 400-410.
114. Griswold, K. E.; Kawarasaki, Y.; Ghoneim, N.; Benkovic, S. J.; Iverson, B. L.; Georgiou, G., Evolution of highly active enzymes by homology-independent recombination. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, (29), 10082-10087.
115. Tawfik, D. S.; Griffiths, A. D., Man-made cell-like compartments for molecular evolution. *Nature Biotechnology* **1998**, *16*, (7), 652-656.
116. Griffiths, A. D.; Tawfik, D. S., Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. *Embo Journal* **2003**, *22*, (1), 24-35.
117. Aharoni, A.; Griffiths, A. D.; Tawfik, D. S., High-throughput screens and selections of enzyme-encoding genes. *Current Opinion in Chemical Biology* **2005**, *9*, (2), 210-216.
118. Bershtein, S.; Tawfik, D. S., Advances in laboratory evolution of enzymes. *Current Opinion in Chemical Biology* **2008**, *12*, (2), 151-158.
119. Griffiths, A. D.; Tawfik, D. S., Man-made enzymes - from design to in vitro compartmentalisation. *Current Opinion in Biotechnology* **2000**, *11*, (4), 338-353.
120. Griffiths, A. D.; Tawfik, D. S., Miniaturising the laboratory in emulsion droplets. *Trends in Biotechnology* **2006**, *24*, (9), 395-402.

121. Kelly, B. T.; Baret, J. C.; Taly, V.; Griffiths, A. D., Miniaturizing chemistry and biology in microdroplets. *Chemical Communications* **2007**, (18), 1773-1788.
122. Miller, O. J.; Bernath, K.; Agresti, J. J.; Amitai, G.; Kelly, B. T.; Mastrobattista, E.; Taly, V.; Magdassi, S.; Tawfik, D. S.; Griffiths, A. D., Directed evolution by in vitro compartmentalization. *Nature Methods* **2006**, 3, (7), 561-570.
123. Lebbink, J. H. G.; Kaper, T.; Bron, P.; van der Oost, J.; de Vos, W. M., Improving low-temperature catalysis in the hyperthermostable *Pyrococcus furiosus* beta-glucosidase CelB by directed evolution. *Biochemistry* **2000**, 39, (13), 3656-3665.
124. Miyazaki, K.; Wintrode, P. L.; Grayling, R. A.; Rubingh, D. N.; Arnold, F. H., Directed evolution study of temperature adaptation in a psychrophilic enzyme. *Journal of Molecular Biology* **2000**, 297, (4), 1015-1026.
125. May, O.; Nguyen, P. T.; Arnold, F. H., Inverting enantioselectivity by directed evolution of hydantoinase for improved production of L-methionine. *Nature Biotechnology* **2000**, 18, (3), 317-320.
126. Fong, S.; Machajewski, T. D.; Mak, C. C.; Wong, C. H., Directed evolution of D-2-keto-3-deoxy-6-phosphogluconate aldolase to new variants for the efficient synthesis of D- and L-sugars. *Chemistry & Biology* **2000**, 7, (11), 873-883.
127. Cherry, J. R.; Lamsa, M. H.; Schneider, P.; Vind, J.; Svendsen, A.; Jones, A.; Pedersen, A. H., Directed evolution of a fungal peroxidase. *Nature Biotechnology* **1999**, 17, (4), 379-384.
128. Nemoto, N.; MiyamotoSato, E.; Husimi, Y.; Yanagawa, H., In vitro virus: Bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *Febs Letters* **1997**, 414, (2), 405-408.
129. Roberts, R. W.; Szostak, J. W., RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, 94, (23), 12297-12302.
130. Roberts, R. W., Totally in vitro protein selection using mRNA-protein fusions and ribosome display. *Current Opinion in Chemical Biology* **1999**, 3, (3), 268-273.
131. Pluckthun, A.; Schaffitzel, C.; Hanes, J.; Jermutus, L., In vitro selection and evolution of proteins. *Advances in Protein Chemistry, Vol 55* **2001**, 55, 367-403.
132. Schaffitzel, C.; Hanes, J.; Jermutus, L.; Pluckthun, A., Ribosome display: an in vitro method for selection and evolution of antibodies from libraries. *Journal of Immunological Methods* **1999**, 231, (1-2), 119-135.
133. Mattheakis, L. C.; Bhatt, R. R.; Dower, W. J., An in-Vitro Polysome Display System for Identifying Ligands from Very Large Peptide Libraries. *Proceedings of the National Academy of Sciences of the United States of America* **1994**, 91, (19), 9022-9026.

134. Yan, X. H.; Xu, Z. R., Ribosome-display technology: applications for directed evolution of functional proteins. *Drug Discovery Today* **2006**, 11, (19-20), 911-916.
135. Sergeeva, A.; Kolonin, M. G.; Molldrem, J. J.; Pasqualini, R.; Arap, W., Display technologies: Application for the discovery of drug and gene delivery agents. *Advanced Drug Delivery Reviews* **2006**, 58, (15), 1622-1654.
136. Kondo, A.; Ueda, M., Yeast cell-surface display - applications of molecular display. *Applied Microbiology and Biotechnology* **2004**, 64, (1), 28-40.
137. Samuelson, P.; Gunneriusson, E.; Nygren, P. A.; Stahl, S., Display of proteins on bacteria. *Journal of Biotechnology* **2002**, 96, (2), 129-154.
138. Daugherty, P. S.; Chen, G.; Iverson, B. L.; Georgiou, G., Antibody selection by Escherichia coli surface display and fluorescence-activated cell sorting. *Abstracts of Papers of the American Chemical Society* **1997**, 213, 90-Biot.
139. Georgiou, G.; Stathopoulos, C.; Daugherty, P. S.; Nayak, A. R.; Iverson, B. L.; Curtiss, R., Display of heterologous proteins on the surface of microorganisms: From the screening of combinatorial libraries to live recombinant vaccines. *Nature Biotechnology* **1997**, 15, (1), 29-34.
140. Olsen, M. J.; Stephens, D. L.; Georgiou, G.; Iverson, B. L., A method for biocatalyst evolution using bacterial surface display. *Abstracts of Papers of the American Chemical Society* **1997**, 214, 35-Catl.
141. Jose, J., Autodisplay: efficient bacterial surface display of recombinant proteins. *Applied Microbiology and Biotechnology* **2006**, 69, (6), 607-614.
142. Boder, E. T.; Wittrup, K. D., Yeast surface display for directed evolution of protein expression, affinity, and stability. *Applications of Chimeric Genes and Hybrid Proteins, Pt C* **2000**, 328, 430-444.
143. Wittrup, K. D., Protein engineering by cell-surface display. *Current Opinion in Biotechnology* **2001**, 12, (4), 395-399.
144. Skerra, A., Engineered protein scaffolds for molecular recognition. *Journal of Molecular Recognition* **2000**, 13, (4), 167-187.
145. Hoess, R. H., Protein design and phage display. *Chemical Reviews* **2001**, 101, (10), 3205-3218.
146. Kehoe, J. W.; Kay, B. K., Filamentous phage display in the new millennium. *Chemical Reviews* **2005**, 105, (11), 4056-4072.

147. Scholle, M. D.; Kehoe, J. W.; Kay, B. K., Efficient construction of a large collection of phage-displayed combinatorial peptide libraries. *Combinatorial Chemistry & High Throughput Screening* **2005**, 8, (6), 545-551.
148. Paschke, M., Phage display systems and their applications. *Applied Microbiology and Biotechnology* **2006**, 70, (1), 2-11.
149. Forrer, P.; Jung, S.; Pluckthun, A., Beyond binding: using phage display to select for structure, folding and enzymatic activity in proteins. *Current Opinion in Structural Biology* **1999**, 9, (4), 514-520.
150. Oneil, K. T.; Hoess, R. H., Phage Display - Protein Engineering by Directed Evolution. *Current Opinion in Structural Biology* **1995**, 5, (4), 443-449.
151. Fernandez-Gacio, A.; Uguen, M.; Fastrez, J., Phage display as a tool for the directed evolution of enzymes. *Trends in Biotechnology* **2003**, 21, (9), 408-414.
152. Riechmann, L.; Winter, G., Novel folded protein domains generated by combinatorial shuffling of polypeptide segments. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, 97, (18), 10068-10073.
153. Hida, K.; Hanes, J.; Ostermeier, M., Directed evolution for drug and nucleic acid delivery. *Advanced Drug Delivery Reviews* **2007**, 59, (15), 1562-1578.
154. Hirose, A.; Esaka, Y.; Ohta, M.; Haraguchi, H., Online Hplc Determination of Enzymatic-Activity of Alkaline-Phosphatase in Natural-Water Using Spectrofluorometric Detection. *Chemistry Letters* **1993**, (2), 307-310.
155. Greenbaum, D.; Medzihradzky, K. F.; Burlingame, A.; Bogoy, M., Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chemistry & Biology* **2000**, 7, (8), 569-581.
156. Gruninger-Leitch, F.; Berndt, P.; Langen, H.; Nelboeck, P.; Dobeli, H., Identification of beta-secretase-like activity using a mass spectrometry-based assay system. *Nature Biotechnology* **2000**, 18, (1), 66-70.
157. Reetz, M. T.; Becker, M. H.; Klein, H. W.; Stockigt, D., A method for high-throughput screening of enantioselective catalysts. *Angewandte Chemie-International Edition* **1999**, 38, (12), 1758-1761.
158. Schrader, W.; Eipper, A.; Pugh, D. J.; Reetz, M. T., Second-generation MS-based high-throughput screening system for enantioselective catalysts and biocatalysts. *Canadian Journal of Chemistry-Revue Canadienne De Chimie* **2002**, 80, (6), 626-632.
159. Reetz, M. T.; Kuhling, K. M.; Deege, A.; Hinrichs, H.; Belder, D., Super-high-throughput screening of enantioselective catalysts by using capillary array electrophoresis. *Angewandte Chemie-International Edition* **2000**, 39, (21), 3891-+.



160. Reetz, M. T.; Kuhling, K. M.; Hinrichs, H.; Deege, A., Circular dichroism as a detection method in the screening of enantioselective catalysts. *Chirality* **2000**, 12, (5-6), 479-482.
161. Janes, L. E.; Lowendahl, A. C.; Kazlauskas, R. J., Quantitative screening of hydrolase libraries using pH indicators: Identifying active and enantioselective hydrolases. *Chemistry-a European Journal* **1998**, 4, (11), 2324-2331.
162. Reetz, M. T.; Becker, M. H.; Kuhling, K. M.; Holzwarth, A., Time-resolved IR-thermographic deflection and screening of enantioselectivity in catalytic reactions. *Angewandte Chemie-International Edition* **1998**, 37, (19), 2647-2650.
163. Taylor, S. J.; Morken, J. P., Thermographic selection of effective catalysts from an encoded polymer-bound library. *Science* **1998**, 280, (5361), 267-270.
164. Bornscheuer, U. T., Methods to increase enantioselectivity of lipases and esterases. *Current Opinion in Biotechnology* **2002**, 13, (6), 543-547.
165. Schmidt, M.; Bornscheuer, U. T., High-throughput assays for lipases and esterases. *Biomolecular Engineering* **2005**, 22, (1-3), 51-56.
166. Cherry, J. R.; Fidantsef, A. L., Directed evolution of industrial enzymes: an update. *Current Opinion in Biotechnology* **2003**, 14, (4), 438-443.
167. Bornscheuer, U. T.; Bessler, C.; Srinivas, R.; Krishna, S. H., Optimizing lipases and related enzymes for efficient application. *Trends in Biotechnology* **2002**, 20, (10), 433-437.
168. Uchiyama, H.; Inaoka, T.; Ohkuma-Soyejima, T.; Togame, H.; Shibana, Y.; Yoshimoto, T.; Kokubo, T., Directed evolution to improve the thermostability of prolyl endopeptidase. *Journal of Biochemistry* **2000**, 128, (3), 441-447.
169. Wintrode, P. L.; Miyazaki, K.; Arnold, F. H., Cold adaptation of a mesophilic subtilisin-like protease by laboratory evolution. *Journal of Biological Chemistry* **2000**, 275, (41), 31635-31640.
170. Tomatis, P. E.; Rasia, R. M.; Segovia, L.; Vila, A. J., Mimicking natural evolution in metallo-beta-lactamases through second-shell ligand mutations. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, 102, (39), 13761-13766.
171. Morillas, M.; McVey, C. E.; Brannigan, J. A.; Ladurner, A. G.; Forney, L. J.; Virden, R., Mutations of penicillin acylase residue B71 extend substrate specificity by decreasing steric constraints for substrate binding. *Biochemical Journal* **2003**, 371, 143-150.
172. Otten, L. G.; Sio, C. F.; Vrieling, J.; Cool, R. H.; Quax, W. J., Altering the substrate specificity of cephalosporin acylase by directed evolution of the beta-subunit. *Journal of Biological Chemistry* **2002**, 277, (44), 42121-42127.

173. Sio, C. F.; Riemens, A. M.; van der Laan, J. M.; Verhaert, R. M. D.; Quax, W. J., Directed evolution of a glutaryl acylase into an adipyl acylase. *European Journal of Biochemistry* **2002**, 269, (18), 4495-4504.
174. Hancock, S. M.; D Vaughan, M.; Withers, S. G., Engineering of glycosidases and glycosyltransferases. *Current Opinion in Chemical Biology* **2006**, 10, (5), 509-519.
175. Shaikh, F. A.; Withers, S. G., Teaching old enzymes new tricks: engineering and evolution of glycosidases and glycosyl transferases for improved glycoside synthesis. *Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire* **2008**, 86, (2), 169-177.
176. McLoughlin, S. Y.; Jackson, C.; Liu, J. W.; Ollis, D., Increased expression of a bacterial phosphotriesterase in Escherichia coli through directed evolution. *Protein Expression and Purification* **2005**, 41, (2), 433-440.
177. Mukhopadhyay, R.; Zhou, Y.; Rosen, B. P., Directed evolution of a yeast arsenate reductase into a protein-tyrosine phosphatase. *Journal of Biological Chemistry* **2003**, 278, (27), 24476-24480.
178. Kong, X. D.; Liu, Y.; Gou, X. J.; Zhang, H. Y.; Wang, X. P.; Zhang, J., Directed evolution of operon of trehalose-6-phosphate synthase/phosphatase from Escherichia coli. *Biochemical and Biophysical Research Communications* **2001**, 280, (1), 396-400.
179. Cirino, P. C.; Arnold, F. H., Protein engineering of oxygenases for biocatalysis. *Current Opinion in Chemical Biology* **2002**, 6, (2), 130-135.
180. Gillam, E. M. J., Engineering cytochrome P450 enzymes. *Chemical Research in Toxicology* **2008**, 21, (1), 220-231.
181. Gillam, E. M. J., Extending the capabilities of nature's most versatile catalysts: Directed evolution of mammalian xenobiotic-metabolizing P450s. *Archives of Biochemistry and Biophysics* **2007**, 464, (2), 176-186.
182. Tee, K. L.; Schwaneberg, U., Directed evolution of oxygenases: Screening systems, success stories and challenges. *Combinatorial Chemistry & High Throughput Screening* **2007**, 10, (3), 197-217.
183. Furukawa, K., Oxygenases and dehalogenases: Molecular approaches to efficient degradation of chlorinated environmental pollutants. *Bioscience Biotechnology and Biochemistry* **2006**, 70, (10), 2335-2348.
184. Kumar, S.; Halpert, J. R., Use of directed evolution of mammalian cytochromes P450 for investigating the molecular basis of enzyme function and generating novel biocatalysts. *Biochemical and Biophysical Research Communications* **2005**, 338, (1), 456-464.
185. Furukawa, K.; Suenaga, H.; Goto, M., Biphenyl dioxygenases: Functional versatilities and directed evolution. *Journal of Bacteriology* **2004**, 186, (16), 5189-5196.

186. Cherry, J. R., Directed evolution of microbial oxidative enzymes. *Current Opinion in Biotechnology* **2000**, 11, (3), 250-254.
187. Broo, K.; Larsson, A. K.; Jemth, P.; Mannervik, B., An ensemble of theta class glutathione transferases with novel catalytic properties generated by stochastic recombination of fragments of two mammalian enzymes. *Journal of Molecular Biology* **2002**, 318, (1), 59-70.
188. Hansson, L. O.; Bolton-Grob, R.; Massoud, T.; Mannervik, B., Evolution of differential substrate specificities in Mu class glutathione transferases probed by DNA shuffling. *Journal of Molecular Biology* **1999**, 287, (2), 265-276.
189. Bhuiya, M. W.; Liu, C. J., Engineering monolignol 4-O-methyltransferases to modulate lignin biosynthesis. *Journal of Biological Chemistry* **2009**.
190. Cohen, H. M.; Tawfik, D. S.; Griffiths, A. D., Altering the sequence specificity of HaeIII methyltransferase by directed evolution using in vitro compartmentalization. *Protein Engineering Design & Selection* **2004**, 17, (1), 3-11.
191. Bernath, K.; Hai, M. T.; Mastrobattista, E.; Griffiths, A. D.; Magdassi, S.; Tawfik, D. S., In vitro compartmentalization by double emulsions: sorting and gene enrichment by fluorescence activated cell sorting. *Analytical Biochemistry* **2004**, 325, (1), 151-157.
192. Juillerat, A.; Gronemeyer, T.; Keppler, A.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K., Directed evolution of O-6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. *Chemistry & Biology* **2003**, 10, (4), 313-317.
193. Langenhan, J. M.; Griffith, B. R.; Thorson, J. S., Neoglycorandomization and chemoenzymatic glycorandomization: Two complementary tools for natural product diversification. *Journal of Natural Products* **2005**, 68, (11), 1696-1711.
194. Griffith, B. R.; Langenhan, J. M.; Thorson, J. S., 'Sweetening' natural products via glycorandomization. *Current Opinion in Biotechnology* **2005**, 16, (6), 622-630.
195. Yang, J.; Hoffmeister, D.; Liu, L.; Fu, X.; Thorson, J. S., Natural product glycorandomization. *Bioorganic & Medicinal Chemistry* **2004**, 12, (7), 1577-1584.
196. Bolt, A.; Berry, A.; Nelson, A., Directed evolution of aldolases for exploitation in synthetic organic chemistry. *Archives of Biochemistry and Biophysics* **2008**, 474, (2), 318-330.
197. Samland, A. K.; Sprenger, G. A., Microbial aldolases as C-C bonding enzymes - unknown treasures and new developments. *Applied Microbiology and Biotechnology* **2006**, 71, (3), 253-264.
198. Jaeger, K. E.; Eggert, T.; Eipper, A.; Reetz, M. T., Directed evolution and the creation of enantioselective biocatalysts. *Applied Microbiology and Biotechnology* **2001**, 55, (5), 519-530.

199. Mueller-Cajar, O.; Morell, M.; Whitney, S. M., Directed evolution of Rubisco in *Escherichia coli* reveals a specificity-determining hydrogen bond in the form II enzyme. *Biochemistry* **2007**, 46, (49), 14067-14074.
200. Akanuma, S.; Kigawa, T.; Yokoyama, S., Combinatorial mutagenesis to restrict amino acid usage in an enzyme to a reduced set. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99, (21), 13549-13553.
201. Kaminski, P. A.; Dacher, P.; Dugue, L.; Pochet, S., In vivo reshaping the catalytic site of nucleoside 2'-deoxyribosyltransferase for dideoxy- and didehydronucleosides via a single amino acid substitution. *Journal of Biological Chemistry* **2008**, 283, (29), 20053-20059.

## CHAPTER II

### DIRECTED EVOLUTION AND *IN VIVO* SCREENING OF HPRT FOR RMP NUCLEOTIDE BIOSYNTHESIS

#### Introduction

The potential obstacles in nucleotide analogue chemical synthesis and the stringent substrate requirements of enzymatic methods highlight the need for using a directed evolution approach to improve biocatalysis. The relatively unexplored area of evolution of salvage enzymes, specifically, evolution of hypoxanthine phosphoribosyl pyrophosphate (HPRT), shows potential as a method for creation of a new biocatalyst for ribavirin monophosphate (RMP) biosynthesis. Random mutation in combination with a novel *in vivo* screening method will be used to isolate an efficient RMP producing mutant. RMP production will inhibit intracellular inosine monophosphate dehydrogenase (IMPDH) of the *E. coli* host, thus preventing further cell growth in minimal media. Prior to a detailed explanation of the screening procedure, a review of structural and biochemical information on HPRT, Ribavirin and RMP, and IMPDH are necessary.

#### Phosphoribosyl Transferases

From *E. coli* to humans, most organisms share similar pathways for intracellular biosynthesis of nucleotides. Nucleotides can be generated through either a stepwise *de novo* pathway, or by individual salvage enzymes. The *de novo* pathways for purines AMP or GMP consist of 17 steps catalyzed by 17 different enzymes, ultimately converting glucose to final nucleotide product. Conversely, salvage pathways use phosphoribosyl transferases to promote the condensation of nucleobase (hypoxanthine, guanine, adenine, xanthine, orotate, or uracil) to phosphoribosyl pyrophosphate (PRPP), yielding monophosphorylated nucleotide and

pyrophosphate. An estimated ninety percent of free purines in humans are recycled via these salvage pathways.<sup>1</sup>

### ***Purine phosphoribosyl transferases***

Purine phosphoribosyl transferases are the most extensively studied PRTs with over 40 published crystal structures of various proteins. Although the purine PRT enzymes have little sequence identity (<15%), all have a common core consisting of four- or five-stranded parallel  $\beta$ -sheets flanked by 3 to 4  $\alpha$ -helices.<sup>2, 3</sup> (Figure II-1) All have a conserved PRPP-binding site associated with a divalent metal, usually  $Mg^{2+}$ , and all have a similar, yet poorly conserved, flexible loop. As evident in Figure II-2, the loop becomes ordered when  $Mg\bullet PRPP$  binds, isolating the reaction from solvent during catalysis.<sup>2, 4</sup> Comparison of free and bound crystal structures reveals widespread conformational flexibility throughout catalysis.<sup>5</sup> An additional feature of PRTases is a non-conserved hood domain, which contributes to the active site. Located above the core fold, the hood includes residues from both the N- and C-termini. The structure of the hood is highly variable across the PRT family, and is partially responsible for regulation of purine binding. In purine or pyrimidine PRTases, the hood recognizes the appropriate base through specific hydrogen bonds, binds the nucleophilic substrate, and ligates the second  $Mg^{2+}$ .<sup>2</sup>

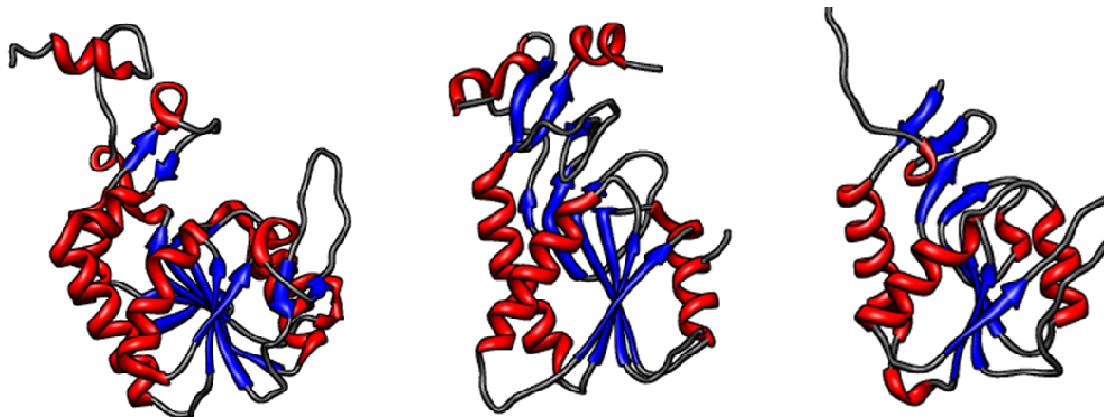
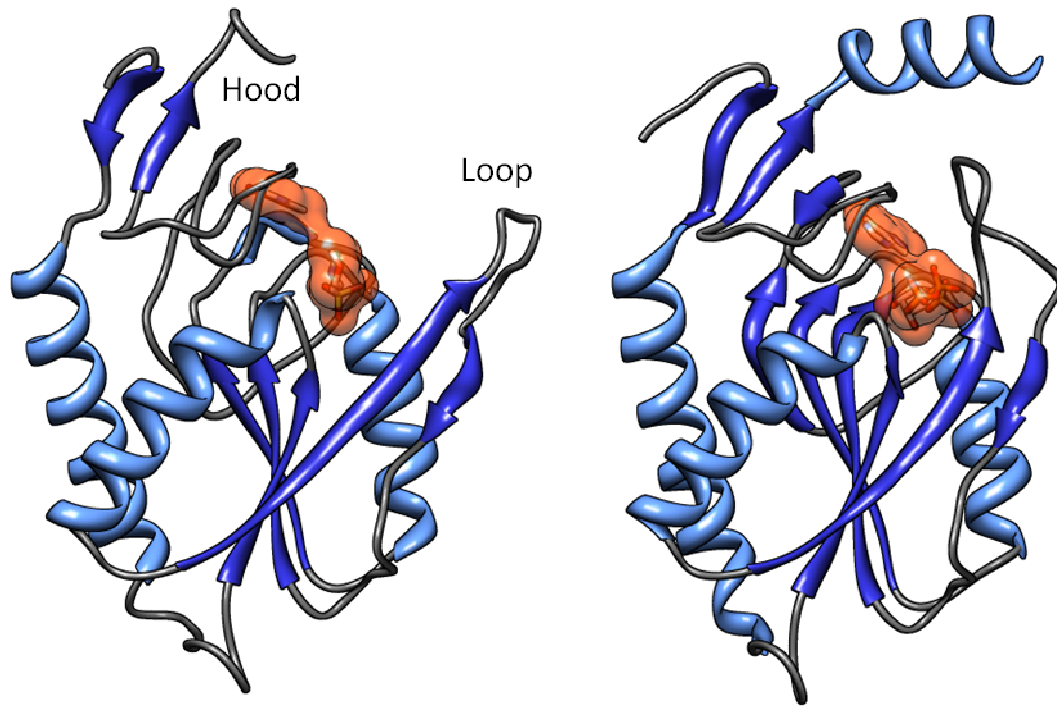


Figure II-1: Comparison of structural similarities in purine PRTs, adapted from Craig, S. P., *et al.*<sup>3</sup> From left to right; APRT from *Leishmani donovani*, HPRT from *Homo sapiens*, and XPRT from *E. coli*. Images were rendered from crystal structures obtained by Keough, D. T.,<sup>5</sup> Phillips, C. L.,<sup>6</sup> and Vos, S.<sup>7</sup> using UCSF Chimera software.<sup>8</sup>

### ***Hypoxanthine Phosphoribosyl Pyrophosphate Transferase (HPRT)***

In addition to the previously mentioned conserved structures of all PRTs, HPRT retains a conserved phosphate binding loop, and single lysine and aspartate residues throughout various species. The phosphate binding loop, consisting of a five amino acid segment, is necessary in forming hydrogen bonds between the phosphate oxygen atoms from the nucleotide and the nitrogen atoms of the peptide backbone. HPRT, as well as other 6-oxopurine PRTs, share a Lys135 residue, which forms a hydrogen bond to the 6-oxo oxygen of hypoxanthine. The conserved Asp107 is within hydrogen bonding distance to the N7 atom of the purine base and is thought to be catalytically important for deprotonation of the nucleobase.<sup>9</sup>



**Figure II-2: Open and closed forms of HPRT from *Trypanosoma cruzi*. Images adapted from crystal structure data obtained by Focia, *et al*,<sup>4,10</sup> using UCSF Chimera software.<sup>8</sup>**

The HPRT product, IMP, is a precursor for either of the two purine nucleotides; AMP or GMP. Although several organisms possess a direct salvage route of adenine to AMP, HPRT also contributes to AMP salvage through the combined actions of adenosine deaminase and PNP. Adenosine deaminase converts adenosine to inosine, and PNP cleaves the hypoxanthine base for later addition to PRPP via HPRT. Some organisms, such as *Plasmodium falciparum*, prefer this method over a direct APRT route.

In some organisms, including humans, a single PRT accepts both hypoxanthine and guanine, aptly named HGPRT (E.C. 2.4.2.8). *E. coli* has two 6-oxo PRTs, one for hypoxanthine (HPRT, E.C. 2.4.2.8) and another which accepts both xanthine and guanine (XGPRT, E.C. 2.4.2.22)<sup>9</sup> (Table II-1). HPRT from *E. coli* has 23% sequence identity with XGPRT, and 33% with



human HPRT. The 182 amino acid-long protein, with a mass of 20.6 kDa, forms a homotetramer in solution, held together through a cluster of six salt bridges at the subunit interfaces.

**Table II-1: Affinity for various purine nucleobases by HPRT and XGPRT from *E. coli*. Adapted from original data by Guddat, L. W., et al.<sup>9</sup>**

Substrate	Specific Activity ( $\mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ )	$K_{M(\text{app})}$ ( $\mu\text{M}$ )	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$k_{\text{cat}}/K_M$ ( $\mu\text{M} \cdot \text{s}^{-1}$ )
<b><i>E. coli</i> HPRT</b>				
hypoxanthine	177	13	59	4.9
guanine	30	294	10	0.03
xanthine	0.02	25	0.0008	0.0003
PRPP	177	192	50	0.26
<b><i>E. coli</i> XGRT</b>				
hypoxanthine	23	91	14	0.2
guanine	95	4	28	6.5
xanthine	114	31	38	1.2
PRPP	95	139	28	0.2

The forward reaction of HPRT proceeds through an  $S_N2$  mechanism with PRPP binding first, followed by hypoxanthine.<sup>3</sup> Deprotonation of N7 by Asp107 allows the concomitant nucleophilic addition of the N9 nitrogen of hypoxanthine to the anomeric position of PRPP. Pyrophosphate is then released followed by the rate-limiting dissociation of the nucleoside monophosphate.<sup>5, 11</sup>

Although the unique PRT enzymes are rather strict in the acceptance of nucleobase substrates, specificity can be significantly altered by single amino acid substitutions.<sup>12, 13</sup> HPRT enzymes have been mutated to GPRTs and vice versa through a single mutation. A chimera was made with *P. falciparum* and human enzymes resulting in a shift in substrate preference.<sup>14</sup> Notably, changes in the hood structure created an HXGPRT with significant reaction rates for all three 6-oxo purines.

## Ribavirin and Ribavirin Monophosphate (RMP)

Ribavirin (1- $\beta$ -D-ribofuranosyl-1,2,3-triazole-3-carboxamide) (Figure II-3) is a broad-spectrum antiviral agent first synthesized in 1972, under the trade name Virazole, by J. T. Witkowski.<sup>15</sup> Antiviral activity of previously synthesized nucleosides pyrazomycin and showdomycin fueled the search for similar 5-membered heterocyclic antiviral nucleosides. The original publication describes two methods of synthesis: an acid-catalyzed fusion procedure and a silylation-glycosylation procedure. Subsequent publications include enzymatic synthesis of the nucleoside through purine nucleoside phosphorylase reactions with ribose-1-phosphate intermediate isolation<sup>16</sup> or without.<sup>17</sup> Several whole cell synthetic examples have also been published<sup>18-21</sup> with only one describing an end product purification.<sup>22</sup>

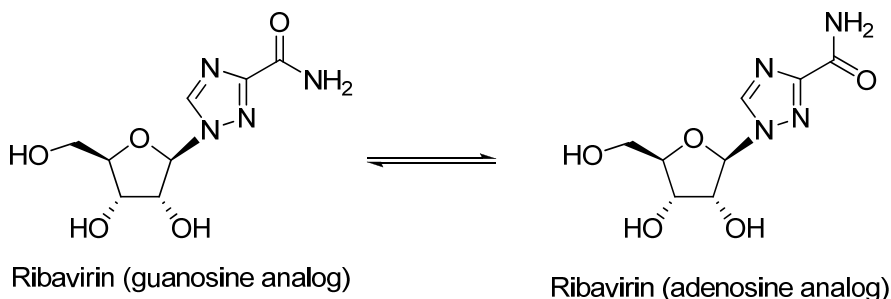


Figure II-3: Ribavirin tautomers as analogues of guanosine and adenosine.

The initial report of ribavirin synthesis included impressive results of broad spectrum antiviral activity *in vitro* for both DNA viruses and RNA viruses.<sup>15</sup> It has shown antiviral activity against influenza, hepatitis B, polio, rabies, measles, smallpox, and respiratory syncytial virus, and is currently approved for treatment of hepatitis C viral infections. Investigations into the mechanism for inhibition revealed antiviral activity was substantially reversed by xanthosine and guanosine. This suggests antiviral activity might be due to inhibition of viral IMPDH, preventing *de novo* GMP biosynthesis, and subsequent viral replication. Purified IMPDH from *E. coli* did not

show inhibition with ribavirin, but rather with the 5' monophosphate metabolite, confirming IMPDH inhibition as a possible mechanism for antiviral activity.<sup>23</sup>

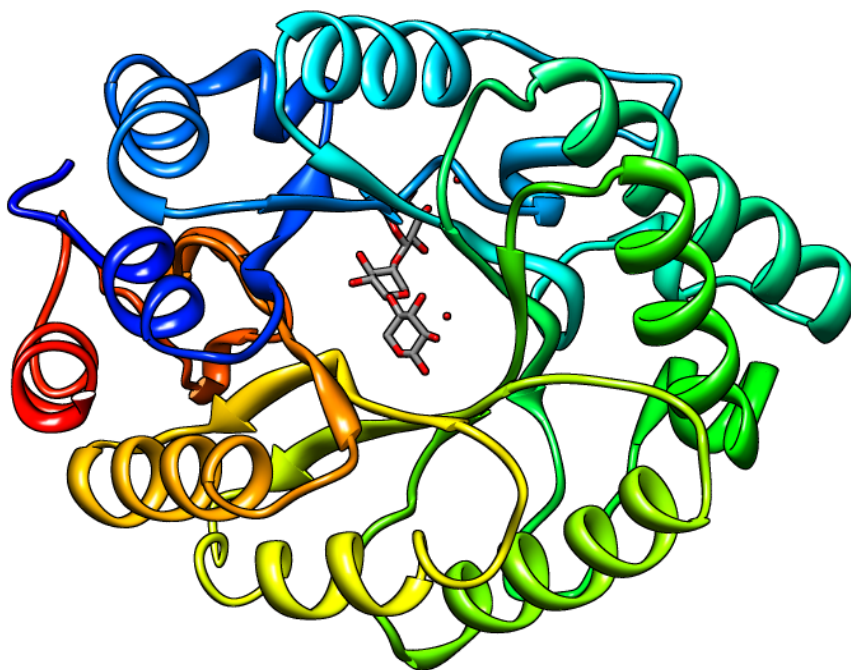
It is now known that ribavirin is a prodrug, converted to the monophosphate (RMP) by adenosine kinase.<sup>24, 25</sup> Phosphorylation then continues beyond the monophosphate to the di- and tri-phosphate nucleotides, with the tri-phosphate being the major metabolite.<sup>25</sup> Other metabolism possible in cells includes dephosphorylation back to ribavirin, occurring mainly in nucleated cells. This is partially responsible for the side effect of hemolytic anemia in patients, as erythrocytes, lacking a nucleus and the necessary enzymes, accumulate RTP. Additional catabolism occurs in the slow deamination of ribavirin by adenosine deaminase, producing inactive triazole carboxylate.<sup>26</sup>

As previously noted, RMP is an inhibitor of IMPDH, with a  $K_i$  of 270 nM for *E. coli* IMPDH<sup>23</sup>, and a  $K_i$  of 650 nM and 390 nM for human IMPDH types I and II.<sup>27</sup> IMPDH inhibition alone, however, does not fully account for the antiviral activity of ribavirin. Complementation of inhibition by XMP or GMP only partly abolishes the antiviral activity, indicating an additional mechanism is involved.<sup>28</sup> Phosphorylated metabolites of ribavirin are also inhibitors of viral RNA polymerases. Specifically, RDP inhibits HIV reverse transcriptase *in vitro*, and RTP has been shown to directly inhibit RNA polymerase of hepatitis C and influenza virus. Most interestingly, RTP is also a substrate for RNA polymerases, and is believed to tip mutation rates of viral genomes into an "error catastrophe." Consequently, RTP was incorporated into a growing RNA strand by poliovirus polymerase, and shown to base pair equally with uridine and cytidine upon rotation of the carboxamide moiety.<sup>29</sup> Furthermore, the resulting increase in mutation rate, from 2 mutations per genome to 7, was shown to decrease infectivity by 95%, as determined by number of plaque forming units per nanogram of viral RNA.<sup>30</sup>

## Inosine Monophosphate Dehydrogenase (IMPDH)

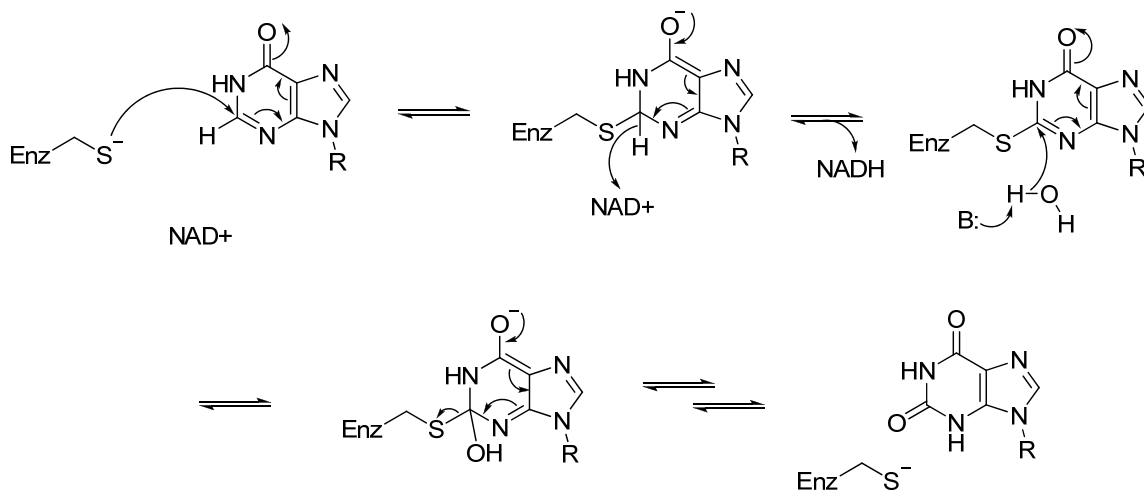
IMPDH (E.C. 1.1.1.205) encoded in *guaB* of the *gua* operon, catalyzes the first committed step in guanine nucleotide biosynthesis. The enzyme catalyzes the oxidation of IMP to XMP with concomitant reduction of  $\text{NAD}^+$  to NADH. XMP is then aminated to GMP by the enzyme GMP synthase, encoded by *guaA* of the same operon.

In *E. coli*, mutations in *guaB*, rendering IMPDH inactive, produces a lethal effect, restored by complementation with xanthosine or guanosine.<sup>31</sup> These knockout species can be distinguished from healthy *E. coli* by choice of media. Rich media allows both types of *E. coli* to grow – the knockouts surviving only via purine salvage. In defined, minimal media, only species with working copies of IMPDH are viable. This phenomenon has been exploited for screening new synthetic compounds for IMPDH inhibition.<sup>32</sup> Similarly, addition of recombinant IMPDH from different species allows for screening of inhibitors of human or other IMPDH.<sup>33</sup>



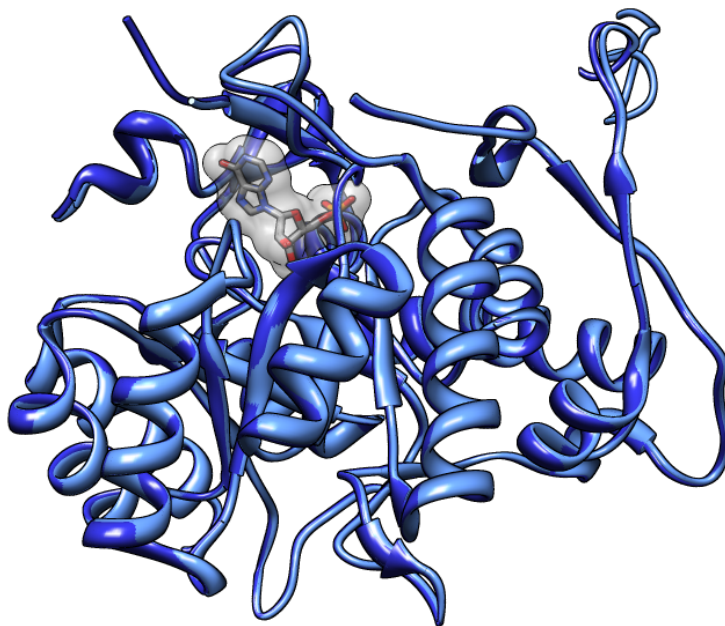
**Figure II-4: Ternary complex of *H. sapiens* IMPDH type II, as adapted from original crystal structures obtained by Colby, T. D., *et al*,<sup>34</sup> using UCSF Chimera software.<sup>8</sup>**

Crystal structures of IMPDH from seven different species have been published, and the mechanism for catalysis has been studied extensively.<sup>35-40</sup> Humans have two isoforms of the enzyme, type I and type II, with 36% sequence identity to that of *E. coli*.<sup>41</sup> The active form of all known IMPDH enzymes is a homotetramer comprised of 55 KDa subunits, with each monomer consisting of an eight-stranded  $\alpha/\beta$  barrel.<sup>35</sup> (Figure II-4) A poorly conserved flap forms part of the active site together with an active site loop, containing a catalytic cysteine residue. The proposed mechanism for catalysis (Scheme II-1) describes the two substrates, IMP and  $\text{NAD}^+$ , binding in random order, followed by the active site Cys attacking the C2 position of IMP. A hydride transfer to  $\text{NAD}^+$  occurs with the formation of a covalent E-XMP intermediate and simultaneous departure of NADH. Water then attacks the same C2 position, expelling the Cys residue and XMP product.<sup>37</sup> The rate of catalysis is  $13 \text{ s}^{-1}$  with a  $K_m$  of 61  $\mu\text{M}$  IMP and 2 mM for  $\text{NAD}^+$ , resulting in a  $k_{\text{cat}}/K_m K_m (\text{sM}^2)^{-1} \times 10^7$ .<sup>40</sup>



**Scheme II-1: Proposed mechanism of IMPDH catalyzed conversion of IMP to XMP, adapted from Gan, L. *et al.*<sup>37</sup>**

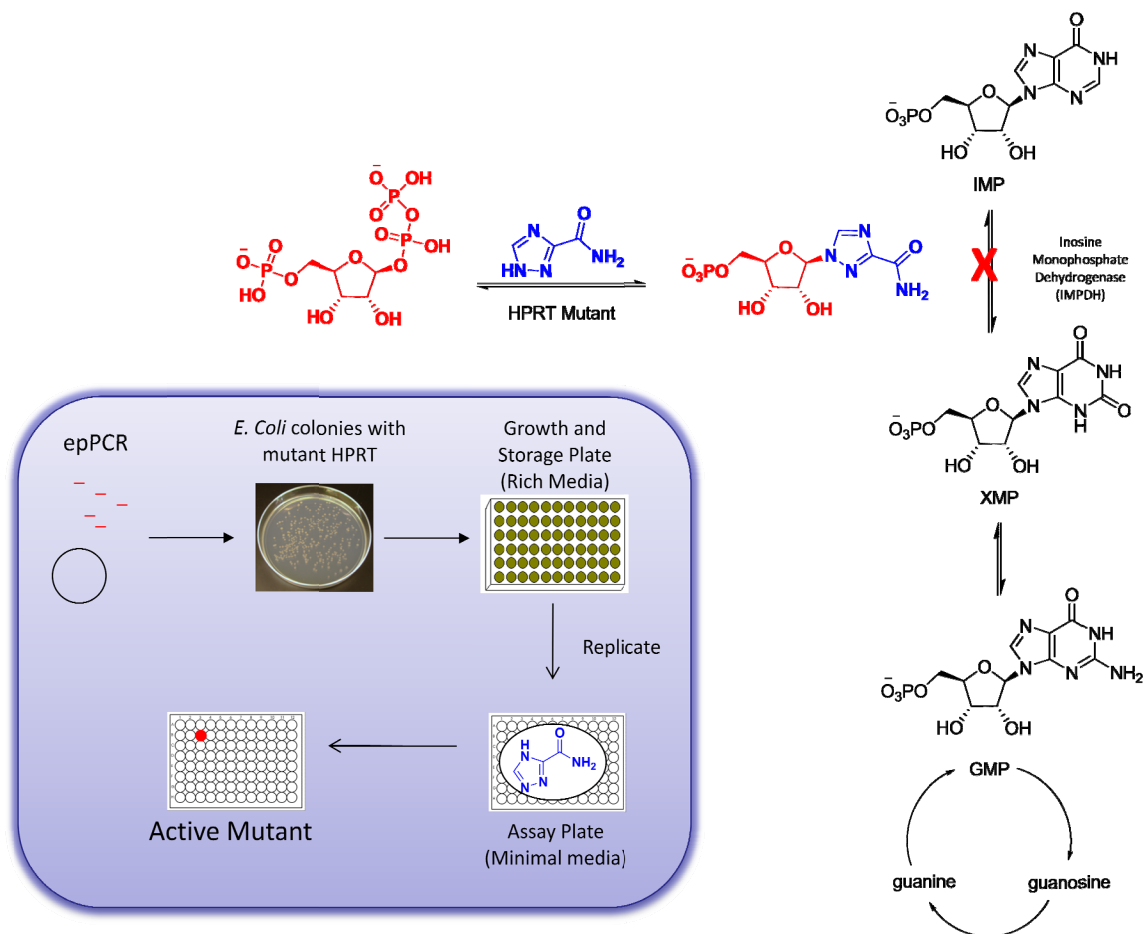
IMPDH is naturally inhibited by GMP ( $K_i$  56  $\mu\text{M}$ ) via feedback inhibition, and numerous synthetic compounds have been created to inhibit the enzyme. Phosphorylated nucleotide analogues, such as ribavirin, competitively inhibit the enzyme via binding to the IMP binding site, while inhibitors such as tiazofurin are active as dinucleotides, inhibiting the  $\text{NAD}^+$  binding pocket.<sup>41</sup> Ribavirin inhibits IMPDH of *E. coli* with a  $K_i$  of 270  $\text{nM}$ <sup>23</sup>, and a  $K_i$  of 650  $\text{nM}$  and 390  $\text{nM}$  for human IMPDH types I and II,<sup>27</sup> and 6  $\mu\text{M}$  for *S. pyogenes*.<sup>42</sup> A crystal structure of *Tritrichomonas foetus* IMPDH in complex with RMP shows the inhibitor occupying the active site in the same orientation as IMP or XMP. (Figure II-5) The phosphate is coordinated with serine and tyrosine residues 317 and 405, and main chain nitrogens. Hydrogen bonding occurs between the RMP amide oxygen and nitrogen backbones at residues 408 and 409.



**Figure II-5: Overlap crystal structures of IMPDH of *T. foetus* bound with IMP and with RMP, adapted from original crystal structures by Prorise, G. L., *et al.*,<sup>35, 43</sup> using UCSF Chimera software.<sup>8</sup> Ligand positions are indiscernable due to identical positioning.**

### **Synopsis of Proposed Screening Method for HPRT Evolution**

The purine salvage enzyme HPRT catalyzes the addition of nucleobase hypoxanthine to the phosphorylated ribose PRPP. Other purine salvage enzymes have shown some leeway in specificity for the nucleobase moiety, providing non-natural nucleotides. Additionally, substrate specificity in HPRT has historically been significantly altered by single amino acid substitutions. Together, these properties suggest HPRT would be a good candidate for directed evolution for the production of the nucleotide analogue ribavirin monophosphate (RMP). RMP is structurally similar to IMP, AMP, and GMP, and is a known inhibitor of IMPDH, with a  $K_i$  of 270 nM for *E. coli* IMPDH. Consequently, *E. coli* harboring an RMP synthesizing mutant HPRT will be distinguishable by media, since IMPDH catalyzes the first committed step in the *de novo* pathway for guanine nucleotide biosynthesis. This provides the basis for the *in vivo* screening method (Scheme II-2) in the directed evolution of HPRT for RMP biosynthesis.



**Scheme II-2:** *In vivo* screening method for RMP production. TCA added to growth media passively transports into the cell where it condenses with endogenous PRPP via HPRT catalysis. Resulting RMP inhibits IMPDH effectively inhibiting GMP production and cell growth. Cells grown with guanine or guanosine supplementation can survive the inhibition. Inset depicts selection process of isolation of host cells, each carrying a unique variant, followed by growth in minimal media with triazole prodrug. Active mutants are identified as wells with little to no growth.

## Results

To begin the evolution of the hypoxanthine phosphoribosyl transferase into a triazole phosphoribosyl transferase, the *hpt* gene was cloned from *E. coli* via PCR, into commercial plasmid pET28a (Novagen). Substrate 1,2,4-triazole-3-carboxamide (TCA) was synthesized from the methyl ester and composition was confirmed by NMR, MS, and elemental analysis. The pET28a plasmid bearing the wild type HPRT gene was transformed into *E. coli* BL21(DE3) and the



toxicity of the synthesized triazole substrate was evaluated to guide the starting concentration of TCA in subsequent screening studies. The growth rate of uninduced cultures of this strain, in minimal growth medium M9, was determined in the presence of increasing concentrations of TCA, indicating an  $IC_{50}$  of 115  $\mu$ M after 6 hours. (Figure II-6) Guanosine, which is also reported to complement genetic knock-outs of IMPDH in *E. coli*, was capable of near complete restoration of growth in triazole induced inhibition in this strain, suggesting that the toxicity of TCA results from its transformation into IMPDH targeting RMP. (Figure II-7)

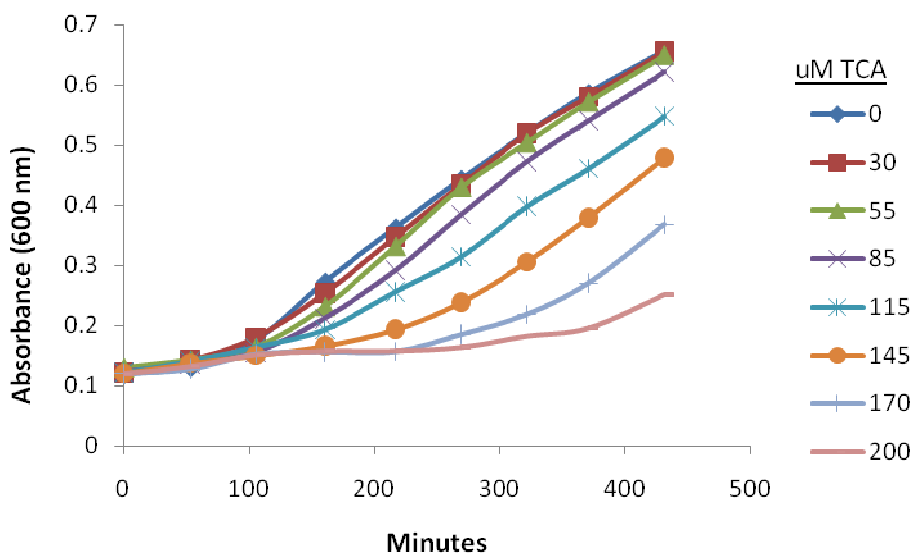
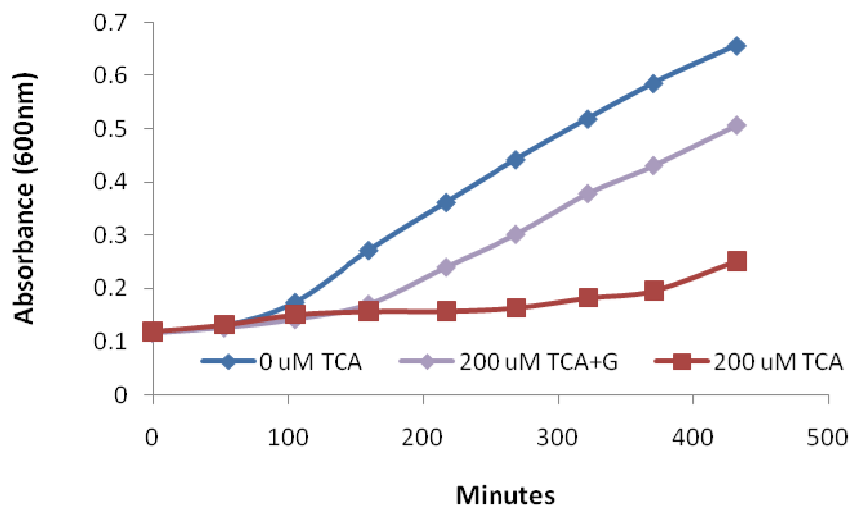


Figure II-6: TCA toxicity to *E. coli* containing wt HPRT in pET28a plasmid.



**Figure II-7: Guanosine complementation in *E. coli* with wt HPRT.**

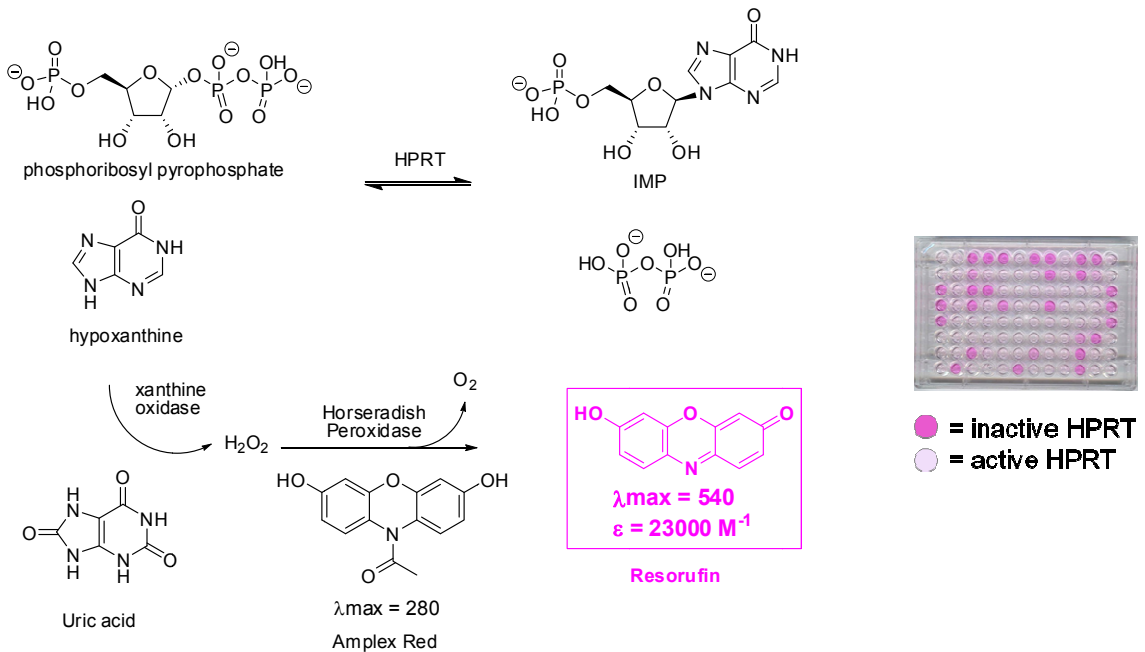
Diversity was introduced to the HPRT gene pool by mutagenic PCR to create a library of HPRT variants. Primers were designed 14 bp from flanking cloning sites in order to maximize efficiency of restriction for subsequent ligation of library product into plasmid vector. Control over mutation rate was managed by addition of  $MnCl_2$ . (Table II-2) Mutagenic PCR product was restricted using NdeI and HindIII and ligated into correspondingly restricted pET28a plasmid vector for subsequent transformation by electroporation into *E. coli* BL21(DE3).

**Table II-2: Correlation of mutation rate with manganese chloride concentration.**

$MnCl_2$ (mM)		0.5	0.3	0.2	0.05
Avg. mutations/gene		8.3	4	2	0

Error rates were determined by sequence comparison of a small sample of randomly selected library members with wild type HPRT sequence. While this provides an exact number of mutations per gene, pecuniary limitations restrict the sample size, consequently decreasing the

statistical accuracy of the mutation estimate. If the assumption can be made that most mutations to the wild type enzyme will have no effect or detrimental effects on catalysis of the natural substrate, a measurement of enzyme activity can be used to correlate to mutation rate of a larger sample size. In this way, hypoxanthine consumption by a microtiter plate of randomly selected clones was used to estimate the proper mutation rate of each epPCR library. Inactive mutants were unable to consume all available substrate in the allotted time. The remaining hypoxanthine was then converted to xanthine and ultimately uric acid, by xanthine oxidase, with concurrent production of peroxide. Together, with horseradish peroxidase, the resultant peroxide reacted with an added dye (Amplex Red) to create a highly visible response. (Figure II-8)

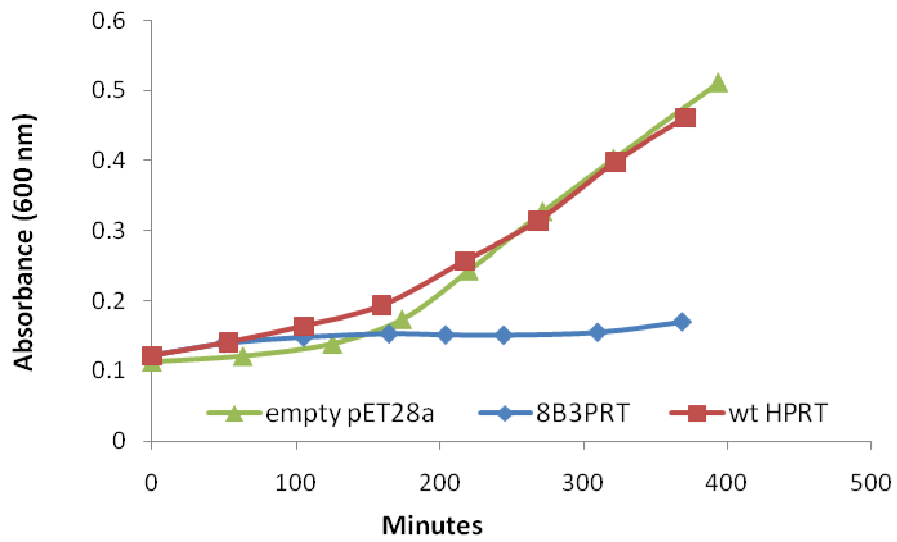


**Figure II-8: Hypoxanthine consumption assay for percent activity of a library.**

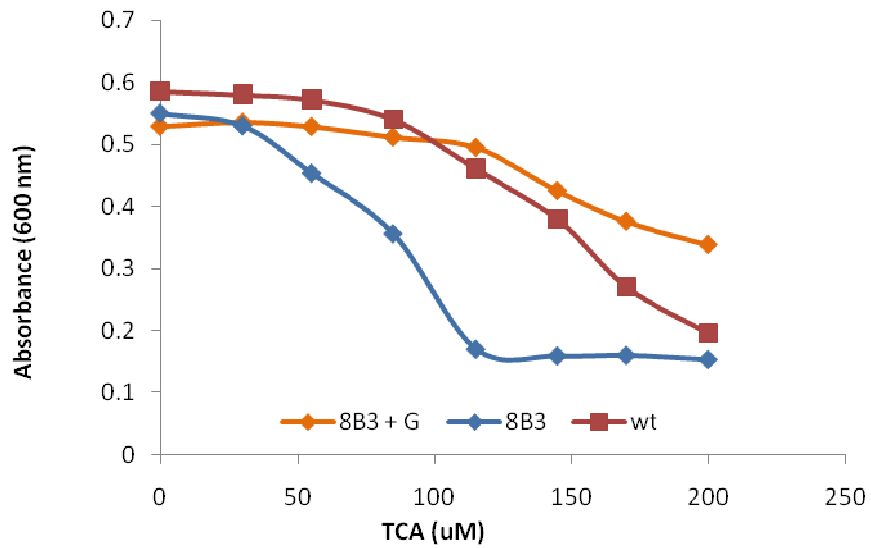
With the mutation rate optimized at a level of 30 to 40% inactivity, or approximately 1-2 mutations per gene, (0.6 kb HPRT coding region) 1000 clones were picked into 96-well plates

containing LB medium for initial screening. Colonies were replicated into minimal medium containing 200  $\mu$ M TCA, covered with parafilm, and incubated with shaking at 37°C for six days to observe growth rates. Visual inspection of wells revealed clones that were apparently more sensitive to triazole, and wells with no growth at six days were selected as candidates for secondary screening. Ten positive clones were rescreened against a dilution series of TCA, and all selected clones demonstrated a consistently increased sensitivity to the triazole protoxin. Plasmids were isolated from the top two growth attenuated strains, strains 8B3 and 10A1, and freshly transformed into *E. coli* BL21(DE3) to rule out auxotrophic effects.

Figure II-9 compares the growth of *E. coli* containing empty pET28a vector, pET28a-wtHPRT and pET28-8B3PRT with a fixed concentration of triazole (115  $\mu$ M). Consistent with earlier observations, there is a marked difference in growth phenotype between the 8B3PRT containing strain and strains containing wild-type HPRT or empty vector. Figure II-10 illustrates dose response curves of these stains demonstrating a strikingly decreased IC<sub>50</sub> in 8B3PRT in comparison to wild type HPRT containing strains. Moreover, supplementation of the 8B3PRT strain with 1 mM guanosine substantially complemented the 8B3PRT containing strains, eliminating triazole induced cytotoxicity and implicating IMPDH as the target of endogenously produced RMP. The growth phenotypes of wild type HPRT and 8B3 transferase (8B3PRT) support the hypothesis that 8B3 accumulates intracellular pools of ribavirin monophosphate and that these pools attenuate growth in triazole containing medium.



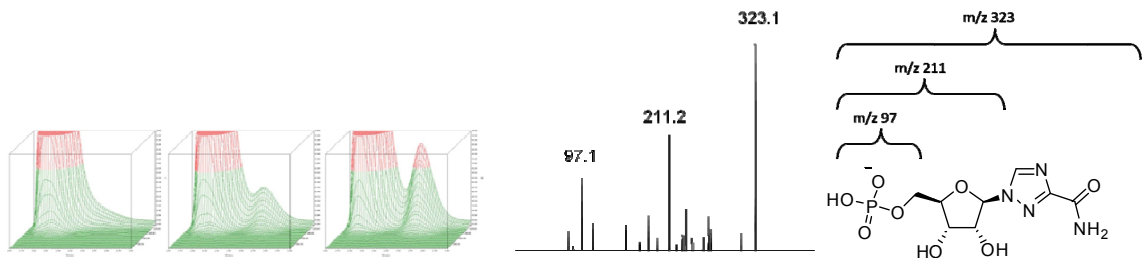
**Figure II-9: Growth rate comparison of 8B3PRT mutant and wt HPRT in 115  $\mu$ M TCA.**



**Figure II-10: Complementation effect of guanosine on 8B3 mutant in comparison with wt HPRT and 8B3 without guanosine.**

To verify the presence of RMP, an HPLC method for comparison of TCA turnover was developed using triethylammonium ion as a pairing agent, and a chemically synthesized RMP standard. Lysate of induced cell cultures, when combined with substrates, produced an RMP

peak with identical retention time and identical fragmentation of the preparative RMP fraction as determined by tandem mass spectrometric analysis. (Figure II-11) Quantitative comparison of the biosynthesized RMP in induced cultures, normalized by  $OD_{600}$ , revealed a five fold increase in RMP production by the 8B3 mutant versus wild type HPRT.



**Figure II-11: HPLC and MS of biosynthesized RMP: Images on left illustrate the appearance of the RMP peak over time. Mass spectrum of the preparatory fraction of the RMP peak is to the right.**

Protein quantification of nonpurified cell lysate was possible by fluorescence imaging after exposing protein gels to InVision stain, specific to histidine tagged proteins. This analysis showed expression levels of 8B3PRT enzyme was 2.8 times higher than wt HPRT cultures of equal density.

Subsequently, mutant transferases were overexpressed by induction with IPTG, purified by  $Ni^{2+}$  affinity chromatography, and desalted by size exclusion gel chromatography. Kinetic parameters were determined by measurement of pyrophosphate release upon nucleobase binding. Results, shown in Figure II-12 and Table II-3, indicate a modest 1.8-fold increase in catalytic turnover, suggesting that the majority of improvement in cellular 8B3PRT turnover of triazole is due to enhanced protein expression levels.

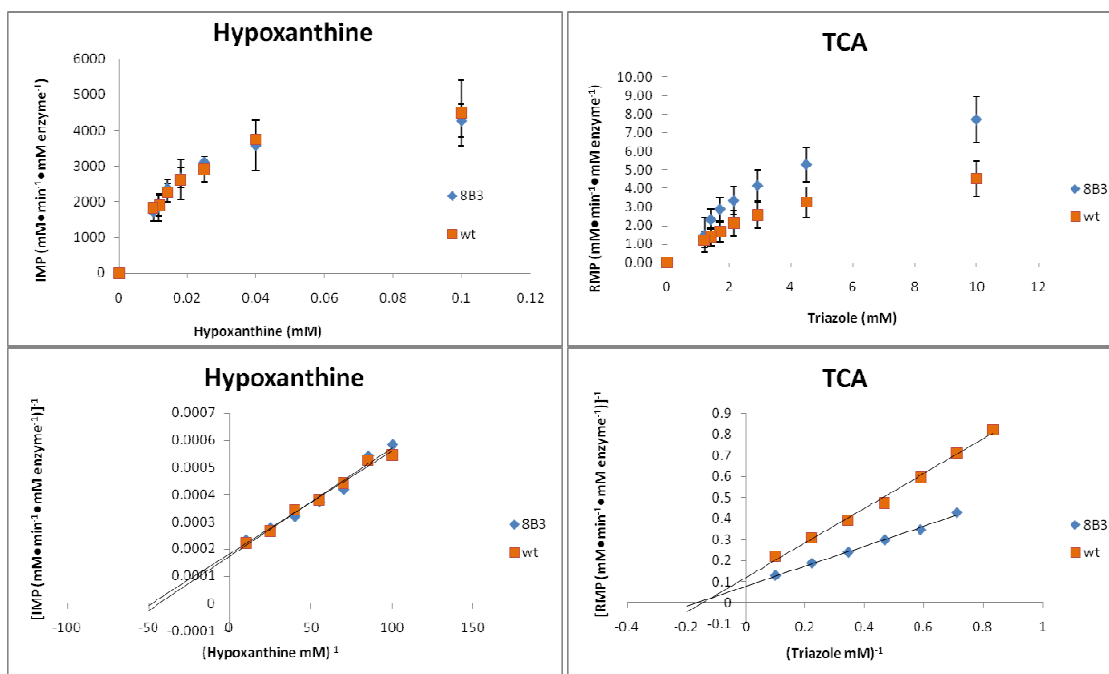


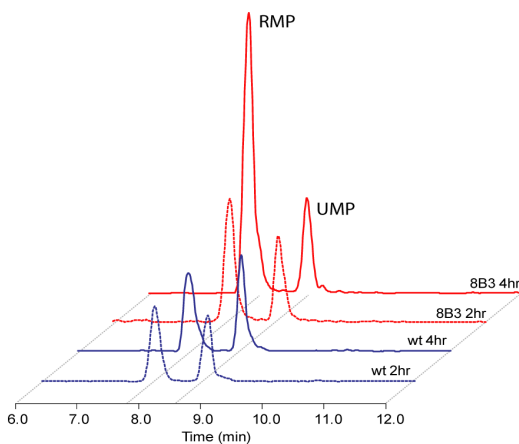
Figure II-12: Kinetic measurements of hypoxanthine and TCA turnover by wt HPRT and 8B3PRT

Table II-3: Kinetic parameters of hypoxanthine and TCA turnover by wt HPRT and 8B3PRT

	Hx			TCA		
	$k_{cat}$ ( $s^{-1}$ )	$K_m$	$k_{cat}/K_m$ ( $mM^{-1}s^{-1}$ )	$k_{cat}$ ( $s^{-1}$ )	$K_m$	$k_{cat}/K_m$ ( $mM^{-1}s^{-1}$ )
<b>WT</b>	5.46E+03	2.04E-02	2.68E+05	6.92	5.16	1.34
<b>error (+/-)</b>	141	1.27E-03		0.217	0.299	
<b>8B3</b>	5.12E+03	1.78E-02	2.87E+05	12.6	6.19	2.03
<b>error (+/-)</b>	194	1.71E-03		0.950	0.825	
<b>Fold</b>	<b>0.9</b>	<b>0.9</b>	<b>1.1</b>	<b>1.8</b>	<b>1.2</b>	<b>1.5</b>

To confirm the hypothesis that 8B3 mediated intracellular RMP accumulation is linked to growth inhibition, intracellular accumulation of RMP in selected strains was quantified using HPLC/ESIMS. Cultures of 8B3PRT and wild type HPRT were inoculated into minimal media containing 150  $\mu$ M triazole and incubated for two and four hours prior to lysis by detergent and protein precipitation with chloroform:methanol (1:2). Nucleotides extracted with water were

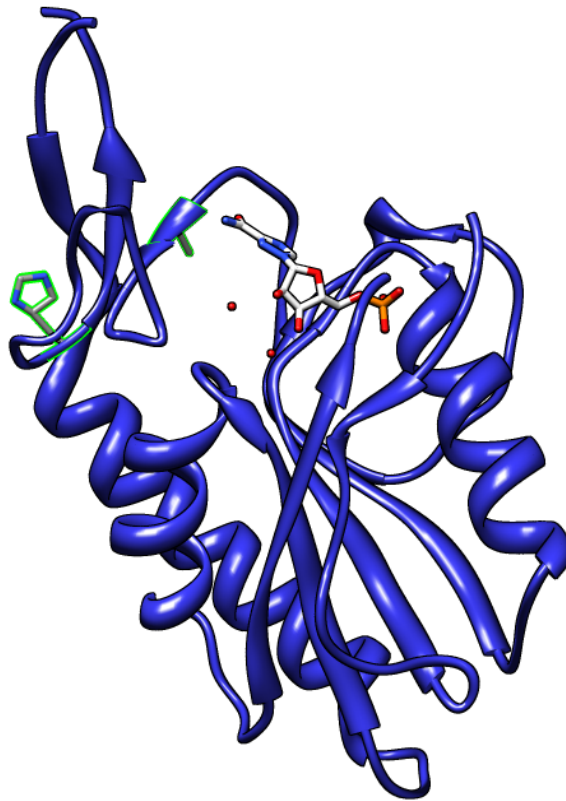
separated using porous graphitic carbon (Hypercarb) and analyzed by ESIMS. Figure II-13 demonstrates that the decreased growth phenotype strain 8B3PRT is linked to elevated intracellular accumulation of RMP. Remarkably, RMP concentrations were over twice as concentrated in 8B3PRT strains as HPRT strains.



**Figure II-13: Intracellular comparison of RMP in wt and 8B3. The second compound to elute is the isobaric UMP naturally present in the cell.**

Mutational analysis revealed 8B3 to contain three silent and two coding mutations - V157A, in the active site nucleobase binding region, and Y170H, distal to the active site. The valine residue, while not oriented toward the active site, is an important residue in the  $\beta$ -sheet of the hood domain. As previously noted, the hood domain regulates recognition and binding of the purine base. It is possible that the V157A mutation alters the native configuration of this domain, consequently changing the nucleobase binding. Additionally, either of the mutations could possibly alter the intrinsic dynamics of the enzyme, ultimately effecting substrate binding and rate of catalysis.





**Figure II-14: Theoretical model of 8B3PRT with bound RMP, generated with UCSF Chimera software,<sup>8</sup> adapted from crystal structure of wild type HPRT with bound IMP, as determined by Guddat, L. W. *et al.*<sup>9</sup> Mutated residues V157A and Y170H are highlighted.**

In addition to the directed evolution experiment using random mutations, several site directed point mutations were individually created in wt HPRT to modify substrate specificity. These point mutations included the original 8B3 mutations V157A and Y170H, and several others (F156N, I162Q, K135Q, V158K, and V158Q). The individually created 8B3 mutations were constructed separately to determine the interaction of these mutations, if any, in 8B3. Based on growth sensitivity to TCA and TCA turnover by cell lysate, both mutations are required to match the RMP formation ability of 8B3. Since the amide moiety of TCA is free to rotate, and the corresponding portion of hypoxanthine is rigid, mutations I162Q and K135Q were created in an effort to accommodate the rotated, adenine analogue rotomer of TCA while simultaneously completely blocking hypoxanthine. Mutations V158K, V158Q, and F156N lie close to the active

site and flank the successful V157A mutation site. Based on the crystal structure of wild type HPRT, it is possible that these mutations could have favorable TCA binding. Interestingly, none of the point mutations matched 8B3 in TCA turnover (data not shown).

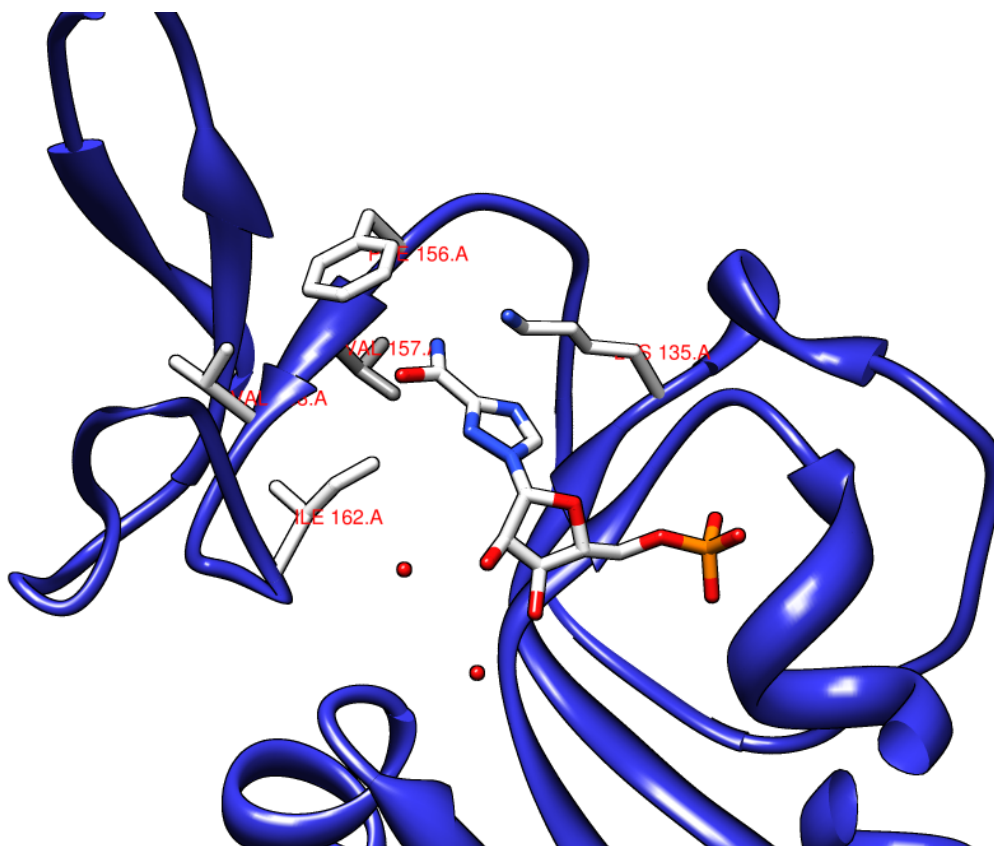


Figure II-15: Active site point mutation locations shown with native amino acid residues and theoretical bound RMP with rotated amide bond, created with UCSF Chimera software.<sup>8</sup> Adapted from original crystal structure of HPRT from *E. coli*, as determined by Guddat, L. W. *et al.*<sup>9</sup>

## Conclusions

Addition of  $MnCl_2$  to the PCR proved a reliable method of introducing mutations to the HPRT gene. The best selected mutant, 8B3, produced around 4.5 fold more RMP than wt HPRT per cell, as determined by LCMS. This improvement was primarily due to an increased expression level, and partially due to improved catalysis. Furthermore, intracellular accumulation of RMP was detected in both types of *E. coli* cultures; mutant and wt HPRT. The

mutant showed twice as much intracellular accumulation of RMP after four hours. Together, this data provides excellent support for the hypothesis that HPRT can produce RMP from TCA, and that the IMPDH inhibiting nucleotide product will be detectable by media selection. Although the improvement in turnover for the triazole substrate is moderate, the ability of the *in vivo* screening method to consistently identify 8B3 from wt serves as a proof of principle for this method of *in vivo* screening, or reverse selection for intracellular target product.

## Experimental Methods

**Preparation of Electrocompetent Cells.** Stock electrocompetent cells were created by inoculating 500 mL of L-broth (LB broth with 50% of the amount of NaCl) in a 2.8 L Fernbach flask with a 5 mL overnight culture of BL21(DE3) *E. coli* cells. These cells were grown at 37°C with shaking, to an OD<sub>600</sub> of 0.6. Cells were then chilled on ice for 20 minutes, and collected by centrifugation (13,000 rpm) at 4°C. Supernatant was discarded, and cells resuspended in ice cold 10% glycerol. Finally, these cells were centrifuged again (13,000 rpm, 4°C) and supernatant decanted. This process provided around forty 50-μL aliquots which were stored at -80°C for later use.

**Electroporation.** Electroporation was performed using Bio-Rad's Gene Pulser Xcell™ Electroporation System with Pulse Control Module. Microcentrifuge tubes containing 40 μL of electrocompetent BL21DE3 cell stock, thawed on ice, were added to 2 μL of resuspended DNA, and incubated on ice for 1 minute. The cell mixture was transferred to a previously chilled electroporation cuvette, and placed in the shock chamber. Cells received brief pulse of 2.5 kV, and were immediately resuspended in SOC media. After an incubation time of one hour at 37°C, and shaking at 225 rpm, cells were spread on Petri-dishes containing LB agar with Kanamycin (50 μg/mL).

**Library generation.** The mutation rate was titrated to adjust the HPRT inactivity of the library to approximately 40%. Biochemical assay of HPRT activity was determined by a hypoxanthine consumption assay as follows. Random colonies were picked into LB in 200  $\mu$ L microtiter plates and incubated overnight. Cells were diluted 1:4 into with LB medium and induced with 1 mM IPTG. After a 2 hour incubation at 37°C, cells were lysed with a cell disruption reagent (BugBuster, Novagen Inc.) and 10  $\mu$ L of the supernatant was added to 90  $\mu$ L Hypoxanthine Master Mix (100 mM Tris-HCl, 12 mM MgCl<sub>2</sub>, 1 mM phosphoribosyl pyrophosphate, and 100  $\mu$ M hypoxanthine). Reactions were allowed to proceed for thirty minutes at 37 °C before 50  $\mu$ L Amplex Red solution (10 mM Amplex Red reagent (Molecular Probes), 2 U horseradish peroxidase, 0.2 U xanthine oxidase) was added. After 30 minutes of incubation at 37°C, activity was assayed by quantifying the remaining hypoxanthine by the coupled oxidation of Amplex Red monitored at 540 nm.

Competent cells were transformed with DNA from PCR conditions with the desired mutation rate resulting in approximately 40% inactive clones. Final mutagenic PCR conditions were performed on a 50  $\mu$ L scale containing 3.75 ng pET28HPRT plasmid template, 125 ng each primer (Fwd. CATCACAGCAGCGGCCTGGTG, Rev. TGGTGCTCGAGTGCGGCCGC), 0.2 mM each dNTP, 5.0 mM MgCl<sub>2</sub>, 0.3 mM MnCl<sub>2</sub>, 5  $\mu$ L 10X PCR Buffer (Roche), 0.5  $\mu$ L DMSO, and 5 U Taq Polymerase (Roche), using 30 thermal cycles (1 min. 94°C, 1 min. 50.5°C, 1 min. 72 °C). Expression vector pET28a and PCR insert were purified (Qiagen Miniprep and PRC purification kits), and digested with NdeI and HindIII. Digested DNA was gel purified and cloned into pET28a using Quick Ligase (NEB), and the new construct was desalted by precipitation prior to transformation into competent *E. coli* BL21(DE3) via electroporation. Seven random colonies were chosen for validation of digestion, confirmation of efficiency of ligation and sequencing after plasmids were purified with Qiagen Miniprep kit.

**Synthesis of 1,2,4-triazole-3-carboxamide.** Methyl-1H-1,2,4-triazole-3-carboxylate (408 mg, 3.21 mmoles) (Aldrich) was added to a 100-mL round bottom flask, with septum, and flushed with argon. Twenty-five mL 7N ammonia in methanol was then added through the septum with a glass syringe, and the solution stirred at room temperature for 2 days. The product (1,2,4-triazole-3-carboxamide) appeared as a white precipitate, collected with a fritted filter funnel, and washed with methanol. Finally, the product was dried under high vacuum, resulting in a mass of 200 mg (56% yield). <sup>1</sup>H-NMR (DMSO): 14.49 (s, 1H), 8.39 (s, 1H), 7.93 (s, 1H), 7.68 (s, 1H). Addition of D<sub>2</sub>O resulted in the subsiding of peaks at 7.93 and 7.68, indicative of exchangeable amide hydrogens. <sup>13</sup>C-NMR (DMSO): 161. Elemental analysis calculated for C<sub>3</sub>H<sub>4</sub>N<sub>4</sub>O: C, 32.15; H, 3.60; N, 49.98; O, 14.27. Analysis found: C, 32.03; H, 3.63; N, 48.97; O, 14.82. Exact mass of C<sub>3</sub>H<sub>4</sub>N<sub>4</sub>O<sup>+</sup> = 112.0380 amu. Observed mass = 112.0378 amu.

**Selection method.** Transformants were transferred into 96-well master plates with 100 μL/well LB medium and incubated overnight. Glycerol was added to a final concentration of 25% and then immediately sealed prior to storage at -80°C. Master plates were replicated into 96 well plates containing 200 μL/well M9 minimal media and 200 μM 1,2,4-triazole-3-carboxamide. Screening plates were sealed with a Parafilm layer (covering the wells) under the lid and incubated at 37°C, with shaking at 275 rpm. Wells were visually inspected each day, and those without growth after 6 days were selected from master plates for further characterization.

**Cell growth assay.** LB cultures (5 mL) were inoculated with glycerol stock cultures of *E. coli* containing wt or mutant HPRT and grown overnight at 37°C. Cells were then centrifuged for two minutes at 13,000 rpm (15,682 rcf) and cell pellets resuspended in minimal media. Resuspended cells (500 μL) were added to 5 mL cultures of minimal media supplemented with various concentrations of 1,2,4-triazole-3-carboxamide (0 to 200 μM). Absorbances at 600 nm were measured every 45 to 60 minutes.

**Expression and purification of wt and mutant HPRT.** Wild type and mutant HPRT were grown in 0.5 L cultures in LB medium to an OD of ca. 0.6 and induced with 1 mM IPTG for 3 hours before centrifugation. Frozen cell pellets were resuspended in 25 mL Buffer A (500 mM NaCl, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 30 mM imidazole, pH 7.4) prior to lysis by double passage through a French Pressure Cell, 20000 psi. Cell lysates were clarified by centrifugation (4°C, 13,000 rpm, 20 min) and applied to previously stripped (500 mM NaCl, 20 mM Na<sub>2</sub>HPO<sub>4</sub>, 50 mM EDTA, pH 7.4) and freshly recharged Ni<sup>2+</sup>-affinity columns (HisTrap FF crude, 5 mL, GE Healthcare, stripped and recharged with 100 mM NiSO<sub>4</sub> according to manufacturers instructions). Proteins were eluted with a linear gradient of 100:0 A:B to 50:50 A:B (Buffer A = 500 mM NaCl, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 20 mM imidazole, pH 7.4, Buffer B = 500 mM NaCl, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 500 mM imidazole, pH 7.4) at 5 mL/min and concentrated by centrifugal 10K MWCO filtration (Millipore). Proteins were exchanged into Desalting Buffer (50 mM Tris-HCl, pH = 7.4, 500 mM NaCl, 10 mM MgCl<sub>2</sub>) via gel filtration (HiTrap Desalting, 5 mL, GE Healthcare). Concentrated enzymes were stored in 100 µL aliquots at -80°C. Proteins were analyzed by 12% polyacrylamide gel, loaded with 7 µL marker (Precision Plus Protein Kaleidoscope Standard, BioRad Inc), and 5 µL of each enzyme and stained using Coomassie blue or InVision stain.

**Protein Quantification with InVision His-Tag In-gel Stain.** Cells were prepared as described in "Expression and Purification" section above. PRT extract (10 µL) was added to Laemmli sample buffer (20 µL) with 5% β-mercaptoethanol and denatured for 5 minutes at 95°C. After cooling on ice for 5 minutes, protein solutions (5 µL) were added to wells of a 12% polyacrylamide gel and separated at 200 V for 40 min. Hexahistidine tagged proteins were visualized with InVision His-Tag In-gel Stain (Invitrogen, Carlsbad, CA) according to manufacturers protocol and quantified using a 532/1064nm laser with 555nm long pass filter (Molecular Imager FX System, Bio-Rad, Hercules, CA). The linear range was established by

measurement of a series of several enzyme dilutions, and relative protein concentration was determined by comparison of gel band intensities within this linear range.

**Intracellular RMP analysis.** Wild type and 8B3 mutant HPRT clones were grown in conditions identical to the Cell Growth Assay with the amount of 1,2,4-triazole-3-carboxamide set at 150  $\mu\text{M}$ . Three separate 8B3 cultures and three WT cultures were inoculated from randomly selected, newly transformed colonies on LB agar and grown overnight in LB broth. Each individual colony culture was individually centrifuged and resuspended in minimal media, and 500  $\mu\text{L}$  used to inoculate two separate 5 mL minimal media cultures. A total of 12 minimal media cultures were grown: 6 wt and 6 8B3. Half were collected after two hours (3 wt and 3 8B3) and the remaining were collected after 4 hours. Upon collection, the cultures were centrifuged, resuspended in 1 mL minimal media, and centrifuged again. Pellets were lysed with 200  $\mu\text{L}$  detergent (Bugbuster) and protein precipitated with 500  $\mu\text{L}$  of a chloroform:methanol (1:2) solution, followed by vortexing and incubation on ice for 5 minutes. Finally, 200  $\mu\text{L}$  water were added, vortexed, centrifuged, and 600  $\mu\text{L}$  of the aqueous portion was removed and stored in 50  $\mu\text{L}$  aliquots at  $-80^{\circ}\text{C}$ .

**HPLC/ESI-MS analysis of intracellular nucleotides.** Seven micro liters of an internal standard of 2'-deoxy inosine monophosphate (0.9  $\mu\text{M}$  final concentration) was added to the 50  $\mu\text{L}$  aliquot prior to mass analysis. Hypercarb porous graphitic carbon was used to separate RMP from other nucleotides using a 10 minute linear gradient from 95:5 20 mM ammonium acetate:acetonitrile to 80:20 using 20  $\mu\text{L}$  injections. RMP levels were compared by dividing the RMP peak area by the dIMP internal standard peak area, and dividing once more by the  $\text{OD}_{600}$  of the culture at time of collection to eliminate differences engendered by growth rates.

**HPLC of cell lysate reaction.** Cells were prepared as described in "Expression and Purification" section above. PRT extract (10  $\mu\text{L}$ ) was added to an aqueous solution containing a

final concentration of 1 mM PRPP, 1 mM TCA, 12 mM TRIS pH 7.3 and 12 mM MgCl<sub>2</sub> for a total reaction volume of 0.5 mL. Reactions were incubated at 37°C and analyzed every 5 to 7 minutes by HPLC. Each HPLC injection was 10 µL in volume, and nucleotides separated with an isocratic mobile phase of 100 mM triethyl ammonium phosphate, pH 7.0, with 3% methanol, flow rate of 1 mL per minute. Column used was XTerra MS C<sub>18</sub>, 5 µm 4.6 x 150 mm with 4.6 x 10 mm XTerra MS C<sub>18</sub> guard column. Separation is complete in 5 minutes.

**Synthesis of 1-β-D-Ribofuranosyl-1,2,4-triazole-3-carboxamide (ribavirin) mono-phosphate.** Nucleotide product was prepared according to the method described by Allen, *et al.*<sup>44</sup> To a 50 mL round bottom flask, 2.5 mL of trimethylphosphate (Aldrich) was added, and cooled to 5°C with a NaCl ice bath. Previously distilled phosphorous oxychloride (Aldrich) was added (220 µL) and moved reaction to 0°C ice bath. A total of 10.8 µL H<sub>2</sub>O was then added, followed by 146 mg Ribavirin (Chem Pacific) and stirred at 0°C for 5 hours. The pH was brought to 2.5 with 2.9 mL 2N NaOH and allowed to set 16 hours at 23°C. The product was twice purified by a passing through 2.5 mL activated charcoal/water slurry, washed with 2 mL H<sub>2</sub>O, and eluted with 2 mL triethylamine/ethanol/water (3:10:7 v/v). This mixture was filtered through a 0.45 µm syringe filter prior to evaporation by negative pressure. <sup>1</sup>H-NMR (D<sub>2</sub>O): 8.65 (s, 1H), 5.89, (d, 1H), 4.53 (t, 1H), 4.36 (t, 1H), 4.19 (m, 1H), 3.92 (m, 2H). Exact mass of C<sub>8</sub>H<sub>12</sub>N<sub>4</sub>O<sub>8</sub>P<sup>-</sup> = 323.04 amu. Observed mass = 323.1 amu.

**Kinetic measurement.** Assay method was an extension of the method described by Erion, M in 1997.<sup>45</sup> Briefly, freshly prepared Assay Mix (1 mL) containing 30 µL of PRT extract and a final concentration of 100 mM Tris-HCl pH 8.0, 100 mM MgCl<sub>2</sub>, 1 mM PRPP and 8.3 U/mL inorganic pyrophosphatase (from Baker's yeast, Sigma Chemical Inc.). Reactions were initiated by adding Assay Mix (95 µL per well) to triplicate wells containing 5 µL nucleobase in DMSO (final concentrations 10 to 100 µM for hypoxanthine and 1.2 to 10 mM TCA) and allowed to



proceed for five minutes before quenching with 80  $\mu$ L of a solution of 200 mM zinc acetate and 30 mM ammonium molybdate, pH 5. After 30 seconds of incubation, 20  $\mu$ L of 20% ascorbic acid, pH 5 was added, and absorbance at 850 nm was recorded after 17 minutes. Background absorbance values from base analogue, cell lysate, and PRPP were subtracted from the reaction measurement and pyrophosphate turnover was determined by comparison to a standard curve.

## References

1. Stout, J. T.; Caskey, C. T., Hprt - Gene Structure, Expression, and Mutation. *Annual Review of Genetics* **1985**, 19, 127-148.
2. Sinha, S. C.; Smith, J. L., The PRT protein family. *Current Opinion in Structural Biology* **2001**, 11, (6), 733-739.
3. Craig, S. P.; Eakin, A. E., Purine phosphoribosyltransferases. *Journal of Biological Chemistry* **2000**, 275, (27), 20231-20234.
4. Focia, P. J.; Craig, S. P.; Eakin, A. E., Approaching the transition state in the crystal structure of a phosphoribosyltransferase. *Biochemistry* **1998**, 37, (49), 17120-17127.
5. Keough, D. T.; Brereton, I. M.; de Jersey, J.; Guddat, L. W., The crystal structure of free human hypoxanthine-guanine phosphoribosyltransferase reveals extensive conformational plasticity throughout the catalytic cycle. *Journal of Molecular Biology* **2005**, 351, (1), 170-181.
6. Phillips, C. L.; Ullman, B.; Brennan, R. G.; Hill, C. P., Crystal structures of adenine phosphoribosyltransferase from *Leishmania donovani*. *Embo Journal* **1999**, 18, (13), 3533-3545.
7. Vos, S.; Parry, R. J.; Burns, M. R.; de Jersey, J.; Martin, J. L., Structures of free and complexed forms of *Escherichia coli* xanthine-guanine phosphoribosyltransferase. *Journal of Molecular Biology* **1998**, 282, (4), 875-889.
8. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, 25, (13), 1605-1612.
9. Guddat, L. W.; Vos, S.; Martin, J. L.; Keough, D. T.; De Jersey, J., Crystal structures of free, IMP-, and GMP-bound *Escherichia coli* hypoxanthine phosphoribosyltransferase. *Protein Science* **2002**, 11, 1626-1683.

10. Focia, P. J.; Craig, S. P.; Nieves-Alicea, R.; Fletterick, R. J.; Eakin, A. E., A 1.4 angstrom crystal structure for the hypoxanthine phosphoribosyltransferase of *Trypanosoma cruzi*. *Biochemistry* **1998**, 37, (43), 15066-15075.
11. Xu, Y. M.; Eads, J.; Sacchettini, J. C.; Grubmeyer, C., Kinetic mechanism of human hypoxanthine-guanine phosphoribosyltransferase: Rapid phosphoribosyl transfer chemistry. *Biochemistry* **1997**, 36, (12), 3700-3712.
12. Lee, C. C.; Craig, S. P.; Eakin, A. E., A single amino acid substitution in the human and a bacterial hypoxanthine phosphoribosyltransferase modulates specificity for the binding of guanine. *Biochemistry* **1998**, 37, (10), 3491-3498.
13. Munagala, N.; Sarver, A. E.; Wang, C. C., Converting the guanine phosphoribosyltransferase from *Giardia lamblia* to a hypoxanthine-guanine phosphoribosyltransferase. *Journal of Biological Chemistry* **2000**, 275, (47), 37072-37077.
14. Subbayya, I. N. S.; Sukumaran, S.; Shivashankar, K.; Balaram, H., Unusual substrate specificity of a chimeric hypoxanthine-guanine phosphoribosyltransferase containing segments from the *Plasmodium falciparum* and human enzymes. *Biochemical and Biophysical Research Communications* **2000**, 272, (2), 596-602.
15. Witkowski, J. T.; Simon, L. N.; Sidwell, R. W.; Robins, R. K., Design, Synthesis, and Broad-Spectrum Antiviral Activity of 1-Beta-Dribofuranosyl-1,2,4-Triazole-3-Carboxamide and Related Nucleosides. *Journal of Medicinal Chemistry* **1972**, 15, (11), 1150.
16. Utagawa, T.; Morisawa, H.; Yamanaka, S.; Yamazaki, A.; Hirose, Y., Enzymatic-Synthesis of Nucleoside Antibiotics .6. Enzymatic-Synthesis of Virazole by Purine Nucleoside Phosphorylase of *Enterobacter-Aerogenes*. *Agricultural and Biological Chemistry* **1986**, 50, (1), 121-126.
17. Hennen, W. J.; Wong, C. H., A New Method for the Enzymatic-Synthesis of Nucleosides Using Purine Nucleoside Phosphorylase. *Journal of Organic Chemistry* **1989**, 54, (19), 4692-4695.
18. Shirae, H.; Yokozeki, K.; Kubota, K., Enzymatic Production of Ribavirin from Orotidine by *Erwinia-Carotovora* Aj-2992. *Agricultural and Biological Chemistry* **1988**, 52, (6), 1499-1504.
19. Shirae, H.; Yokozeki, K.; Kubota, K., Enzymatic Production of Ribavirin from Pyrimidine Nucleosides by *Enterobacter-Aerogenes* Aj-11125. *Agricultural and Biological Chemistry* **1988**, 52, (5), 1233-1237.
20. Shirae, H.; Yokozeki, K.; Kubota, K., Enzymatic Production of Ribavirin. *Agricultural and Biological Chemistry* **1988**, 52, (1), 295-296.

21. Shirae, H.; Yokozeki, K.; Uchiyama, M.; Kubota, K., Enzymatic Production of Ribavirin from Purine Nucleosides by *Brevibacterium-Acetylicum* Atcc-954. *Agricultural and Biological Chemistry* **1988**, 52, (7), 1777-1783.
22. Barai, V. N.; Zinchenko, A. I.; Eroshevskaya, L. A.; Kalinichenko, E. N.; Kulak, T. I.; Mikhailopulo, I. A., A universal biocatalyst for the preparation of base- and sugar-modified nucleosides via an enzymatic transglycosylation. *Helvetica Chimica Acta* **2002**, 85, (7), 1901-1908.
23. Streeter, D. G.; Witkowsk.Jt; Khare, G. P.; Sidwell, R. W.; Bauer, R. J.; Robins, R. K.; Simon, L. N., Mechanism of Action of 1-Beta-D-Ribofuranosyl-1,2,4-Triazole-3-Carboxamide (Virazole) - New Broad-Spectrum Antiviral Agent. *Proceedings of the National Academy of Sciences of the United States of America* **1973**, 70, (4), 1174-1178.
24. Willis, R. C.; Carson, D. A.; Seegmiller, J. E., Adenosine Kinase Initiates Major Route of Ribavirin Activation in a Cultured Human Cell Line. *Proceedings of the National Academy of Sciences of the United States of America* **1978**, 75, (7), 3042-3044.
25. Page, T.; Connor, J. D., The Metabolism of Ribavirin in Erythrocytes and Nucleated Cells. *International Journal of Biochemistry* **1990**, 22, (4), 379-383.
26. Wu, J. Z.; Walker, H.; Lau, J. Y. N.; Hong, Z., Activation and deactivation of a broad-spectrum antiviral drug by a single enzyme: Adenosine deaminase catalyzes two consecutive deamination reactions. *Antimicrobial Agents and Chemotherapy* **2003**, 47, (1), 426-431.
27. Hager, P. W.; Collart, F. R.; Huberman, E.; Mitchell, B. S., Recombinant Human Inosine Monophosphate Dehydrogenase Type-I and Type-Ii Proteins - Purification and Characterization of Inhibitor Binding. *Biochemical Pharmacology* **1995**, 49, (9), 1323-1329.
28. Lau, J. Y. N.; Tam, R. C.; Liang, T. J.; Hong, Z., Mechanism of action of ribavirin in the combination treatment of chronic HCV infection. *Hepatology* **2002**, 35, (5), 1002-1009.
29. Crotty, S.; Maag, D.; Arnold, J. J.; Zhong, W. D.; Lau, J. Y. N.; Hong, Z.; Andino, R.; Cameron, C. E., The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. *Nature Medicine* **2000**, 6, (12), 1375-1379.
30. Crotty, S.; Cameron, C.; Andino, R., Ribavirin's antiviral mechanism of action: lethal mutagenesis? *Journal of Molecular Medicine-Jmm* **2002**, 80, (2), 86-95.
31. Gilbert, H. J.; Drabble, W. T., Complementation Invitro between Guab Mutants of Escherichia-Coli-K12. *Journal of General Microbiology* **1980**, 117, (Mar), 33-45.
32. Collart, F. R.; Huberman, E. Method to Identify Specific Inhibitors of IMP Dehydrogenase. US Patent 6,153,398, Dec. 24, 1997, 2000.

33. Farazi, T.; Leichman, J.; Harris, T.; Cahoon, M.; Hedstrom, L., Isolation and characterization of mycophenolic acid-resistant mutants of inosine-5'-monophosphate dehydrogenase. *Journal of Biological Chemistry* **1997**, 272, (2), 961-965.
34. Colby, T. D.; Vanderveen, K.; Strickler, M. D.; Markham, G. D.; Goldstein, B. M., Crystal structure of human type II inosine monophosphate dehydrogenase: Implications for ligand binding and drug design. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, 96, (7), 3531-3536.
35. Prorise, G. L.; Wu, J. Z.; Luecke, H., Crystal structure of *Tritrichomonas foetus* inosine monophosphate dehydrogenase in complex with the inhibitor ribavirin monophosphate reveals a catalysis-dependent ion-binding site. *Journal of Biological Chemistry* **2002**, 277, (52), 50654-50659.
36. Schlippe, Y. V. G.; Riera, T. V.; Seyedsayamdost, M. R.; Hedstrom, L., Substitution of the conserved Arg-Tyr dyad selectively disrupts the hydrolysis phase of the IMP dehydrogenase reaction. *Biochemistry* **2004**, 43, (15), 4511-4521.
37. Gan, L.; Petsko, G. A.; Hedstrom, L., Crystal structure of a ternary complex of *Tritrichomonas foetus* inosine 5'-monophosphate dehydrogenase: NAD(+) orients the active site loop for catalysis. *Biochemistry* **2002**, 41, (44), 13309-13317.
38. Kerr, K. M.; Cahoon, M.; Bosco, D. A.; Hedstrom, L., Monovalent cation activation in *Escherichia coli* inosine 5'-monophosphate dehydrogenase. *Archives of Biochemistry and Biophysics* **2000**, 375, (1), 131-137.
39. Kerr, K. M.; Digits, J. A.; Kuperwasser, N.; Hedstrom, L., Asp338 controls hydride transfer in *Escherichia coli* IMP dehydrogenase. *Biochemistry* **2000**, 39, (32), 9804-9810.
40. Kerr, K. M.; Hedstrom, L., The roles of conserved carboxylate residues in IMP dehydrogenase and identification of a transition state analog. *Biochemistry* **1997**, 36, (43), 13365-13373.
41. Sintchak, M. D.; Nimmegern, E., The structure of inosine 5'-monophosphate dehydrogenase and the design of novel inhibitors. *Immunopharmacology* **2000**, 47, (2-3), 163-184.
42. Zhang, R. G.; Evans, G.; Rotella, F. J.; Westbrook, E. M.; Beno, D.; Huberman, E.; Joachimiak, A.; Collart, F. R., Characteristics and crystal structure of bacterial inosine-5'-monophosphate dehydrogenase. *Biochemistry* **1999**, 38, (15), 4691-4700.
43. Prorise, G. L.; Luecke, H., Crystal structures of *Tritrichomonas foetus* inosine monophosphate dehydrogenase in complex with substrate, cofactor and analogs: A structural basis for the random-in ordered-out kinetic mechanism. *Journal of Molecular Biology* **2003**, 326, (2), 517-527.
44. Allen, L. B.; Boswell, K. H.; Khwaja, T. A.; Meyer, R. B.; Sidwell, R. W.; Witkowski, J. T.; Christensen, L. F.; Robins, R. K., Synthesis and Anti-Viral Activity of Some Phosphates of

Broad-Spectrum Anti-Viral Nucleoside, 1-Beta-D-Ribofuranosyl-1,2,4-Triazole-3-Carboxamide (Ribavirin). *Journal of Medicinal Chemistry* **1978**, 21, (8), 742-746.

45. Erion, M. D.; Stoeckler, J. D.; Guida, W. C.; Walter, R. L.; Ealick, S. E., Purine nucleoside phosphorylase .2. Catalytic mechanism. *Biochemistry* **1997**, 36, (39), 11735-11748.

## CHAPTER III

### APPLICATION OF AN IMPROVED BIOCATALYST FOR NUCLEOTIDE ANALOGUE BIOSYNTHESIS

#### Introduction

Nucleoside analogues are important tools in chemotherapy and in combating parasitic infections, but are most important as fundamental agents in the protection against viral diseases. *In vivo*, these pharmaceuticals are metabolized into their 5'-phosphorylated counterparts by various kinases. Access to these activated forms of nucleotides is essential for metabolic, kinetic, and structure-activity relationship studies of new and potential nucleoside therapeutics. Chemical synthesis of nucleotide analogues is most commonly performed via 5'-phosphorylation of the corresponding nucleoside analogue, and often requires multiple protective group manipulations to control the regiochemistry of ribose phosphorylation. Moreover, these nucleoside precursors themselves are the products of multistep pathways that must address several potential synthetic challenges – selective activation of the anomeric position for nucleobase addition, control of stereochemistry (or resolution of diastomeric products) and purification of highly polar water soluble end products. Intermediates in these pathways are often unstable, particularly the anomerically activated ribosides. Each nucleoside analogue presents its own distinct synthetic challenges requiring the development of a variety of non-generalizable strategies.

Chapter I detailed several biosynthetic alternatives to chemical synthesis by purine nucleoside phosphorylase and nucleoside 2'-deoxyribosyltransferase. PNP and NdRT provide an alternative to chemical synthesis for production of nucleoside analogues, and have shown flexibility in substrate recognition. But nucleotide products still require an additional step of

phosphorylation of the nucleoside through chemical process, or by enzymatic phosphorylation by a nonspecific kinase. Direct enzymatic biosynthesis of phosphorylated nucleotide analogue end products is much less common, and only a few examples have been reported, due to a higher level of discrimination by these enzymes.

Modification of discriminating active sites has been achieved by directed evolution and applied for biosynthesis of nucleoside and nucleotide analogues. Chapter II describes the evolution of the purine salvage enzyme, HPRT, for a higher turnover of nucleobase substrate 1,2,4-triazole-3-carboxamide. The IMPDH inhibitory activity of the resulting nucleotide, ribavirin monophosphate (RMP), and subsequent cell death, was used as a means for selection of the improved biocatalyst. The resulting improvement of the mutant 8B3 for the TCA substrate raised interest into the activity for other nucleobase analogues. It was hypothesized that the mutations which relaxed specificity for TCA substrate, might also allow catalytic turnover of other substrates. To test this hypothesis, 8B3PRT was screened against several nucleobase analogues to test for activity. This content in this chapter was published in 2007 by Scism, R. A. *et al.*<sup>1</sup> is subject to copyright by American Chemical Society, and has been adapted with permission.

## Results

To prepare the purine transferase biocatalyst, 8B3PRT, the plasmid pRAS1001 (pET28a-based plasmid-containing the mutant *hpt* gene) was transformed into electrocompetent *E. coli* BL21(DE3) cells via electroporation. Cells were grown to an OD<sub>600</sub> of 0.6, and 8B3PRT protein was over expressed by induction with IPTG (1 mM). After a three-hour incubation at 37 °C, cultures were distributed into 2 mL aliquots, pelleted by centrifugation and lysed with

detergent. To allow the catalytic method to be more accessible, any further purification of the enzyme was omitted, and reactions were performed with the crude cell lysate.

To determine the utility of this newly evolved biocatalyst as a general nucleobase transferase, a high throughput assay was needed to screen potential substrate candidates. The common assay for HPRT activity, which relies on a colorimetric response in relation to hypoxanthine consumption, would not be applicable with nucleobase analogues. Therefore, a microtiter assay, based on coupled pyrophosphate release, was developed from an existing method for phosphate detection.<sup>2</sup> Briefly, a buffered solution of 8B3PRT biocatalyst was added, in combination with an excess of inorganic pyrophosphatase, to a microtiter plate containing a different nucleobase analogue in each well. After a 5 minute incubation period, reactions were quenched with the addition of a solution of zinc acetate and ammonium molybdate. Final colorimetric response was induced by addition of ascorbate solution, pH 5, and absorbance at 850 nm read after 17 minutes.

Interestingly, lysate of *E. coli* with 8B3PRT showed significantly different substrate acceptance profile selectivity than wild type HPRT. Figure III-1 and Figure III-2 show turnover of various base analogues with wild type and 8B3PRT on different scales for low and high turnover respectively. In addition to the efficient turnover of the triazole carboxamide (8), several other base analogues showed increased turnover. The greatest improvement in turnover was observed for bases 9, 10, and 11; hypoxanthine pyrazolo (3,4-*d*) pyrimidine (allopurinol), 6-bromoguanine and 6-thioguanine, with increases in activity of 13, 11, and 17 fold, respectively.



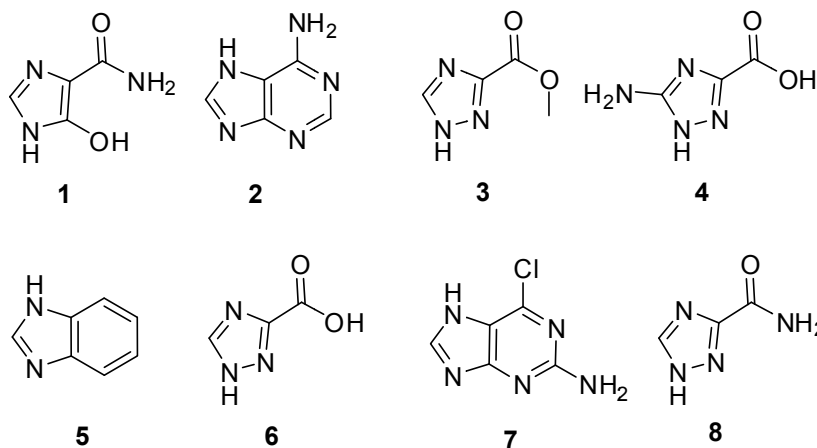
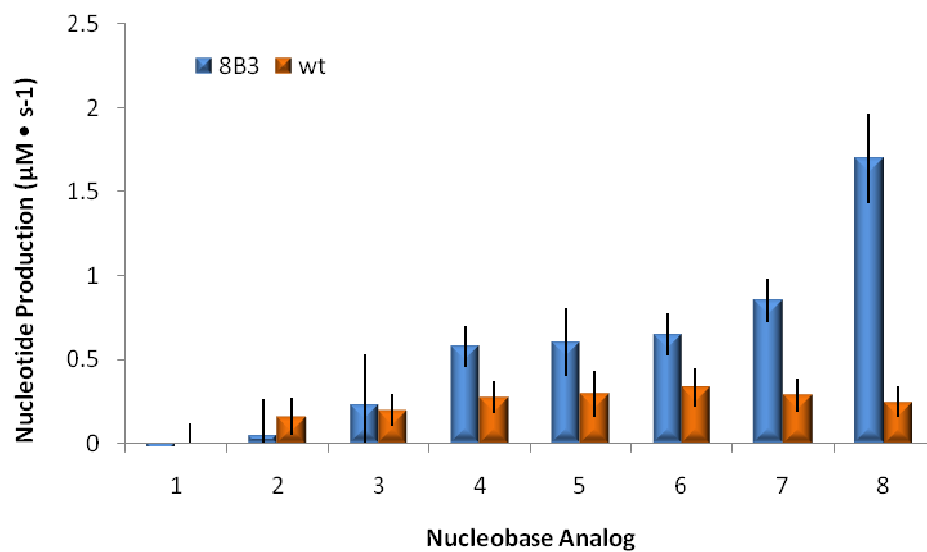
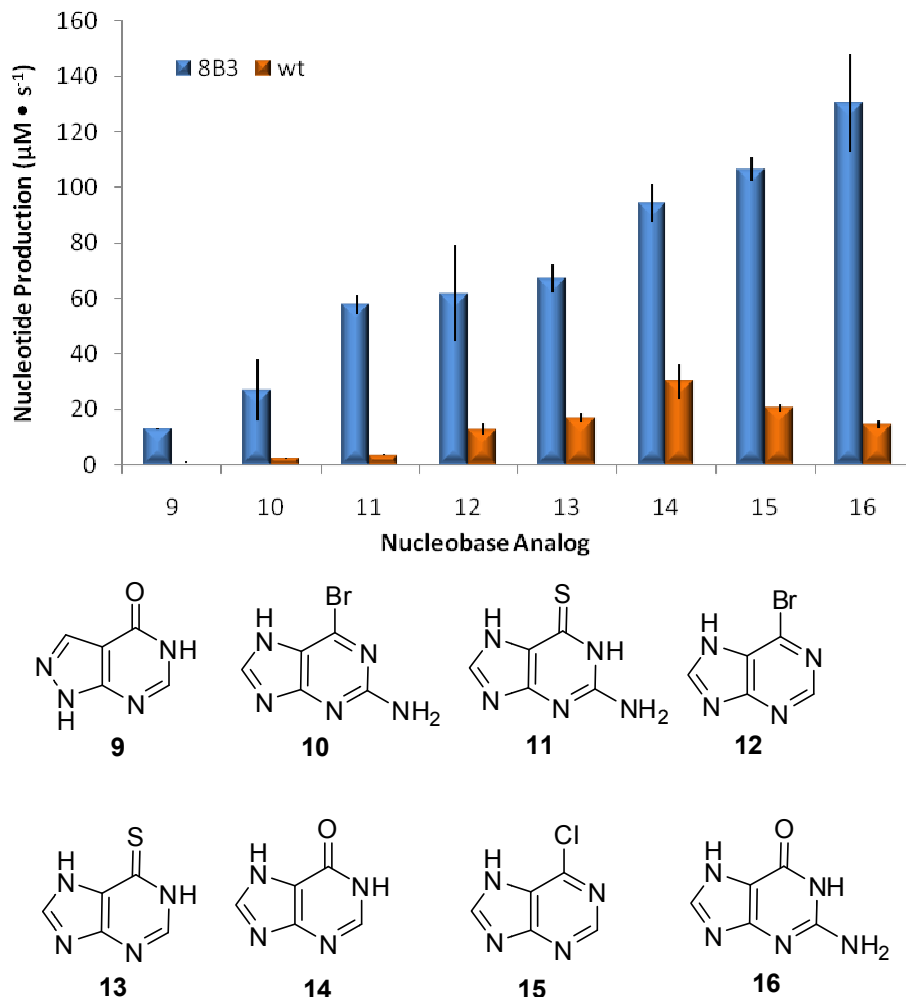


Figure III-1: Turnover of nucleobase analogues from the substrate library with relatively lower turnover, plotted on a smaller scale for visibility. Adapted from Scism, R. A. *et al.*<sup>1</sup>



**Figure III-2: Turnover of nucleobase analogues from substrate library: Highly active substrates listed in order of activity with 8B3PRT. Adapted from Scism, R. A. *et al.*<sup>1</sup>**

Since the pyrophosphate turnover assay is an indirect indication of product formation, percent conversion was determined via <sup>1</sup>H NMR, assaying for the disappearance of PRPP and formation of product, as shown in Table III-1. The catalyst was prepared as described previously, and 2 mL aliquots of cells were pelleted by centrifugation and treated with a commercial detergent (BugBuster, Novagen, Inc.), each providing enough catalyst for 30 one milligram-scale reactions. The cell preparation was added to buffered solutions of phosphoribosyl pyrophosphate (PRPP) and purine base analogues to catalyze the formation of nucleotides in less than 2 hours at 37 °C. To determine percent conversion, 10% D<sub>2</sub>O was added to reactions,

and a W5-WATERGATE water suppression pulse sequence was employed to eliminate the water resonance and accurately measure the ratio of the anomeric protons.<sup>3</sup> Often, conversion of PRPP was nearly quantitative, and since PRPP is a highly unstable technical grade compound containing contaminating inorganic phosphates and degradation products, percent conversion was used to determine reaction efficiency in place of percent yield. NMR images used in determining percent conversion are found in Appendix A.

**Table III-1: Percent conversion of select nucleobase analogues. Adapted from Scism, R. A. *et al.*<sup>1</sup>**

Nucleobase	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>
Percent Conversion	38	68	90	66	96	34	72	87	60	N.D.

Preparative scale reactions of the best substrates were repeated for subsequent purification. As nucleotides are not easily resolved under standard reversed-phase HPLC conditions, the nucleotide analogues were purified by a rapid anion exchange method. The entire reaction mixtures (1 mL) were applied to a quaternary amine column (Mono Q, GE Biosciences), and eluted with an increasing concentration gradient of triethyl ammonium carbonate or triethyl ammonium formate. These volatile buffers can be largely removed *in vacuo* subsequent to pH neutralization. Purified nucleotide monophosphates were confirmed by <sup>1</sup>H NMR, <sup>13</sup>C NMR, and mass spectrometry. NMR images of purified products can be found in Appendix B. The enzymatically catalyzed reactions were completely  $\beta$ -selective, as evidenced by the formation of a single anomeric product by NMR and by comparison of coupling constants and chemical shifts of biocatalytically generated nucleotides to known nucleotides.

## Discussion

### Activity Assay

Measurement of HPRT activity has been long established through various spectroscopic techniques. Most commonly, reaction progress is followed at an absorbance of 245 nm, the wavelength at which the maximum difference in absorbance of hypoxanthine and IMP can be measured.<sup>4-6</sup> The reported  $\Delta\epsilon$  ranges from 1900 to 2439  $M^{-1} \text{ cm}^{-1}$ , ostensibly allowing enough difference in absorbance for kinetic measurement. Other methods include scintillation counts with radio labeled substrates such as 8-<sup>14</sup>C hypoxanthine, or 8-<sup>3</sup>H IMP for analysis of forward and reverse reactions, respectively.<sup>6, 7</sup> And various colorimetric assays coupled with xanthine oxidase consumption of hypoxanthine have also been reported. These all involve the reaction of the resultant peroxide species with a chromophore or fluorophore such as Amplex Red.<sup>8</sup> These established methods are well suited for standard hypoxanthine turnover and kinetic determinations, but are inadequate for high throughput screening of non-native substrates with HPRT. For this application, a measurement of P<sub>Pi</sub> release upon nucleobase condensation is necessary.

Very few quantitative high throughput pyrophosphate quantification methods have been reported. One of the earliest methods for P<sub>Pi</sub> detection used gravimetric analysis of cadmium pyrophosphate precipitate after P<sub>Pi</sub> was combined with a solution of cadmium acetate.<sup>9</sup> Aside from mass spectroscopy, the predominant P<sub>Pi</sub> quantification method uses a combination of an ATP generating enzyme and firefly luciferase to yield a luminescent, quantifiable product.<sup>10-12</sup> In this process, also used in DNA pyrosequencing, ATP sulfurylase catalyses the transfer of P<sub>Pi</sub> with sulfate of adenylyl sulfate, to yield ATP. Firefly luciferase then catalyzes the addition of the *in situ* produced ATP with the biological pigment, luciferin, to

create luciferyl adenylate. Finally, luciferase catalyzed oxidation of luciferyl adenylate produces an amount of light proportional to the original amount of pyrophosphate present.

Although HPRT produces pyrophosphate during the reaction course, addition of inorganic pyrophosphatase, which cleaves PPI to 2 Pi, allows reaction progress to be monitored by Pi detection rather than PPI detection. Enzymatic methods for Pi detection are also available. One such example uses maltose phosphorylase to combine orthophosphate with maltose, giving glucose1-phosphate, which then combines with glucose to form gluconolactone with the help of glucose oxidase. This latter step produces hydrogen peroxide which combines with Amplex Red to form a fluorescent compound, resorufin.

Any of the above described methods for PPI or Pi detection could have been attempted in the high throughput screen for base analogue turnover. However, the multistep requirements of the coupled enzymatic methods added unnecessary complexity in comparison to the chosen method of zinc acetate and ammonium molybdate. In combination with inorganic pyrophosphatase, the zinc acetate method was preferred for its simplicity, reproducibility, and dynamic range, capable of Pi detection from 6  $\mu$ M to nearly 1 mM (compared to 2 to 150  $\mu$ M for the commercially available EnzChek kit from Molecular Probes).

### **Requirements of a Nucleobase**

Upon inspection of nucleobase preference by 8B3PRT, common structural requirements can be found. The natural substrate hypoxanthine, as well as the other bases accepted, all have an electronegative atom present at the 6 position of the purine ring. Adenine (2) and benzimidazole (5), which does not meet this requirement, showed no activity. This observation is consistent with previous suggestions that hypoxanthine hydrogen bonds with a conserved lysine residue, Lys135, shown in Figure III-3. It has also been suggested that Asp163 could

provide hydrogen bonding with NH<sub>2</sub> of guanine.<sup>13</sup> This aspartate residue would move out of position as the enzyme adjusts to a slightly different conformation with hypoxanthine-like substrates. For this reason, guanine-like analogues possessing a NH<sub>2</sub> substituent at the C2 position appear to be equally as favorable as their NH<sub>2</sub> lacking counterparts. This portion of the active site is apparently the most relaxed in regards to the nucleobase recognition, as indicated, not only by the optional NH<sub>2</sub> substituent at C2, but also due to atoms C2 and N3 being completely optional components of the substrate, as evident in TCA (8). A second aspartate residue, Asp107, is thought to deprotonate N7, driving forward the addition of the nucleobase at N9. However, allopurinol (9) has nitrogen at position N6 rather than N7 and interestingly, the nitrogen is not protonated at that position. Also, comparison of compound 6 to 8 underlines the importance of nitrogen occurring at position 1. Backbone carbonyl of Val157 is within hydrogen bonding distance of this N1, and provides the driving force for this particular selectivity. Finally, all substrates provide some degree of  $\pi$ -stacking interaction with Phe156.<sup>1</sup>

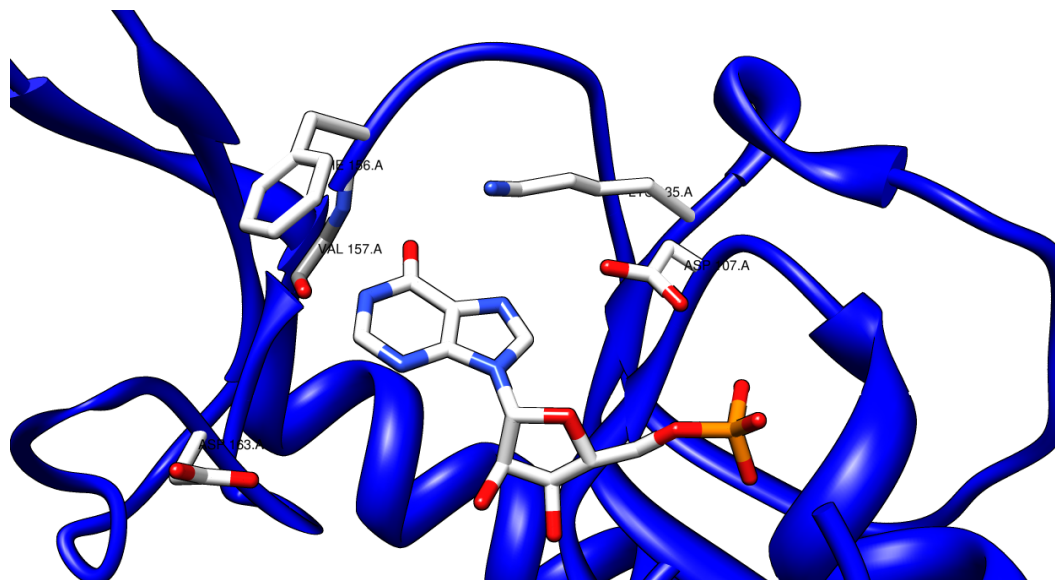


Figure III-3.: Key active site residues of HPRT, adapted from crystal structure of wt HPRT from *E. coli*, obtained by Guddat *et al.*<sup>13</sup> Modeled with UCSF Chimera.<sup>14</sup>

Ultimately, 8B3PRT showed a completely shifted substrate preference in comparison to wt HPRT. The evolved biocatalyst showed significantly increased performance for TCA (8), and for bases 9, 10, and 11; hypoxanthine pyrazolo (3,4-*d*) pyrimidine (allopurinol), 6-bromoguanine and 6-thioguanine, with increases in activity of 13, 11, and 17 fold, respectively. Application of the WATERGATE water peak suppression method provides an insight into reaction progress when complications of product or substrate decomposition prevent usual methods of evaporation and dissolving compounds for NMR analysis. Finally, the use of nonpurified biocatalyst in the one-step synthetic reaction, and the simple one-step purification by anion exchange, provides a convenient alternative for nucleotide analogue synthesis over traditional chemical means.

## Experimental Methods

**Preparation of Phosphoribosyltransferase Biocatalyst.** An *E. coli* strain BL21(DE3) containing pET28a(+) (Novagen Inc, San Diego, CA) based plasmids, pRAS1002 and pRAS1001, express wild-type HPRT and 8B3PRT respectively as N-terminal hexahistidine fusion proteins. Freshly transformed strains were incubated in LB medium containing 50 µg/mL kanamycin, at 37°C, with shaking at 225 rpm. Cultures were induced with 1 mM IPTG after reaching an OD<sub>600</sub> of 0.6. Following an additional three hours of incubation, the OD<sub>600</sub> of the cultures was measured and cell pellets were generated by centrifugation of 2 mL aliquots (4 °C, 13,000 rpm, 5 min). Pellet aliquots were stored at -80°C prior to use. Directly prior to PRT reactions, frozen cell pellets were resuspended in 300 µL BugBuster protein extraction reagent (Novagen Inc, San Diego, CA) and gently agitated via orbital shaking at room temperature for 10 minutes and centrifuged (4°C, 13,000 rpm for 5 min). Extract was used directly without purification.

**HPRT/8BPRT Protein Quantification.** PRT extract (10  $\mu$ L) was added to Laemmli sample buffer (20  $\mu$ L) with 5%  $\beta$ -mercaptoethanol and denatured for 5 minutes at 95°C. After cooling on ice for 5 minutes, protein solutions (5  $\mu$ L) were added to wells of a 12% polyacrylamide gel and separated at 200 V for 40 min. Hexahistidine tagged proteins were visualized with InVision His-Tag In-gel Stain (Invitrogen, Carlsbad, CA) according to manufacturers protocol. His-tagged protein bands were quantified using a 532/1064nm laser with 555nm long pass filter (Molecular Imager FX System, Bio-Rad, Hercules, CA). The linear range was established by measurement of a series of several enzyme dilutions, and relative protein concentration was determined by comparison of gel band intensities within this linear range.

**Microplate HPRT Turnover Assay.** Assay Mix (1 mL) was freshly prepared containing 30  $\mu$ L of PRT extract and a final concentration of 100 mM Tris-HCl, pH 7.3, 100 mM  $MgCl_2$ , 1 mM PRPP and 8.3 U/mL inorganic pyrophosphatase (from Baker's yeast, Sigma Chemical Inc.). Reactions were initiated by adding Assay Mix (95  $\mu$ L per well) to triplicate wells containing 5  $\mu$ L concentrated base analogue stock solutions (Table III-2) and allowed to proceed for five minutes before quenching with 80  $\mu$ L of a solution of 200 mM zinc acetate and 30 mM ammonium molybdate, pH 5. After 30 seconds of incubation, 20  $\mu$ L of 20% ascorbic acid, pH 5 was added, and absorbance at 850 nm was recorded after 17 minutes. Background absorbance values from base analogue, cell lysate, and PRPP were subtracted from the reaction measurement and pyrophosphate turnover was determined by comparison to a standard curve.



**Table III-2: Concentration of nucleobase stock solutions and final concentrations in wells.**  
Adapted from Scism, R. A. *et al.*<sup>1</sup>

Base analogue	Analogue number	Stock soln (mM)	Solvent	Final concn (mM)
5-hydroxy-1H-imidazole-4-carboxamide	1	177.8	DMSO	8.89
adenine	2	100	0.1M NaOH	5
methyl 1H-1,2,4-triazole-3-carboxylate	3	723.8	DMSO	36.19
5-amino-1H-1,2,4-triazole-3-carboxylic acid	4	100	0.1M NaOH	5
benzimidazole	5	200	DMSO	10
1H-1,2,4-triazole-3-carboxylic acid	6	75	0.1M NaOH	3.75
6-chloroguanine	7	100	DMSO	5
1H-1,2,4-triazole-3-carboxamide	8	100	DMSO	5
allopurinol	9	172	DMSO	8.6
6-bromoguanine	10	389	DMSO	19.45
6-thioguanine	11	80	DMSO	4
6-bromopurine	12	268.8	DMSO	13.44
mercaptapurine monohydrate	13	20	DMSO	1
hypoxanthine	14	100	DMSO	5
6-chloropurine	15	284.6	DMSO	14.23
guanine	16	20	1M NaOH	1

**Reaction Conditions for Nucleotide Analogue Preparation and Purification.** PRT extract (10  $\mu$ L) was added to an aqueous solution containing a final concentration of 10 mM PRPP, 10 mM base analogue, 12 mM TRIS pH 7.3, and 12 mM MgCl<sub>2</sub>, for a total reaction volume of 1 mL. Reactions were incubated at 37°C for two hours and filtered through a 0.25  $\mu$ m syringe filter (CA, Whatman) prior to purification via anion exchange chromatography (Tricorn MonoQ 5/50 GL, 5 mL anion exchange column, GE Biosciences, Piscataway, NJ). Nucleotide analogues were eluted using a 20 column volume gradient of 0 - 40 mM triethylammonium formate, pH 3.5 (or 0-300 mM triethylammonium carbonate, pH 7.0) at 2 mL/min. Fractions of 1 mL, with UV absorbance monitored at 254 nm, were assayed for the presence of ribosides by spotting

fractions on silica TLC plates and staining with *p*-anisaldehyde (9.2 mL *p*-anisaldehyde, 63.75 mL acetic acid, 338 mL 95% ethanol, and 12.5 mL H<sub>2</sub>SO<sub>4</sub>). Fractions containing nucleotide analogues were pooled and evaporated *in vacuo*. Residue was then redissolved in 1 mL water, 200  $\mu$ L (14.8 M) ammonium hydroxide added, and evaporated a second time *in vacuo*.

**Synthesis of 1,2,4-triazole-3-carboxamide (8).** Methyl-1*H*-1,2,4-triazole-3-carboxylate (1.9733 g, 15.5 mmol) (Aldrich) was added to a solution of 7N ammonia in methanol (50 mL) and stirred at room temperature for 2 days. 1,2,4-triazole-3-carboxamide, formed as a white precipitate, was collected by vacuum filtration, washed with anhydrous methanol and dried to yield 646 mg (5.8 mmol 37% yield). <sup>1</sup>H-NMR (400 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  8.39 (s, 1H), 7.93 (s, 1H), 7.68 (s, 1H). <sup>13</sup>C-NMR (150 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  159.9, 153.7, 147.5. Analysis calcd for C<sub>3</sub>H<sub>4</sub>N<sub>4</sub>O: C, 32.15; H, 3.60; N, 49.98; O, 14.27. Found: C, 32.03; H, 3.63; N, 48.97; O, 14.82. HRMS(ESI+) calcd for C<sub>3</sub>H<sub>4</sub>N<sub>4</sub>O<sup>+</sup> = 112.037962 *m/z*. Observed 112.0378 *m/z*.

### Nucleotide Characterization

**Inosine MP (17)** <sup>1</sup>H-NMR (400 MHz, D<sub>2</sub>O):  $\delta$  8.43 (s, 1H), 8.10 (s, 1H), 6.03 (d, J=5.8 Hz, 1H), 4.39 (m, 1H), 4.26 (m, 1H), 3.93 (m, 2H). <sup>13</sup>C-NMR (125 MHz, D<sub>2</sub>O):  $\delta$  158.6, 148.8, 146.0, 139.9, 123.6, 87.2, 84.6, 74.5, 63.6. LRMS (ESI-) calcd for C<sub>10</sub>H<sub>12</sub>N<sub>4</sub>O<sub>8</sub>P<sup>-</sup> = 347.0 *m/z*. Observed mass = 347.1 *m/z*.

**Mercaptopurine MP (18)** <sup>1</sup>H-NMR (400 MHz, D<sub>2</sub>O):  $\delta$  8.58 (s, 1H), 8.26 (s, 1H), 6.05 (d, J=5.6 Hz, 1H), 4.41 (m, 1H), 4.27 (m, 1H), 3.96 (m, 2H). <sup>13</sup>C-NMR (125 MHz, D<sub>2</sub>O):  $\delta$  175.1, 145.9, 144.5, 142.3, 135.1, 87.6, 84.7, 74.6, 70.5, 63.8. LRMS (ESI-) calcd for C<sub>10</sub>H<sub>12</sub>N<sub>4</sub>O<sub>7</sub>PS<sup>-</sup> = 363.0 *m/z*. Observed mass = 363.0 *m/z*.

**Thioguanine MP (19)** <sup>1</sup>H-NMR (400 MHz, D<sub>2</sub>O):  $\delta$  8.20 (s, 1H), 5.83 (d, J=6.1 Hz, 1H), 4.38 (m, 1H), 4.22 (m, 1H), 3.93 (m, 2H). <sup>13</sup>C-NMR (125 MHz, D<sub>2</sub>O):  $\delta$  173.6, 153.5, 148.1, 140.4, 128.3,

86.8, 84.4, 73.7, 70.5, 63.74. LRMS (ESI-) calcd for  $C_{10}H_{13}N_5O_7PS^-$  = 378.0 *m/z*. Observed mass = 378.1 *m/z*.

**Ribavirin MP (20)**  $^1H$ -NMR (400 MHz,  $D_2O$ ):  $\delta$  8.77 (s, 1H), 5.91 (d,  $J=4.3$  Hz, 1H), 4.58 (m, 1H), 4.39 (m, 1H), 4.23 (m, 1H), 3.87 (m, 2H).  $^{13}C$ -NMR (150 MHz,  $D_2O$ ):  $\delta$  162.8, 156.2, 145.8, 92.0, 84.6, 74.7, 70.3, 63.4. LRMS (ESI-) calcd for  $C_8H_{12}N_4O_8P^-$  = 323.0 *m/z*. Observed mass = 323.0 *m/z*.

**Allopurinol MP (21)**  $^1H$ -NMR (400 MHz,  $D_2O$ ):  $\delta$  8.17(s, 1H), 8.13 (s, 1H), 6.19 (d,  $J=5.2$  Hz, 1H), 4.79 (t,  $J=5.4$  Hz, 1H), 4.44 (t,  $J=4.7$ , 1H), 4.20 (dd,  $J=4.6$ , 1H), 3.85 (m, 2H).  $^{13}C$ -NMR (125 MHz,  $D_2O$ ):  $\delta$  159.8, 153.6, 148.8, 136.9, 106.4, 87.4, 83.8, 72.7, 70.4, 64.1. LRMS (ESI-) calcd for  $C_{10}H_{12}N_4O_8P^-$  = 347.0 *m/z*. Observed mass = 347.1 *m/z*.

**6-chloropurine MP (22)**  $^1H$ -NMR (400 MHz,  $D_2O$ ):  $\delta$  8.88 (s, 1H), 8.68 (s, 1H), 6.18 (d,  $J=5.2$  Hz, 1H), 4.72 (m, 1H), 4.43 (m, 1H), 4.29 (m, 1H), 3.96 (m, 2H).  $^{13}C$ -NMR (150 MHz,  $D_2O$ ):  $\delta$  151.8, 151.3, 150.1, 145.5, 131.1, 87.8, 84.7, 74.6, 70.4, 63.5. LRMS (ESI-) calcd for  $C_{10}H_{11}ClN_4O_7P^-$  = 365.0 *m/z*. Observed mass = 364.9 *m/z*.

**6-bromopurine MP (23)**  $^1H$ -NMR (400 MHz,  $D_2O$ ):  $\delta$  8.82 (s, 1H), 8.63 (s, 1H), 6.17 (d,  $J=5.4$  Hz, 1H), 4.42 (m, 1H), 4.30 (m, 1H), 4.02 (m, 1H).  $^{13}C$ -NMR (150 MHz,  $D_2O$ ):  $\delta$  151.8, 150.0, 145.1, 142.1, 133.9, 87.9, 84.1, 74.4, 70.2, 64.2. LRMS (ESI-) calcd for  $C_{10}H_{11}BrN_4O_7P^-$  = 409.0 *m/z*. Observed mass = 408.9 *m/z*.

**NMR Estimation of Percent Conversion.** PRT extracts (6  $\mu$ L) were added to an aqueous solution of 1 mM PRPP, 30  $\mu$ L base analogue (Table III-3), 12 mM TRIS pH 7.3, and 12 mM  $MgCl_2$  to a final reaction volume of 600  $\mu$ L. After incubation at 37°C for one hour, 60  $\mu$ L  $D_2O$  was added to facilitate NMR analysis. Integrated values of the anomeric peak of nucleotides (5.83 - 6.94 ppm) and PRPP (5.7 ppm) were determined using a WATERGATE-W5 water suppression pulse sequence.<sup>3</sup>  $^1H$  NMR experiments were acquired using a Bruker AV-III spectrometer

equipped with a 5 mm cryogenically cooled NMR probe operating at a proton Larmor frequency of 600.13 MHz. Water suppression was performed with the W5-watergate sequence (zgpgw-5, Bruker pulse sequence nomenclature) using a double gradient echo<sup>1</sup>. Prior to acquiring water suppression experiments, the proton 90° pulse (p1) was accurately calibrated using a nutation experiment as was the distance of the next null used for the binomial water suppression delay (d19). The data was processed using sinusoidal function to remove the residual water signal. Percent conversion was calculated by dividing the integration of the anomeric nucleotide peak (5.83 - 6.94 ppm) by the sum of the integrated values of the anomeric PRPP and degradation products. Additional peaks of unknown compounds appear in NMR for nucleotides **22** and **23** (from bases **15** and **12**). The values of these additional peaks were included in the sum of degradation peaks for calculation of percent conversion of nucleotides **22** and **23**. These unidentified peaks are likely degradation products of **22** and **23** rather than of PRPP.

**Table III-3: Percent conversion values and base analogue concentrations used in reactions, adapted from Scism, R. A. *et al.*<sup>1</sup>**

Analogue Number	Final Concn (mM)	Percent Conversion
<b>5</b>	10	0
<b>7</b>	3.33	0.38
<b>8</b>	10	0.68
<b>9</b>	8.63	0.9
<b>10</b>	12.98	0.66
<b>11</b>	5	0.96
<b>12</b>	24.85	0.34
<b>13</b>	11.02	0.72
<b>14</b>	5	0.87
<b>15</b>	28.47	0.6

### 8B3PRT Protein Sequence.

MGSSHHHHHSSGLVPRGSHMMKHTVEVMIPAEIKARIAELGRQITERYKDSGSDMVLVGLLRGSFMFM  
ADLCREVQVSHEVDFMTASSYSGSMSTTRDVKILKDLDEDIRGKDVLIVEDIIDSGNTLSKVREILSLREPKSLAI  
CTLLDKPSRREVNVPVEFIGFSIPDEF**A**VGYGIDYAQRYRHLP**I**IGKVILLDE

Underlined sequence represents the 6 X His tag and additional amino acids resulting from expression in pET 28a.

### 8B3PRT Nucleotide Sequence.

ATGATGAAACATACTGTAGAAGTAATGATCCCCGAAGCGGAGATTAAGCGCGTATCGCCGAACGGGT  
CGTCAGATTACTGAGCGTTACAAAGACAGCGGCAGCGATATGGT**A**CTGGTGGGTCTGCTGCGTGGCTCA  
TTTATGTTTATGGCGGACCTGTGCCGTGAAGTTCAGGTATCTCATGAAGTCGACTTTATGACCGCCTCCAG  
CTACGGTAGCGGCATGTCCACCACCCGTGATGTGAAAATCCTCAAAGATCTGGATGAAGATATCCG**G**GG  
CAAGGACGTGCTGATTGTTGAAGATATCATCGACTCGGGGAATACTACTGTGCGAAAGTGCCTGAGATCTT  
AAGCCTGCGCGAACCGAAGTCGCTGGCGATTTGTACGCTGCTGGATAAACCGTCCCCTCGTGAAGTGAA  
CGTCCCGGTAGAATTTATCGGTTTCTCGATCCCGGATGAGTTT**G**GGTGGGTTACGGCATTGATTACGCA  
CAGCGTTACCGTCATCTGCCG**C**ATATCGGCAA**G**GTGATTCTGCTGGACGAGTAA

**Bold white** represents a mutation from the wild type *hpt* gene and translated enzyme.

## References

1. Scism, R. A.; Stec, D. F.; Bachmann, B. O., Synthesis of nucleotide analogues by a promiscuous phosphoribosyltransferase. *Organic Letters* **2007**, 9, (21), 4179-4182.
2. Erion, M. D.; Stoeckler, J. D.; Guida, W. C.; Walter, R. L.; Ealick, S. E., Purine nucleoside phosphorylase .2. Catalytic mechanism. *Biochemistry* **1997**, 36, (39), 11735-11748.
3. Liu, M. L.; Mao, X. A.; Ye, C. H.; Huang, H.; Nicholson, J. K.; Lindon, J. C., Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy. *Journal of Magnetic Resonance* **1998**, 132, (1), 125-129.
4. Vos, S.; deJersey, J.; Martin, J. L., Crystal structure of Escherichia coli xanthine phosphoribosyltransferase. *Biochemistry* **1997**, 36, (14), 4125-4134.
5. Subbaya, I. N. S.; Sukumaran, S.; Shivashankar, K.; Balaram, H., Unusual substrate specificity of a chimeric hypoxanthine-guanine phosphoribosyltransferase containing segments from the Plasmodium falciparum and human enzymes. *Biochemical and Biophysical Research Communications* **2000**, 272, (2), 596-602.

6. Xu, Y. M.; Eads, J.; Sacchettini, J. C.; Grubmeyer, C., Kinetic mechanism of human hypoxanthine-guanine phosphoribosyltransferase: Rapid phosphoribosyl transfer chemistry. *Biochemistry* **1997**, 36, (12), 3700-3712.
7. Hochstadt, J., Hypoxanthine Phosphoribosyltransferase and Guanine Phosphoribosyltransferase from Enteric Bacteria. *Methods in Enzymology* **1978**, 11.
8. Zhou, M. J.; Diwu, Z. J.; PanchukVoloshina, N.; Haugland, R. P., A stable nonfluorescent derivative of resorufin for the fluorometric determination of trace hydrogen peroxide: Applications in detecting the activity of phagocyte NADPH oxidase and other oxidases. *Analytical Biochemistry* **1997**, 253, (2), 162-168.
9. Cohn, G.; Kolthoff, I. M., Determination of Pyrophosphate. *Industrial and Engineering Chemistry* **1942**, 14, (11), 886-890.
10. Nyren, P.; Lundin, A., Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. *Analytical Biochemistry* **1985**, 151, (2), 504-509.
11. Jansson, V.; Jansson, K., Enzymatic chemiluminescence assay for inorganic pyrophosphate. *Analytical Biochemistry* **2002**, 304, (1), 135-137.
12. Arakawa, H.; Karasawa, K.; Igarashi, T.; Suzuki, S.; Goto, N.; Maeda, M., Detection of cariogenic bacteria genes by a combination of allele-specific polymerase chain reactions and a novel bioluminescent pyrophosphate assay. *Analytical Biochemistry* **2004**, 333, (2), 296-302.
13. Guddat, L. W.; Vos, S.; Martin, J. L.; Keough, D. T.; De Jersey, J., Crystal structures of free, IMP-, and GMP-bound *Escherichia coli* hypoxanthine phosphoribosyltransferase. *Protein Science* **2002**, 11, 1626-1683.
14. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, 25, (13), 1605-1612.

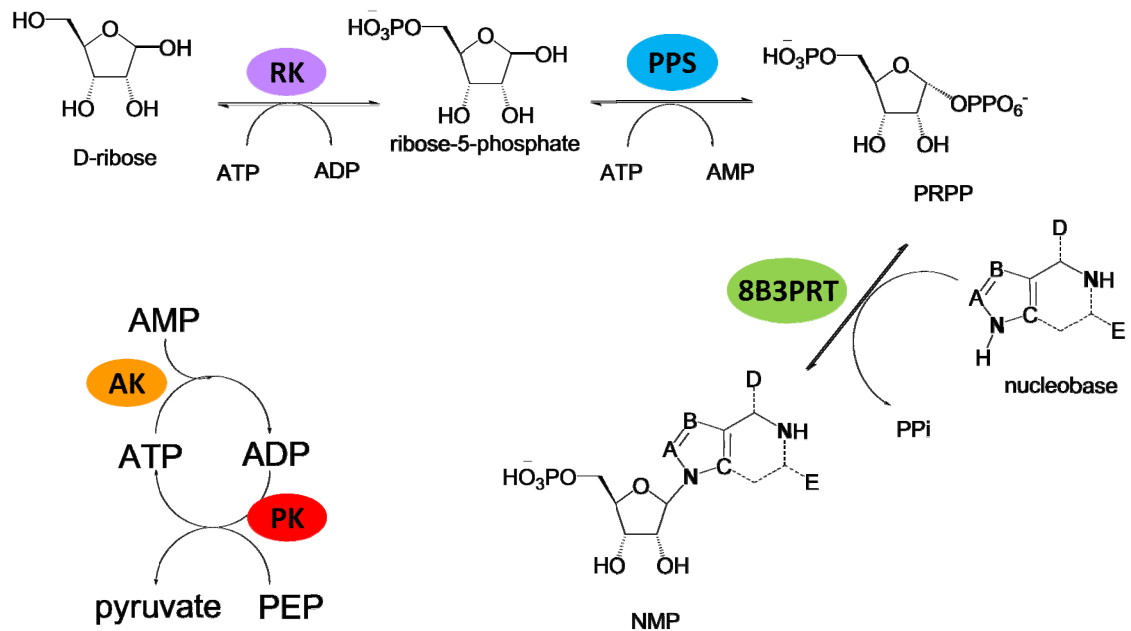
## CHAPTER IV

### PATHWAY ENGINEERING FOR CASCADE SYNTHESIS OF NUCLEOTIDE ANALOGUES IN A CROSS-LINKED ENZYME AGGREGATE

#### Introduction

Chemical synthesis of nucleotide analogues suffers from potential obstacles of non regio- and stereospecific steps, need for isolation of unstable intermediates, and a need for multiple protecting group manipulations. Chapter II describes the directed evolution of HPRT to create a biocatalyst with relaxed substrate specificity, enabling biosynthesis of nucleotide analogue ribavirin monophosphate. The evolved 8B3PRT enzyme has an even greater utility as a general nucleotide analogue biocatalyst. The wide substrate profile of this biocatalyst and a simple method for nucleotide synthesis and single step purification is described in Chapter III.

The evolved 8B3PRT biocatalyst greatly simplifies the task of synthesizing nucleotide analogues, but is not without limitations. The required PRPP substrate has a half life of less than one hour at 37 °C, and its availability is often limited and costly, with a current market price of approximately 1.2 million dollars per mole. Thus, a multistep enzymatic pathway which forms nucleotide analogues from ribose, via *in situ* PRPP biosynthesis, is of obvious utility. To this end, ribokinase (RK) and phosphoribosyl pyrophosphate synthetase (PPS) were cloned into expression vectors for nucleotide synthesis. (Scheme IV-1) Additional enzymes adenylate kinase (AK) and pyruvate kinase (PK) were also added for ATP regeneration. The work presented in this chapter is an adaptation of a previous publication by Scism, R. A. *et al*, 2010,<sup>1</sup> copyright of Wiley-VCH Verlag GmbH & Co. KGaA. Figures, schemes, tables and text have been reproduced with permission.



**Scheme IV-1: Three-step pathway from ribose to nucleotide utilizing ribokinase (RK), phosphoribosyl pyrophosphate synthetase (PPS), and 8B3PRT, with ATP regeneration enzymes adenylate kinase (AK) and pyruvate kinase (PK). Adapted from Scism, R. A. *et al.*<sup>1</sup>**

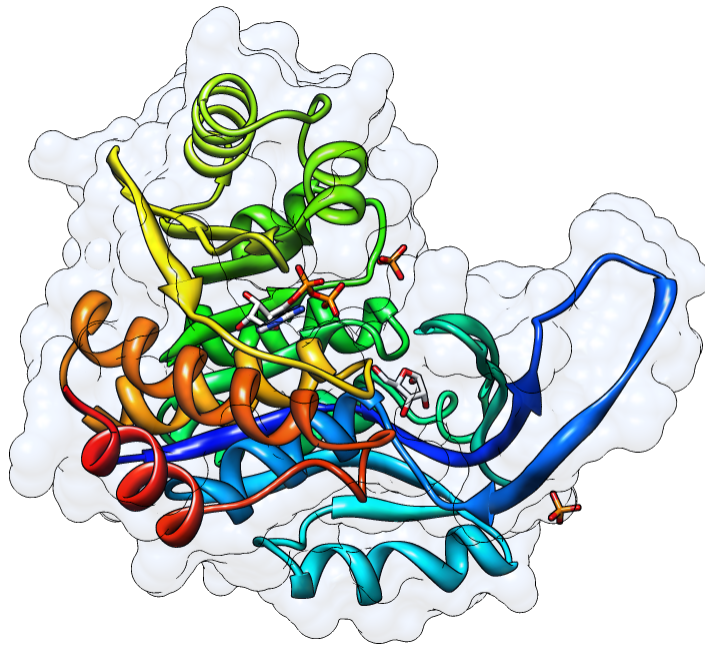
The proposed three step pathway described will alleviate problems associated with PRPP expense and instability. However an additional complication to be addressed is the stability of the catalytic pathway proteins themselves. Most enzyme preparations, purified or crude, have a relatively short lifespan at temperatures necessary for catalysis. Enzyme immobilization provides a potential solution for potentially increasing the lifetime of these biocatalysts.

## Ribokinase

Ribokinase (E.C. 2.7.1.15), the first enzyme in the ribose to nucleotide pathway, is responsible for converting D-ribose to D-ribose-5-phosphate (r-5-p) with the aid of ATP and a divalent cation,  $Mg^{2+}$ . In cells, phosphorylation of ribose by ribokinase acts to trap ribose within the cell, and allows the incorporation of ribose into the pentose phosphate pathway, and for use



in the synthesis of nucleotides, histidine, and tryptophan.<sup>2</sup> The enzyme exists as a homodimer in solution, with each 33 kDa subunit comprised of 309 amino acids, and two domains. The larger of the two domains consists of a nine-stranded  $\beta$ -sheet flanked by  $\alpha$ -helices, and provides the majority of specific binding interactions with the substrates. (Figure IV-1) The second domain, a small  $\beta$ -sheet, acts as a lid over the active site during catalysis, binding only through hydrophobic interactions with the ribose substrate. This smaller domain also provides the dimer interface, with the lids of the two subunits interwoven, each contributing to the active site lid of the other subunit.<sup>2,3</sup>



**Figure IV-1: Ribokinase with bound ADP and ribose. Adaptation of crystal structure obtained by Sigrell, J. A. *et al*,<sup>4</sup> rendered by UCSF Chimera software.<sup>5</sup>**

The *apo* form of the enzyme is relatively open before substrates are bound. Upon ribose binding, the  $\beta$ -sheet lid closes and simultaneously opens the nucleotide active site for Mg/ATP complex binding.<sup>3, 6</sup> The divalent cation counteracts the multiple negative charges of the

phosphate groups, and arranges the ATP into a favorable conformation for binding.<sup>7</sup> An aspartate residue then deprotonates the ribose O5 hydroxyl, setting up the substrate for nucleophilic attack upon the  $\gamma$ -phosphate of the nucleotide. An anion hole, activated by a monovalent cation ( $K^+$ ), helps stabilize the transition state of the reaction.<sup>7</sup> The ADP by-product is then released, and the enzyme re-opens to its original conformation with simultaneous r-5-p release.<sup>6</sup>

ATP and  $Mg^{2+}$  directly bind with the enzyme and are essential for catalysis. Additionally, inorganic phosphate has been shown to increase catalysis. At a pH of 6.2, addition of 20 mM inorganic phosphate increased  $V_{max}$  by 23 fold, and decreased  $K_M$  of ribose from 0.61 mM to 0.21 mM. Phosphate apparently induces a lesser effect on the  $K_M$  of ATP, holding at 2.5 mM at pH 6.2. (Others report a  $K_M$  for ATP of 213  $\mu$ M, but under different reaction conditions, and higher pH).<sup>8</sup> At higher pH, the phosphate activation effects were slightly diminished, and at pH 7.4, 20 mM phosphate lowered  $K_M$  of ribose from 0.18 mM to 0.11 mM, and increased  $V_{max}$  by a factor of less than six. The latter results are more relevant, as the pH optimum for the enzyme is between 8 and 9, at temperatures between 30 and 50 °C. The mechanism for activation by phosphate is unclear, although the authors speculate phosphate invokes and stabilizes a certain open configuration of the protein, allowing ribose to bind with a reduced  $K_M$ .<sup>6</sup>

The interaction of monovalent cations is less ambiguous, as a crystal structure of ribokinase with bound cesium places the ion in close proximity to the active site.<sup>7</sup> Cesium was used in the crystal structure due to its greater electron density, but ammonium, and to much greater extent potassium cations have a more profound effect on catalysis.<sup>7, 8</sup> The crystal structure shows a monovalent cation would bind between two loops immediately adjacent to the anion hole of the active site, which helps stabilize the transition state of the reaction. The buried nature of the cation binding site suggests conformational changes upon binding, as it is

completely sequestered from solvent when bound. Addition of 5 mM potassium increased catalysis of ribokinase by at least 60 fold.<sup>7</sup> These catalytic improvements via addition of phosphate and potassium, as well as the pH optimum and requirement for  $Mg^{2+}$ , were carefully considered when designing the reaction buffer for the CLEA experiment.

### **Phosphoribosyl Pyrophosphate Synthetase (PPS)**

Phosphoribosyl pyrophosphate synthetase (PPS) (E.C. 2.7.6.1), also referred to as phosphoribosyl diphosphate synthetase, or ribose-5-phosphate pyrophosphotransferase, catalyzes the formation of PRPP and AMP by transfer of a diphosphate group from ATP to the 1-hydroxyl of ribose-5-phosphate. PRPP product is an essential precursor of purine and pyrimidine nucleotides, and the aromatic amino acids histidine and tryptophan.<sup>9</sup> The key position in the metabolism of the cell justifies the strict regulation of PPS by ADP, and in by GDP in mammalian systems.<sup>10</sup> The 33 to 34 KDa PPS of *E. coli* spans 314 amino acids, and has a strict requirement for  $Mg^{2+}$  and inorganic phosphate.<sup>11</sup> Different values have been reported for pH optima, ranging from 9.8 for *E. coli*, 9.2 for *Salmonella typhimurium*, and pH 8.8 and 7.4 for human isozymes I and II.<sup>10, 11</sup> Michaelis-Menten constants for r-5-p are 138  $\mu$ M, 160  $\mu$ M, 52  $\mu$ M and 83  $\mu$ M for *E. coli*, *S. typhimurium*, and human isozymes I and II, respectively. Measured  $K_M$  for MgATP complex is 94  $\mu$ M, 43  $\mu$ M, and 21  $\mu$ M for *E. coli*, *S. typhimurium*, and human type I.<sup>10, 11</sup>

Crystal structures for PPS have been solved for *Homo sapiens*, *Methanocaldococcus jannaschii*, and *Bacillus subtilis*, the latter being the focus of this discussion. The protein is functional as a propeller-like hexamer structure exhibiting C3 symmetry. Each subunit is comprised of two domains, and the active site is located between them, including residues from two subunits. The six subunits are arranged with the N-termini clustered near the central axis. In this quaternary form, 30% of each subunit surface is embedded. The binding site for r-5-p is

located exclusively by a single subunit, near the C-terminus. The ATP binding site is situated between the two domains of each subunit, and includes residues from N-terminal portions of neighboring subunits.<sup>12</sup>

As previously mentioned,  $Mg^{2+}$  and  $P_i$  are required for PPS catalysis, and ADP is a potent inhibitor of the enzyme. Steady state kinetic analysis, in combination with substrate placement in crystal structures, reveals a complicated mechanism of activation and inactivation through a combination of both allosteric and competitive inhibition. The reaction proceeds by attack of the 1-hydroxyl of r-5-p upon the  $\beta$ -phosphate of ATP, resulting in the transfer of the  $\beta,\gamma$ -diphosphoryl moiety to r-5-p.<sup>13</sup> Substrates bind in the order of  $Mg^{2+}$ , MgATP, and r-5-p. Inorganic phosphate can bind randomly to the enzyme, either before the other substrates, or after. At saturating concentrations of phosphate,  $P_i$  binds first, lowering the  $K_M$  for  $Mg^{2+}$  and the remaining substrates. ADP inhibition is competitive at lower concentrations, or allosteric in nature at saturating concentrations of r-5-p. Allosteric  $P_i$  activation and ADP inhibition are thought to occur by competing for binding to the same site.<sup>9</sup>

Studies in substrate specificity of PPS indicate some flexibility in the ATP binding pocket. A PPS from spinach accepted dATP, GTP, CTP, and UTP in addition to ATP as alternate diphosphoryl donors.<sup>14</sup> PPS was strictly specific for r-5-p, however, the authors report no activity detected with deoxyribose-5-phosphate. The lack of specificity in the ATP active site has been employed in reverse reactions for biosynthesis of ATP analogues via diphosphorylation of corresponding monophosphate nucleotides.<sup>15,16</sup>

## **Immobilization Techniques**

There are several other reasons for immobilization of enzymes, in addition to increasing stability. Immobilization facilitates biocatalyst separation from product, to recover enzymes for

reuse, and presents the possibility for continuous flow processes and automation. Applications of enzyme immobilization include laboratory scale synthesis<sup>17</sup> or therapeutic biotechnology<sup>18</sup>, although the primary applications are for industrial processes and products.<sup>19</sup>

There is no one best way to immobilize an enzyme and the optimal method varies per enzyme, based on the desired outcome of immobilization, and the properties of the enzyme to be used. Most methods of immobilization can be organized in to one of four main categories – covalent binding, entrapment or encapsulation, adsorption, and cross-linking.<sup>20</sup>

### ***Covalent Binding***

Covalent binding of an enzyme to a solid support normally occurs between a functional group of an amino acid on the enzyme surface with an activated residue on a solid support. Many different types of solid supports are available, though it has been shown that the more hydrophilic supports show greater activities than hydrophobic supports. Consequently, cellulose, dextran (Sephadex), starch, and agarose (Sepharose) are commonly used. Additional supports include porous glass and silica, and commercially available supports such as Sepabeads,<sup>21</sup> and Amberzyme (polymethylmethacrylate).<sup>22</sup> The chemistry involved in linkage is often by formation of an isourea linkage, diazo linkage, peptide bond, or alkylation reaction. Glyoxyl-agarose activated with aldehyde groups react with primary amines on the enzymes (Lys, Arg) and hydrophobic acrylic epoxy-supports attach by interaction between the nucleophilic groups of the enzyme and the epoxy groups of the support.<sup>17</sup> CNBr is often used to activate the hydroxyl functional groups in polysaccharide support materials, such as sepharose.<sup>20, 23</sup> Covalent linkage has the advantage of high retention of protein, and little to no leakage into reaction vessel. The disadvantage is that care must be taken not to deactivate the enzyme by reacting important residues in the active site upon immobilization with support in the coupling rxn.<sup>20</sup>

### ***Entrapment and Encapsulation***

Entrapment and encapsulation are very similar techniques in which enzyme molecules are free in solution, but restricted by the lattice structure of a gel (entrapment) or within a semi permeable membrane (encapsulation). Gels can be created by ionotropic gelation (alginate), temperature induced gel formation (agarose, gelatin), or via organic polymerization (polyacrylamide)<sup>24</sup>. In the latter case, pore size of the gel can be adjusted by relative amounts of monomer and cross linking agent, with the goal of allowing free movement of substrate and product, yet preventing leakage of enzyme.<sup>20</sup> Other examples include polyvinyl alcohol<sup>25</sup> and hydroxy(propylmethyl) cellulose.<sup>26</sup> Nylon and cellulose nitrate have been used for encapsulation, as well as a unique dendrimer silica nanocomposite.<sup>27</sup> These methods have the advantage of bypassing potentially harmful chemical modification of the enzyme through covalent attachment. However, problems with diffusion and separation of the protein from bulk solution can be difficult since the immobilized particle may have a density similar to that of the buffered solvent.

### ***Adsorption***

Adsorption through electrostatic forces such as van der Waals, ionic, and hydrogen bonding provides an additional method of immobilization that does not require perturbation of the enzyme by chemical modification, due to the utilization of existing surface chemistry. Adsorption methods are rather simple, consisting of mixing the enzyme with support, followed by a washing step to remove nonbound material. These methods, however are highly reversible, causing unwanted desorption, or leakage. This same disadvantage, however, allows regeneration of spent catalyst with fresh enzyme. Examples include ionic adsorption,

hydrophobic attachment to solid supports such as polypropylene<sup>28</sup> or porous polystyrene resin<sup>29</sup> and Ni affinity with hexa-histidine tagged proteins.<sup>30-33</sup>

### ***Cross-linking***

Cross-linking enzymes is unique in being a support-free immobilization. Proteins are joined together into a large three dimensional complex structure, without attachment to other materials. Cross-linking reactions are often used to enhance other methods of immobilization by reducing leakage, and typically consist of a glutaraldehyde connection with primary amines. Unwanted reactions with catalytically important residues are the biggest disadvantage, and amount of cross-linking reagent used must be carefully manipulated to prevent enzyme deactivation. Lysine rich albumins and gelatin have been used to provide spacers to minimize close proximity problems due to over cross linking.

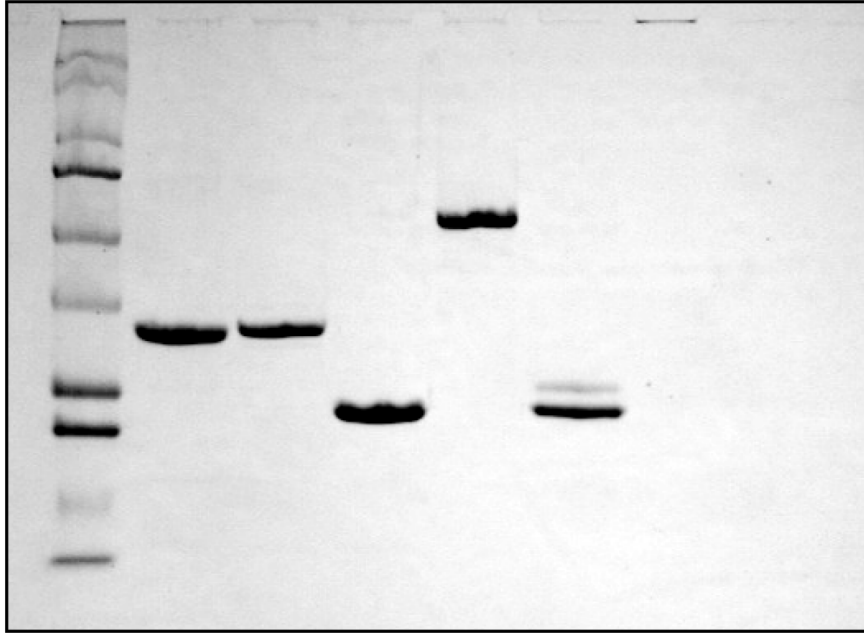
Enzyme cross linking with glutaraldehyde has been practiced for several decades for purposes of immobilization and stabilization.<sup>34</sup> Cross-linking with glutaraldehyde is often used in combination with other support methods such as gelatin-fibrinogen,<sup>35</sup> polystyrene nanofibers,<sup>36</sup> and polyvinyl alcohol, due to the soluble nature of enzymes without additional support.<sup>37</sup> Accounts of glutaraldehyde cross-linking of ammonium sulfate precipitated proteins can be found as early as 1967, with several other accounts leading into the latter part of the 21<sup>st</sup> century.<sup>38-40</sup> Ammonium sulfate has long been used as a means to purify proteins by precipitation and subsequent reconstitution. Cross-linked enzyme aggregates, now commonly referred to as CLEAs, are proteins precipitated with salts or organic solvent, followed by cross linking of the insoluble material. The term CLEA was coined in 2000 by Sheldon, who later registered the now prevalent term as the trademark for CLEA Technologies of the Netherlands.<sup>41</sup>

Self-immobilization by the CLEA method has been a very successful method for increasing enzyme stability for a variety of different applications. Thermal stability was significantly improved upon CLEA formation of  $\beta$ -galactosidase, glutaryl acylase, feruloyl esterase, penicillin G acylase, and a fungal laccase.<sup>40, 42-45</sup> CLEAs have also been used for stabilization in extremes of pH<sup>46, 47</sup> and for use in organic solvents, including acetone, alcohols, and various ethers, that would inactivate free enzymes.<sup>37, 41, 45, 48</sup> Finally, CLEA self-immobilization is a very simple and economical procedure, eliminating the need for bulky and expensive solid supports, and in some cases, purified enzyme is unnecessary. It is for these reasons that self immobilization by the CLEA method was determined superior to other methods, and could provide the needed stability for the proposed ribose to nucleotide pathway.

## Results

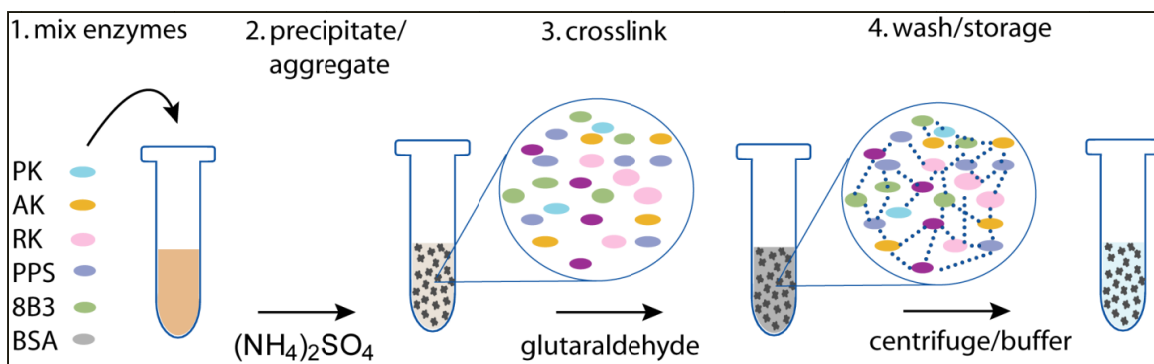
Ribokinase (*rbsk*) and pyruvate kinase (*pykF*) genes were cloned from *Escherichia coli* K12 via PCR amplification. Phosphoribosyl pyrophosphate synthetase (*prs*) and adenylate kinase (*adk*) were similarly cloned from *Bacillus cereus*. Cloned genes were inserted into commercial vectors pCDF Duet-1, pET28a, or pET22b (EMD Chemicals Inc., New Jersey USA) to allow for the addition of either *N*-terminal or *C*-terminal hexahistidine tags. All plasmids were then separately transformed into *E. coli* K12 and protein expression was induced with IPTG at log phase for 3 - 4 hours. Cells were collected by centrifugation, resuspended, and disrupted using a French pressure cell. Enzymes were purified by Ni<sup>2+</sup>-affinity chromatography (Figure IV-2) and imidazole was removed via gel filtration.<sup>1</sup>



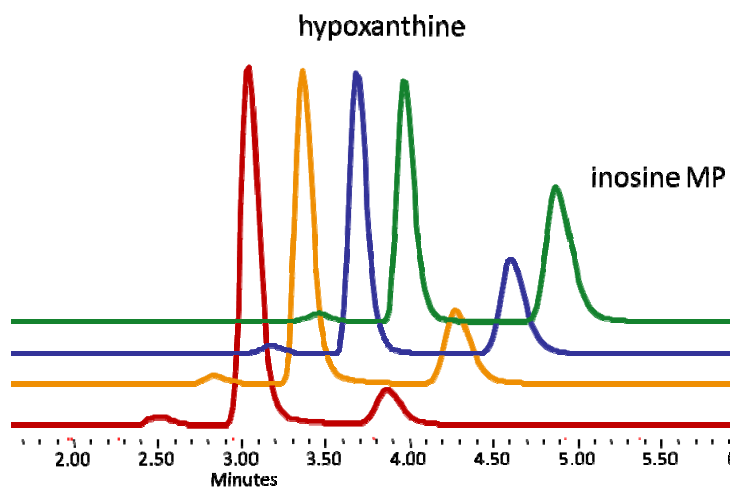


**Figure IV-2: SDS-PAGE gel of purified enzymes and CLEA; lane 1, marker; lane 2, 1.1  $\mu\text{g}$  RK; lane 3, 0.9  $\mu\text{g}$  PPS; lane 4, 1.8  $\mu\text{g}$  8B3; lane 5, 1.5  $\mu\text{g}$  PK; lane 6, 1.3  $\mu\text{g}$  AK; lane 7, 690  $\mu\text{g}$  CLEA aggregate. This image copyright of Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission. Originally published by Scism, R. A. *et al.*<sup>1</sup>**

Aliquots of CLEA particles were generated (Figure IV-3) by combining 50  $\mu\text{L}$  of each of the five pathway enzyme preparations and bovine serum albumin (10 mg/mL). The six protein mixture was precipitated by adding ammonium sulfate and the aggregated enzyme suspension was cross linked by adding glutaraldehyde and incubating for two hours with stirring at 4  $^{\circ}\text{C}$ . Cross-linked aggregate particles were readily collected by centrifugation and washed three times by resuspension in a storage buffer. Ratios of enzymes were optimized to maintain highest production level of nucleotide product, and prevent accumulation of AMP or ADP, as determined by NMR and HPLC.<sup>1</sup> (Figure IV-4)



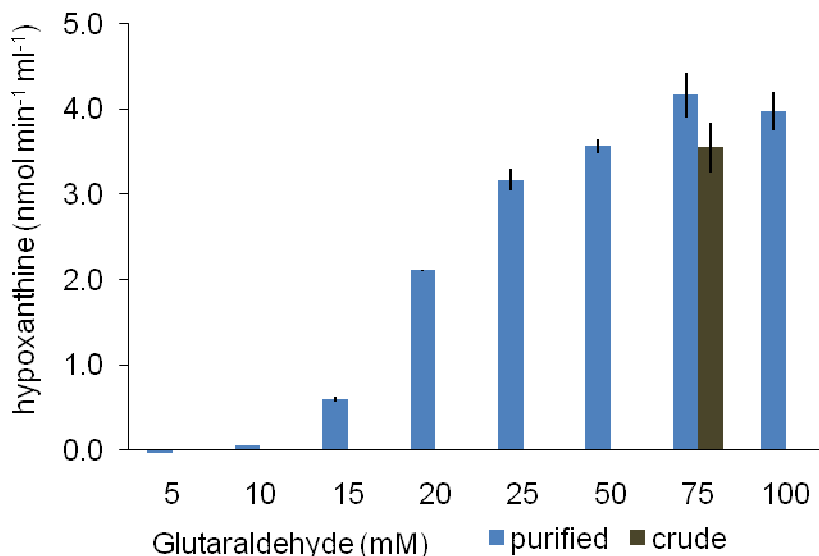
**Figure IV-3: CLEA preparation** - PK (pyruvate kinase), AK (adenylate kinase), RK (ribokinase), PPS (phosphoribosyl pyrophosphate synthetase), 8B3 (8B3PRT evolved biocatalyst), BSA (bovine serum albumin). This image copyright of Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission. Originally published by Scism, R. A. *et al.*<sup>1</sup>



**Figure IV-4: Reaction progress by HPLC** using ion pairing reagent and reverse phase chromatography. Adapted from Scism, R. A. *et al.*<sup>1</sup>

To determine the optimal concentration of glutaraldehyde used in nucleotide analogue pathway CLEA synthesis, activities of particles prepared with increasing glutaraldehyde concentrations were assayed by monitoring the conversion of D-ribose to inosine monophosphate (IMP). Preparations were magnetically stirred in reaction buffer with 5 mM ribose, 5 mM hypoxanthine, 6 mM ATP and 25 mM PEP. After 15 minutes, the reaction was centrifuged to remove enzymes and supernatants were assayed for hypoxanthine consumption

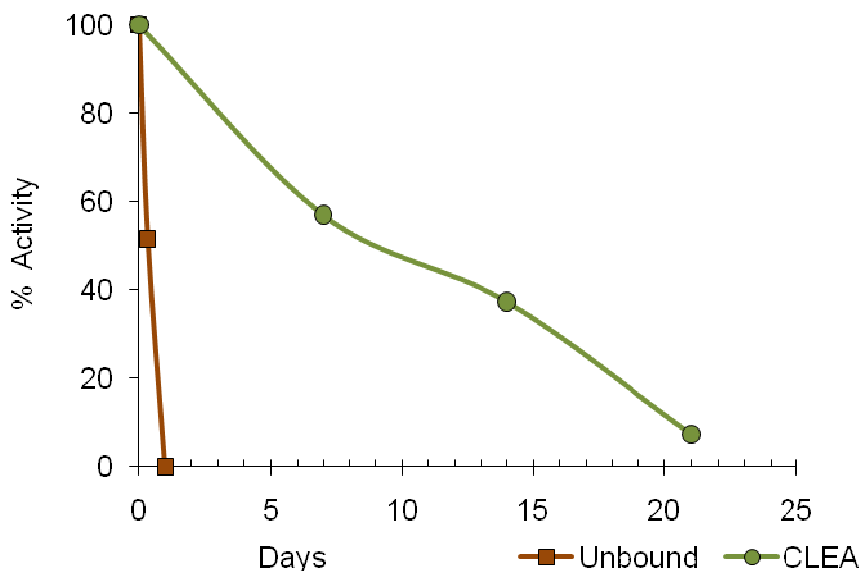
via a xanthine oxidase/INT coupled assay. Enzyme aggregates cross-linked with glutaraldehyde concentrations at 75 mM (Figure IV-5) demonstrated optimal activity of  $67.9 \text{ nmol}\cdot\text{min}^{-1}\text{mL}^{-1}$  ( $4.17 \text{ nmol}\cdot\text{min}^{-1}\text{mL}^{-1}\text{mg}^{-1}$  protein). Thus, the aggregate turnover rate was approximately 30% less than pathway reactions using equivalent amounts of non aggregated enzymes (data not shown), which demonstrated a turnover of  $96.0 \text{ nmol}\cdot\text{min}^{-1}\text{mL}^{-1}$  ( $5.89 \text{ nmol}\cdot\text{min}^{-1}\text{mL}^{-1}\text{mg}^{-1}$  protein). At glutaraldehyde concentrations below 15 mM, aggregates redissolved in the wash steps and demonstrated little or no detectable activity.<sup>1</sup>



**Figure IV-5: Effect of glutaraldehyde concentration on activity with purified enzymes. Shown in dark brown is activity of a CLEA particle created with non-purified cell lysate. Adapted from Scism, R. A. *et al.*<sup>1</sup>**

The integrity of CLEA particles was confirmed by assaying for resolubilization of pathway enzymes. Turnover was undetectable in CLEA wash supernatants and no soluble protein was observed in Coomassie Blue stained polyacrylamide gels in CLEAs formulated with 25 to 100 mM glutaraldehyde. (Figure IV-2, lane 7) The stability of immobilized enzymes generated using 75 mM glutaraldehyde was assessed by extended incubation at 37 °C. Aliquots were removed at

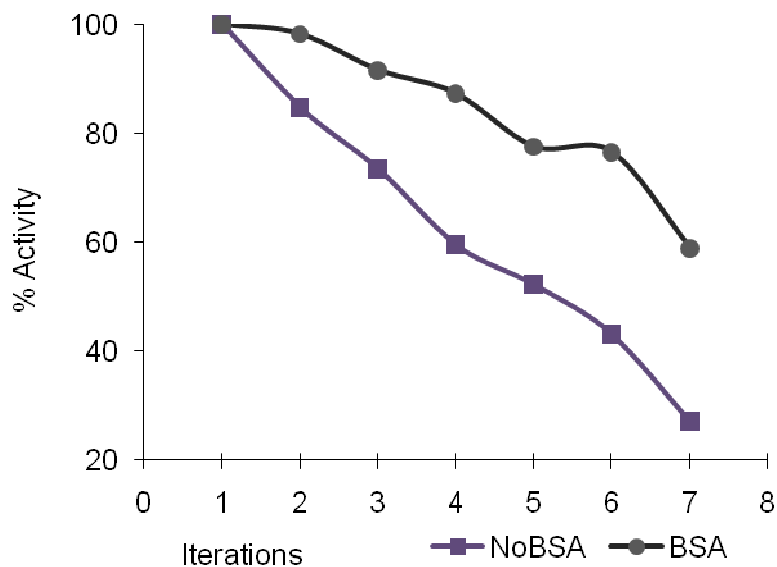
increasing time points and assayed using the hypoxanthine consumption assay as described above. As a benchmark, equivalent amounts of purified non-immobilized enzymes were mixed and incubated under comparable conditions. Remarkably, the immobilized pathway retained over 50% activity after 7 days and, at up to three weeks of accelerated aging, retained significant pathway activity. (Figure IV-6) In comparison, the equivalent soluble enzyme mixture retained no traces of pathway activity after 24 hours.<sup>1</sup>



**Figure IV-6: Stability of CLEA vs. unbound enzymes at 37°C. Adapted from Scism, R. A. *et al.*<sup>1</sup>**

CLEA residual activity, or reusability, was evaluated by assay of centrifugally recovered enzyme pellets. IMP formation was assayed and recovered pellets were subsequently washed three times in resuspension buffer and re-assayed. Following seven cycles, CLEA pellets retained ca. 30% of their original activity. (Figure IV-7) Loss of activity is presumably due to disintegration of the pellet with extensive handling rather than loss of enzyme, as no protein is detected in the washings. The addition of lysine-rich bovine serum albumin (BSA) to cross-linking reactions to generate CLEAs has been previously shown to increase enzyme activity, thermal stability,

residual activity, and to protect against proteolysis. The addition of BSA is proposed to enhance the formation of the reticulate, providing additional structure and preventing excessive cross-linking of catalytic enzyme. Consistent with these observations, inclusion of BSA in the nucleotide generation pathway significantly enhanced stability in reuse studies. Particles generated with BSA retained up to 60% of their original activity after seven uses, while retaining similar initial activity.<sup>1</sup>

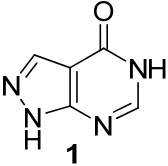
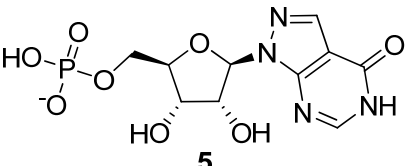
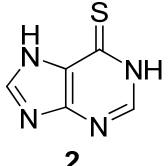
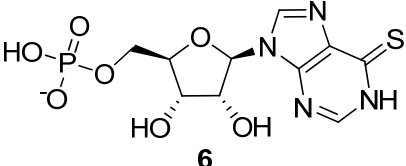
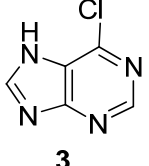
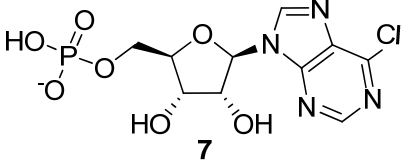
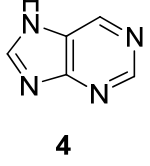
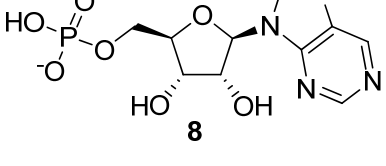


**Figure IV-7: CLEA residual activity after catalysis and subsequent washing. CLEAs prepared with and without added bovine serum albumin. Adapted from Scism, R. A. *et al.*<sup>1</sup>**

Given the known substrate flexibility of 8B3PRT, as described in Chapter III, a large complement of purine nucleotide analogues should be accessible by using this catalyst. The synthetic utility of the immobilized pathway catalyst for nucleotide analogue generation was demonstrated with selected base analogue substrates (Table IV-1). Nucleotide analogues were synthesized in 500 mL reactions by stirring CLEA pellets prepared with BSA (corresponding to 50 mL each enzyme) with 10 mM nucleobase, 12 mM ribose, 6 mM ATP and 30 mM PEP for two

hours. Reaction progress was monitored by WATERGATE NMR. Analogues were purified by DEAE-sepharose anion exchange resin eluted with an ammonium carbonate gradient and characterized by NMR. (See Appendix C for NMR images of purified products.) Substrates 1–3 showed complete or nearly complete conversion, while purine (4) conversion was 48% without optimization.<sup>1</sup> Products 5–7 were previously generated from PRPP using the terminal transferase enzyme 8B3PRT, described in Chapter III.

**Table IV-1: Nucleotide analogues prepared with the 5-enzyme CLEA, purified by anion exchange. This table is copyright of Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission. Originally published by Scism, R. A. *et al.*<sup>1</sup>**

<u>Substrate</u>	<u>Product</u>	<u>% Conversion</u>
 <p>1</p>	 <p>5</p>	98
 <p>2</p>	 <p>6</p>	100
 <p>3</p>	 <p>7</p>	100
 <p>4</p>	 <p>8</p>	48

## Discussion

### ATP Regeneration

The potent inhibition of PPS by Mg/ADP poses a potential dilemma in the proposed ribose to nucleotide pathway, when considering that ADP is a byproduct of ribose phosphorylation by ribokinase. Addition of phosphoenol pyruvate (PEP) and pyruvate kinase (PK) eliminates the ADP hindrance by conversion to ATP and pyruvate. To avoid a buildup of AMP after the PPS step, addition of adenylate kinase, in combination with ATP, phosphorylates AMP back to ADP for another recycling step by PK. This recycling scheme not only avoids the PPS inhibition issue, but also allows use of catalytic amounts of ATP. Furthermore, maintaining a triphosphate species of ATP allow ease in purification of the monophosphate nucleotide analogue by anion exchange chromatography. It was also essential to balance the AK to PK ratio and provide an excess of PEP to ensure ATP was the dominant form of ATP over ADP or AMP.

### Nucleotide Purification

Purification of biological components from cellular lysate can involve multiple steps and intricate methods.<sup>49</sup> Anion exchange by DEAE-sepharose anion exchange resin in combination with an ammonium carbonate gradient provided a relatively simple method for purification of nucleotides for this work. Other reported methods for purification of nucleotides or similarly charged compounds were also considered.

Various chromatographic methods have been used for nucleotide purification with end product purity ranging from crude preparations to ultrapure purifications. In some instances, removal of aqueous solvent, followed by protein precipitation and filtration is sufficient.<sup>50</sup> Most purifications involve a protein removal step followed by additional chromatography for

nucleotide separation. Protein removal is achieved via desalting column,<sup>51, 52</sup> filtration of the precipitate formed by heating<sup>53-55</sup> or addition of solvent.<sup>56</sup> After protein removal, chromatographic separation of nucleotides is achieved using a single anion exchange resin<sup>52, 56</sup> or dual separations by both cation and anion exchange<sup>51, 53-55</sup>

A uniquely different method for nucleotide purification uses boronate affinity resin, which binds only *cis*-diols, exemplified in the 2',3' hydroxyls of the ribose moiety. This method has been used to purify nucleotides in similarly complex mixtures of charged species.<sup>57-62</sup> Several attempts were made to purify the nucleotide analogues in this study using this method, but a few caveats prevented this from being an efficient method. The bed volume, as noted by Scott, et al,<sup>57</sup> tends to swell upon application of the gradient, to the extent of double its original size. This dramatic change in volume often proved fatal to the integrity of the column vessel. It has also been documented that TRIS buffers bind to the resin, thus decreasing the column capacity.<sup>60</sup> A final, yet significant problem with this method, is that the adenosine nucleotide cofactors (AMP, ADP, and ATP) also bind and coelute with the end product nucleotide. Surprisingly, this issue was not addressed in the majority of instances utilizing this method. In the two reports that do address this issue, additional steps were necessary to bypass the issue. The enzymes used for the biosynthesis of 8-azaguanine monophosphate required ATP cofactors, but the authors used dATP and pyruvate kinase/adenylate kinase recycling system to avoid coelution of ATP. Others using ATP in the enzymatic method required the addition of a second chromatographic step using anion exchange resin subsequent to boronate affinity chromatography.

In instances of isotopic nucleotide biosynthesis involving supplementation of bacterial cultures with isotopic precursors, purification of cellular RNA provides another route for nucleotide purification. RNA is often collected via aqueous extraction after a protein



precipitation step, followed by RNA digestion by RNase P<sub>1</sub>. This provides nucleotide monophosphates which may be used without further purification,<sup>63</sup> or subsequently purified by the previously discussed methods of ion exchange<sup>64,65</sup> or boronate affinity resin.<sup>66,67</sup>

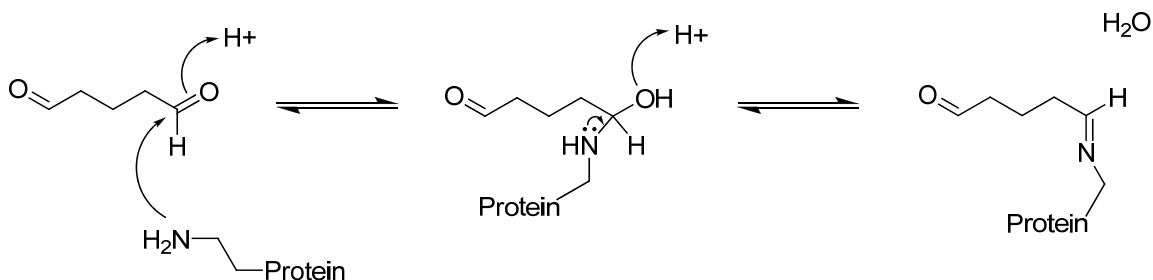
Finally, chromatographic separation of nucleotides may be achieved with reverse phase solid support in combination with an ion pairing reagent. Commonly, ammonium salts of amines such as triethylamine, tetrabutylamine, or dimethylhexylamine are used.<sup>68-72</sup> Classically, this is reserved for analytical techniques, and rarely for preparative separations. More recently porous graphitic carbon has proven to be a superior method for analytical separations of nucleotides, and preparative scale columns are now available.<sup>73</sup>

### **CLEA Stability Rationale**

Protein precipitation as a means for concentration or partial purification has been known for many years. Precipitation by salting out with saturated ammonium sulfate solutions is useful due to its high solubility, low density, and kosmotropic effects<sup>74</sup>. For proteins in solution, a substantial number of hydrophobic groups reside on the surface, often in patches and are forced into contact with the aqueous solvent. The salt solution competes with the protein for water and eventually strips the protein of its aqueous shell.<sup>75</sup> The exposed hydrophobic patches on one protein molecule then interact with those on another, resulting in aggregation.<sup>76</sup> The aggregates are compact clusters of concentrated active protein in native form. The cross-linking of these aggregates with glutaraldehyde then locks these active conformations in place, and rigidifies the molecule. Since denaturing and subsequent inactivation of a protein is attributable to unfolding of its structure, cross-linking prevents conformational changes towards the denatured state caused by external chemical, thermal or mechanical forces,<sup>77</sup> and provides steric protection from proteolysis.<sup>19</sup>

## Glutaraldehyde Cross-linking Mechanism

The mechanism of the cross-linking of glutaraldehyde with protein has been a matter of speculation for many years. Glutaraldehyde interacts with functional groups of tyrosine, histidine, tryptophan, and cysteine, although the most favorable interaction is with the primary amine side chain of lysine. At first thought, it may appear obvious that aldehyde interaction with primary amines is occurring via Schiff base formation. (Scheme IV-2) Some CLEA preparations go so far as to include addition of sodium borohydride to reduce the suspected Schiff bases formed to prevent the reverse reaction of hydrolysis.<sup>42</sup> In the CLEA preparation, no comparison was made to CLEAs made without reduction, but in other examples of cross-linking with glutaraldehyde to solid supports, addition of sodium borohydride did not increase stability.<sup>78-80</sup>



**Scheme IV-2: Cross-linking by Schiff base formation, condensed adaptation from Farris, S. *et al.*<sup>81</sup>**

Schiff base formation is reversible, and the equilibrium constant is usually unfavorable for non-conjugated imine formation in aqueous solutions. Conversely, immobilization by the CLEA method is irreversible, and survives treatment with urea, semicarbazide, and wide ranges of pH, ionic strength and temperature. Furthermore, the  $pK_a$  of a Schiff base is no higher than 5 or 6 whereas the  $pK_a$  of the modified lysine residue is 8 to 8.5.<sup>82</sup> This observed stability of glutaraldehyde cross-linked proteins refutes Schiff base formation as the only mechanism.

NMR analysis of commercial glutaraldehyde solutions has revealed primarily polymeric material.<sup>34, 82</sup> (Figure IV-8) It is suspected that these polymers could allow covalent, irreversible Michael addition like adducts, (Figure IV-9) which would account for the observed stability.<sup>82</sup> The existence of these adducts were later confirmed in x-ray analysis of protein crystal structures. Additional covalent adducts of anti-Markownikov products (Figure IV-9) were also observed under acidic conditions.<sup>34</sup>

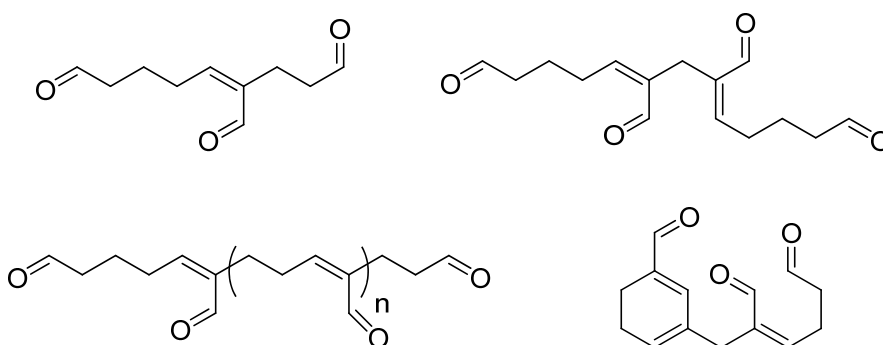


Figure IV-8: Glutaraldehyde polymers<sup>34, 82</sup>

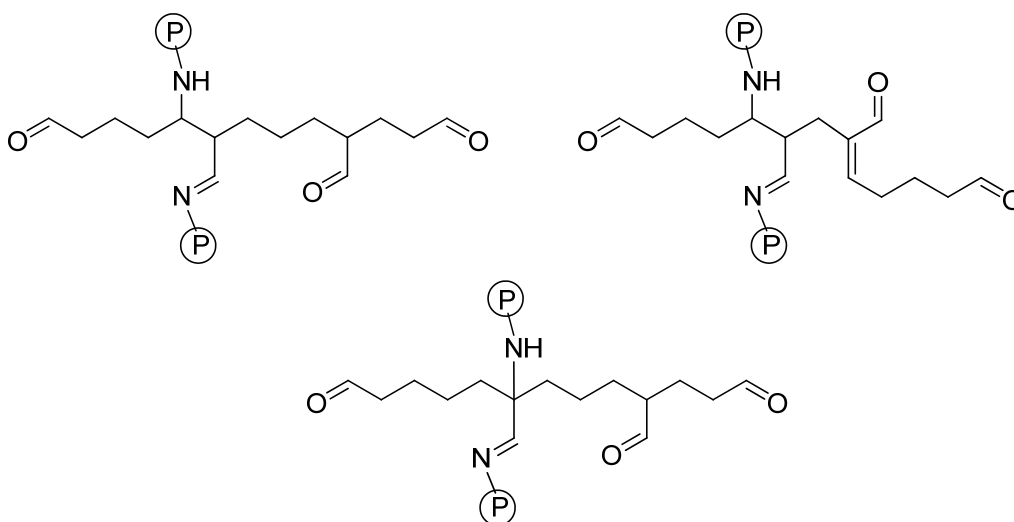


Figure IV-9: Protein adducts with glutaraldehyde polymers including Michael addition and anti-Markownikoff products.<sup>34, 82</sup>

The addition of BSA in the formation of CLEAs has been shown to increase enzyme activity, thermal stability, residual activity<sup>83</sup>, and to protect against proteolysis.<sup>45</sup> SEM images of CLEAs with BSA as a proteic feeder indicate a more structured, ball-like formation compared to the amorphous aggregates found without BSA.<sup>83</sup> The additional, lysine rich protein adds to the reticulate, providing additional structure and prevents excessive cross-linking of catalytic enzyme.<sup>40</sup>

## Conclusion

The multiple enzyme CLEA pathway described is a substantial improvement over the single step PRPP method. It simultaneously addresses the PRPP and enzyme stability problems and provides a catalyst that can be removed by centrifugation, or filtration, and reused. The development of this cascade process, without intermediate recovery, drastically reduces the time and cost required in similar single step processes. The realization that efficient CLEAs can be made with crude cell lysate, (Figure IV-5) rather than non-purified enzymes opens up the possibilities for replacement of whole cell catalysts. Ideally, this method could be applied to other biocatalytic pathways currently impeded by problems with toxicity of substrate, product, or intermediates, allowing greener synthetic alternatives to current production methods.

## Experimental Methods

**General.** All chemicals were purchased from Sigma-Aldrich Inc, Acros Organics, or Fisher Scientific unless otherwise noted. NMR experiments were acquired using a 9.3 T Bruker magnet equipped with a Bruker DRX console operating at 400.13 MHz. Protein and nucleotide purification was performed on AKTA FPLC, Amersham Biosciences. HPLC assays utilized a Waters HPLC with 2996 PDA detector.

**Bacterial strains and plasmids.** Bacterial strains and plasmids are listed in Table IV-2.

**Table IV-2: Bacterial strains and plasmids used. Adapted from Scism, R. A. *et al.*<sup>1</sup>**

Strain or plasmid	Relevant properties
<b>Strain</b>	
<i>E. coli</i> BL21(DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm</i> (DE3)
<b>Plasmids</b>	
pET28a(+)	
pRAS1001	0.54-kb NdeI-HindIII insert from epPCR product of <i>hpt</i> in pET-28a(+); Kan <sup>r</sup> , T7lac promoter
pRAS1005	1.41-kb NdeI-NotI insert from PCR product <i>pykF</i> in pET-28a(+); Kan <sup>r</sup> , T7lac promoter
pRAS1006	0.65-kb NdeI-XhoI insert from PCR product <i>adk</i> in pET-28a(+); Kan <sup>r</sup> , T7lac promoter
pET22b(+)	
pRAS1003	0.95-kb NdeI-XhoI insert from PCR product <i>prs</i> in pET-22b(+); Ap <sup>r</sup> , T7lac promoter
pCDFDuet-1	
pRAS1004	0.93-kb BamHI-HindIII insert from PCR product <i>rbsk</i> in pCDFDuet-1; Sm <sup>r</sup> , T7lac promoter

**Media and bacteriological techniques.** Each strain (Table IV-2) was maintained in LB medium containing either 50 µg/mL kanamycin (pRAS1001, pRAS1005, pRAS1006), 50 µg/mL streptomycin (pRAS1004) or 60 µg/mL ampicillin (pRAS1003) at 37 °C, with shaking at 225 rpm. Protein over expression was induced with 0.75 mM IPTG after reaching an OD<sub>600</sub> of 0.6. Following an additional four hours of incubation (3 hours for pRAS1001), the OD<sub>600</sub> of the cultures were measured and cell pellets were generated by centrifugation of 0.5 L aliquots (4 °C, 2800 x g, 30 min). Pelleted cell aliquots were stored at - 80 °C until immediately prior to use.

**Oligonucleotide primers and PCR conditions.** The proposed nucleotide analogue pathway biosynthetic genes encoding 8B3PRT, PRPP synthetase, ribokinase and cofactor recycling genes pyruvate kinase and adenylate kinase, were cloned into pET vectors using

standard cloning techniques. Primers, restriction enzymes and vectors are listed in Tables IV-2 and IV-3.

**Table IV-3: Primers used in cloning of nucleotide biosynthetic genes. Adapted from Scism, R. A. *et al.*<sup>1</sup>**

Gene	5' primer (5'-3') 3' primer (5'-3')	Restriction sites	Genbank Accession
<i>prs</i>	GAATTCCATATGTCGACTCAATATCTAAATTCTAATTTGAAAG GAATTCCTCGAGATTGAATAATACACTCACAGATTCTTCTTCG	NdeI-XhoI	NC_004722
<i>rbsk</i>	GAATTCGGATCCGCAAACGCAGGCAG GAATTC AAGCTTTCACCTCTGCCTGTCT	BamHI- HindIII	NC_012971
<i>pykF</i>	GAATTCATATGAAAAAGACCAAATTGTTTGCACC GAATTCGCGGCCGCTTACAGGACGTGAACAGATGCG	NdeI-NotI	NC_012971
<i>adk</i>	GAATTCATATGAACTTAATTTAATGGGGCTTCTCTGG GAATTCCTCGAGTTACGCTAAGCCTCCGATGAGAAC	NdeI-XhoI	NC_004722

**Transformation.** Transformations were performed using electroporation of electrocompetent cells generated by centrifugal harvest of BL21(DE3) cells grown in LB medium (10 g/L bacto-tryptone, 5 g/L bacto-yeast extract and 5 g/L NaCl) to log phase, followed by three washings in ice cold 10 % glycerol. Cells were transformed by providing a single electric pulse of 2.5 kV to 50  $\mu$ L electrocompetent cells mixed with 0.8 ng/  $\mu$ L plasmid DNA in a 2 mm electroporation cuvette using GenePulser Xcell (BioRad).

**Protein purification.** Frozen cell pellets were resuspended in 25 mL Buffer A (500 mM NaCl, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 30 mM imidazole, pH 7.4) prior to lysis by double passage through a French Pressure Cell, 20000 psi. Cell lysates were clarified by centrifugation (4 °C, 13,000 rpm, 20 min) and applied to previously stripped (500 mM NaCl, 20 mM Na<sub>2</sub>HPO<sub>4</sub>, 50 mM EDTA, pH 7.4) and recharged (100 mM NiSO<sub>4</sub>) Ni<sup>2+</sup>-affinity columns (HisTrap FF crude, 5 mL, GE

Healthcare). Proteins were eluted with a linear gradient of 100:0 A:B to 50:50 A:B (Buffer A = 500 mM NaCl, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 20 mM imidazole, pH 7.4, Buffer B = 500 mM NaCl, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 500 mM imidazole, pH 7.4) at 5 mL/min and concentrated by 10K molecular weight filter (Millipore). Proteins were exchanged into Reaction Buffer (12 mM Tris-HCl, pH = 8, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 10 mM MgCl<sub>2</sub>) via gel filtration (HiTrap Desalting, 5 mL, GE Healthcare). Concentrated enzyme was stored in 100 uL aliquots at -80°C. Final purified enzyme protein concentration was determined by bicinchoninic acid assay (Thermo Scientific) following the manufacturer's protocol against BSA standards, diluted with Reaction Buffer. Proteins were analyzed by 12 % polyacrylamide gel, loaded with 7 uL marker (Precision Plus Protein Kaleidoscope Standard, BioRad Inc), 5 uL of each enzyme, diluted to 2% of original concentration (1.06, 0.9, 1.82, 1.46, 1.29 µg for RK, PPS, 8B3, PK, AK respectively), and 690 µg centrifugally recovered CLEA. Gels were stained with Bio Safe Coomassie (BioRad, Inc).

### **General CLEA formation Protocols**

**CLEAs from Purified Enzymes.** CLEA aliquots were generated from purified enzymes by combining 50 µL of each of the five pathway enzyme preparations (determined by the bicinchoninic acid assay to be 18.2, 9.0, 10.6, 12.9, and 14.6 mg/mL for 8B3PRT, PPS, RK, AK, and PK respectively), and bovine serum albumin (10 mg/mL) in an Eppendorf tube. The six protein mixture was precipitated by adding an equal volume of ice cold saturated ammonium sulfate solution with rapid magnetic stirring at 4 °C. An additional 348 mg/mL of solid ammonium sulfate was added and allowed to equilibrate for 15 min to ensure complete protein precipitation. After 15 minutes, the aggregated enzyme/salt suspension was cross linked by adding 5 - 100 mM glutaraldehyde (Fluka Inc.) and incubated for two hours while stirring at 4 °C.

Cross-linked aggregate particles were readily collected by centrifugation (10,000 rpm) and washed three times by resuspension in ice cold Reaction Buffer.

**CLEAs from crude cell lysates.** CLEA aliquots were generated from unpurified cell lysates as follows. Aliquots of cell pellets from 50 mL cultures, grown and induced as previously described, were combined and resuspended in 10 mL Reaction Buffer prior to lysis via double passage through a French pressure cell. Aliquots of 500  $\mu$ L were removed for CLEA formation, to which 696 mg/mL ammonium sulfate was added. After 15 minutes, the aggregated enzyme suspension was cross linked by adding a final concentration of 75 mM glutaraldehyde and incubated for two hours while stirring at 4 °C. Cross-linked aggregate particles were readily collected by centrifugation (10,000 rpm) and washed three times by resuspension in ice cold Reaction Buffer.

### **General procedures for turnover measurements**

**Turnover as measured by HPLC.** Progress of CLEA reactions was monitored by HPLC assay using a C<sub>18</sub> column (Waters X-Bridge 4.6 x 150 mm) and an isocratic elution (1 mL/min) of 100 mM triethylammonium phosphate, pH 7 and 5% methanol. Peak identities were confirmed by injections of standard solutions (5 mM) of hypoxanthine, IMP, AMP, ADP and ATP.

**Turnovers as measured by hypoxanthine consumption.** To CLEA aliquots, prepared as described above, was added reaction buffer containing 12 mM Tris-HCl, pH = 8, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 10 mM MgCl<sub>2</sub>, 2.5 mM hypoxanthine, 5 mM D-ribose, 6 mM ATP and 25 mM phosphoenol pyruvate. After incubation with magnetic stirring at 23°C for 15 minutes, reactions were quenched by centrifugal removal of the catalyst followed by heating for 5 minutes at 95 °C. Remaining hypoxanthine was determined by adding 5 mL of reaction to 95 mL Assay Mix in triplicate in a 96 well flat bottom polystyrene plate (Nunc). Assay Mix was freshly prepared and



contained 12 mM Tris-HCl pH 7.0, 0.25 U/mL xanthine oxidase (from bovine milk) and 1 mM iodonitrotetrazolium chloride. Absorbance at 546 nm was measured after 15 minutes and hypoxanthine consumption was calculated the extinction coefficient and subtracting background absorbance of Assay Mix. For assays of non-aggregated enzymes, each of the five pathway enzyme preparations (50  $\mu$ L each) were combined with 2.5 mM hypoxanthine, 5 mM D-ribose, 6 mM ATP, and 25 mM phosphoenol pyruvate in Reaction Buffer for a final total volume of 300  $\mu$ L.

### **General Procedures for Preparation, Purification and Characterization of Nucleotide Analogues**

Nucleotide analogues were synthesized in 500  $\mu$ L reactions by stirring CLEA pellets with 10 mM nucleobase, 12 mM D (-) ribose, 6 mM ATP and 30 mM PEP for 2 hours at 23 °C. Reaction progress was monitored by WATERGATE NMR.

**Reaction Progress by WATERGATE NMR.** 60  $\mu$ L D<sub>2</sub>O was added to reactions to facilitate subsequent NMR quantification of conversion. Integrated values of the anomeric peak of nucleotides (5.83 - 6.94 ppm) were determined from spectra using using a WATERGATE-W5 water suppression pulse sequence. <sup>1</sup>H NMR experiments were acquired using a Bruker AV-III spectrometer operating at 500.13 MHz. Water suppression was performed with the W5-watergate sequence (zggpw-5, Bruker pulse sequence nomenclature) using a double gradient echo.<sup>84</sup> Prior to acquiring water suppression experiments, the proton 90° pulse (p1) was accurately calibrated using a nutation experiment as was the distance of the next null used for the binomial water suppression delay (d19). The data was processed using sinusoidal function to remove the residual water signal.

**Nucleotide Purification.** Analogues were purified by DEAE-sepharose anion exchange resin (80 mL column volume, 2.5 cm × 20 cm) and eluted with an ammonium carbonate gradient (0 to 100% 500 mM ammonium carbonate) at 5 mL/min over 5 column volumes. Fractions (14 mL) were measured by UV absorbance at 254 nm. Fractions of individual peaks were pooled and evaporated *in vacuo*. Residue was then redissolved in D<sub>2</sub>O and evaporated prior to <sup>1</sup>H NMR analysis.

**Percent conversion.** Sample residue was redissolved in 1 mL D<sub>2</sub>O with internal DMSO standard (0.833 mM). Percent conversion calculated by dividing the integration of the anomeric nucleotide peak (5.83 - 6.94 ppm) by the integrated value of the DMSO singlet. Experimental conditions included 32K data points, 13 ppm sweep width, a recycle delay of 10 seconds and 32 scans. <sup>1</sup>H NMR spectra of prepared nucleotides was identical to previously published synthesized nucleotides.

**6-choropurine 5'-phosphate.** <sup>1</sup>H-NMR (400 MHz, D<sub>2</sub>O): δ 8.65 (s, 1H), 8.49 (s, 1H), 5.97 (d, J=5.0 Hz, 1H), 4.49 (t, 1H), 4.20 (t, 1H), 4.08 (m, 1H), 3.72 (m, 2H).

**Mercaptopurine 5'-phosphate.** <sup>1</sup>H-NMR (400 MHz, D<sub>2</sub>O): δ 8.28 (s, 1H), 8.00 (s, 1H), 5.75 (d, J=3.4 Hz, 1H), 4.37 (m, 1H), 4.08 (m, 1H), 3.96 (m, 1H), 3.61 (m, 2H).

**Allopurine 5'-phosphate.** <sup>1</sup>H-NMR (400 MHz, D<sub>2</sub>O): δ 7.95 (s, 1H), 7.94 (s, 1H), 5.98 (d, J=5.0 Hz, 1H), 4.53 (t, 1H), 4.20 (t, 1H), 3.98 (m, 1H), 3.60 (m, 2H).

**Purine 5'-phosphate.** <sup>1</sup>H-NMR (400 MHz, D<sub>2</sub>O): δ 8.96 (s, 1H), 8.76 (s, 1H), 8.69 (s, 1H), 6.08 (d, J=5.0 Hz, 1H), 4.60 (t, 1H), 4.29 (t, 1H), 4.16 (m, 1H), 3.80 (m, 2H). HRMS (ESI-) calcd for C<sub>10</sub>H<sub>12</sub>N<sub>4</sub>O<sub>7</sub>P<sup>-</sup> = 331.0449 m/z. Observed mass = 331.0462 m/z.

## References

1. Scism, R. A.; Bachmann, B. O., Five-Component Cascade Synthesis of Nucleotide Analogues in an Engineered Self-Immobilized Enzyme Aggregate. *Chembiochem* **2010**, *11*, (1), 67-70.
2. Sigrell, J. A.; Cameron, A. D.; Jones, T. A.; Mowbray, S. L., Purification, characterization, and crystallization of Escherichia coli ribokinase. *Protein Science* **1997**, *6*, (11), 2474-2476.
3. Sigrell, J. A.; Cameron, A. D.; Mowbray, S. L., Induced fit on sugar binding activates ribokinase. *Journal of Molecular Biology* **1999**, *290*, (5), 1009-1018.
4. Sigrell, J. A.; Cameron, A. D.; Jones, T. A.; Mowbray, S. L., Structure of Escherichia coli ribokinase in complex with ribose and dinucleotide determined to 1.8 angstrom resolution: insights into a new family of kinase structures. *Structure* **1998**, *6*, (2), 183-193.
5. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, *25*, (13), 1605-1612.
6. Maj, M. C.; Gupta, R. S., The effect of inorganic phosphate on the activity of bacterial ribokinase. *Journal of Protein Chemistry* **2001**, *20*, (2), 139-144.
7. Andersson, C. E.; Mowbray, S. L., Activation of ribokinase by monovalent cations. *Journal of Molecular Biology* **2002**, *315*, (3), 409-419.
8. Chuvikovskiy, D. V.; Esipov, R. S.; Skoblov, Y. S.; Chupova, L. A.; Muravyova, T. I.; Miroshnikov, A. I.; Lapinjoki, S.; Mikhailopulo, I. A., Ribokinase from E-coli: Expression, purification, and substrate specificity. *Bioorganic & Medicinal Chemistry* **2006**, *14*, (18), 6327-6332.
9. Willemoes, M.; Hove-Jensen, B.; Larsen, S., Steady state kinetic model for the binding of substrates and allosteric effectors to Escherichia coli phosphoribosyl-diphosphate synthase. *Journal of Biological Chemistry* **2000**, *275*, (45), 35408-35412.
10. Nosal, J. M.; Switzer, R. L.; Becker, M. A., Overexpression, Purification, and Characterization of Recombinant Human 5-Phosphoribosyl-1-Pyrophosphate Synthetase Isozyme-I and Isozyme-II. *Journal of Biological Chemistry* **1993**, *268*, (14), 10168-10175.
11. Hovejensen, B.; Harlow, K. W.; King, C. J.; Switzer, R. L., Phosphoribosylpyrophosphate Synthetase of Escherichia-Coli - Properties of the Purified Enzyme and Primary Structure of the Prs Gene. *Journal of Biological Chemistry* **1986**, *261*, (15), 6765-6771.
12. Eriksen, T. A.; Kadziola, A.; Bentsen, A. K.; Harlow, K. W.; Larsen, S., Structural basis for the function of Bacillus subtilis phosphoribosylpyrophosphate synthetase. *Nature Structural Biology* **2000**, *7*, (4), 303-308.

13. Willemoes, M.; HoveJensen, B., Binding of divalent magnesium by Escherichia coli phosphoribosyl diphosphate synthetase. *Biochemistry* **1997**, 36, (16), 5078-5083.
14. Krath, B. N.; Hove-Jensen, B., Class II recombinant phosphoribosyl diphosphate synthase from spinach - Phosphate independence and diphosphoryl donor specificity. *Journal of Biological Chemistry* **2001**, 276, (21), 17851-17856.
15. Balzarini, J.; Declercq, E., 5-Phosphoribosyl 1-Pyrophosphate Synthetase Converts the Acyclic Nucleoside Phosphonates 9-(3-Hydroxy-2-Phosphonylmethoxypropyl)Adenine and 9-(2-Phosphonylmethoxyethyl)Adenine Directly to Their Antivirally Active Diphosphate Derivatives. *Journal of Biological Chemistry* **1991**, 266, (14), 8686-8689.
16. Nave, J. F.; Eschbach, A.; Wolffkugel, D.; Halazy, S.; Balzarini, J., Enzymatic Phosphorylation and Pyrophosphorylation of 2',3'-Dideoxyadenosine-5'-Monophosphate, a Key Metabolite in the Pathway for Activation of the Anti-Hiv (Human-Immunodeficiency-Virus) Agent 2',3'-Dideoxyinosine. *Biochemical Pharmacology* **1994**, 48, (6), 1105-1112.
17. Terreni, M.; Ubiali, D.; Bavaro, T.; Pregnolato, M.; Fernandez-Lafuente, R.; Guisan, J. M., Enzymatic synthesis of cephalosporins. The immobilized acylase from *Arthrobacter viscosus*: A new useful biocatalyst. *Applied Microbiology and Biotechnology* **2007**, 77, (3), 579-587.
18. Kaar, J. L.; Oh, H.; Russell, A. J.; Federspiel, W. J., Towards improved artificial lungs through biocatalysis. *Biomaterials* **2007**, 28, 3131-3139.
19. Taylor, R. F., *Protein Immobilization Fundamentals and Applications*. Marcel Dekker, Inc.: Cambridge, 1991.
20. Bickerstaff, G. F., *Immobilization of Enzymes and Cells*. Humana Press: Totowa, 1997.
21. Hilterhaus, L.; Minow, B.; Muller, J.; Berheide, M.; Quitmann, H.; Katzer, M.; Thum, O.; Antranikian, G.; Zeng, A. P.; Liese, A., Practical application of different enzymes immobilized on sepabeads. *Bioprocess and Biosystems Engineering* **2008**, 31, (3), 163-171.
22. Chen, B.; Hu, J.; Miller, E. M.; Xie, W. C.; Cai, M. M.; Gross, R. A., *Candida antarctica* lipase B chemically immobilized on epoxy-activated micro- and nanobeads: Catalysts for polyester synthesis. *Biomacromolecules* **2008**, 9, (2), 463-471.
23. Bolivar, J. M.; Cava, F.; Mateo, C.; Rocha-Martin, J.; Guisan, J. M.; Berenguer, J.; Fernandez-Lafuente, R., Immobilization-stabilization of a new recombinant glutamate dehydrogenase from *Thermus thermophilus*. *Applied Microbiology and Biotechnology* **2008**, 80, (1), 49-58.

24. Pollak, A.; Blumenfeld, H.; Wax, M.; Baughn, R. L.; Whitesides, G. M., Enzyme Immobilization by Condensation Copolymerization into Cross-Linked Polyacrylamide Gels. *Journal of the American Chemical Society* **1980**, 102, (20), 6324-6336.
25. Groger, H.; Capan, E.; Barthuber, A.; Vorlop, K. D., Asymmetric synthesis of an (R)-cyanohydrin using enzymes entrapped in lens-shaped gels. *Organic Letters* **2001**, 3, (13), 1969-1972.
26. Zoumpanioti, M.; Parmaklis, P.; de Maria, P. D.; Stamatis, H.; Sinisterra, J. V.; Xenakis, A., Esterification reactions catalyzed by lipases immobilized in organogels: effect of temperature and substrate diffusion. *Biotechnology Letters* **2008**, 30, (9), 1627-1631.
27. Miller, S. A.; Hong, E. D.; Wright, D., Rapid and efficient enzyme encapsulation in a dendrimer silica nanocomposite. *Macromolecular Bioscience* **2006**, 6, (10), 839-845.
28. Pencreach, G.; Leullier, M.; Baratti, J. C., Properties of free and immobilized lipase from *Pseudomonas cepacia*. *Biotechnology and Bioengineering* **1997**, 56, (2), 181-189.
29. Chen, B.; Miller, M. E.; Gross, R. A., Effects of porous polystyrene resin parameters on *Candida antarctica* Lipase B adsorption, distribution, and polyester synthesis activity. *Langmuir* **2007**, 23, (11), 6467-6474.
30. Nahalka, J.; Liu, Z. Y.; Chen, X.; Wang, P. G., Superbeads: Immobilization in "sweet" chemistry. *Chemistry-a European Journal* **2003**, 9, (2), 373-377.
31. Auge, C.; Malleron, A.; Tahrat, H.; Marc, A.; Goergen, J. L.; Cerutti, M.; Steelant, W. F. A.; Delannoy, P.; Lubineau, A., Outstanding stability of immobilized recombinant alpha(1 -> 3/4)-fucosyltransferases exploited in the synthesis of Lewis a and Lewis x trisaccharides. *Chemical Communications* **2000**, (20), 2017-2018.
32. Chen, X.; Fang, J. W.; Zhang, J. B.; Liu, Z. Y.; Shao, J.; Kowal, P.; Andreana, P.; Wang, P. G., Sugar nucleotide regeneration beads (superbeads): A versatile tool for the practical synthesis of oligosaccharides. *Journal of the American Chemical Society* **2001**, 123, (9), 2081-2082.
33. Luo, Y. X.; Yuan, Z. Y.; Luo, G. M.; Zhao, F. K., Expression of secreted His-tagged S-adenosylmethionine synthetase in the methylotrophic yeast *Pichia pastoris* and its characterization, one-step purification, and immobilization. *Biotechnology Progress* **2008**, 24, (1), 214-220.
34. Wine, Y.; Cohen-Hadar, N.; Freeman, A.; Frolow, F., Elucidation of the mechanism and end products of glutaraldehyde crosslinking reaction by X-ray structure analysis. *Biotechnology and Bioengineering* **2007**, 98, (3), 711-718.
35. Dainiak, M. B.; Allan, I. U.; Savina, I. N.; Cornelio, L.; James, E. S.; James, S. L.; Mikhailovsky, S. V.; Jungvid, H.; Galaev, I. Y., Gelatin-fibrinogen cryogel dermal matrices for wound repair: Preparation, optimisation and in vitro study *Biomaterials* **2010**, 31, (1), 67-76.

36. Lee, S.; Jin, L. H.; Kim, J. H.; Han, S. O.; Na, H. B.; Hyeon, T.; Koo, Y.; Kim, J.; Lee, J., B-Glucosidase coating on polymer nanofibers for improved cellulosic ethanol production. *Bioprocess Biosyst Eng* **2009**, 33, 141-147.
37. Wilson, L.; Illanes, A.; Pessela, B. C. C.; Abian, O.; Fernandez-Lafuente, R.; Guisan, J. M., Encapsulation of crosslinked penicillin G acylase aggregates in lentikats: Evaluation of a novel biocatalyst in organic media. *Biotechnology and Bioengineering* **2004**, 86, (5), 558-562.
38. Habeeb, A. F. S., Preparation of Enzymically Active Water-Insoluble Derivatives of Trypsin. *Archives of Biochemistry and Biophysics* **1967**, 119, (1-3), 264.
39. Khan, S. S.; Siddiqi, A. M., Studies on Chemically Aggregated Pepsin Using Glutaraldehyde. *Biotechnology and Bioengineering* **1985**, 27, (4), 415-419.
40. Khare, S. K.; Gupta, M. N., An Active Insoluble Aggregate of Escherichia-Coli Beta-Galactosidase. *Biotechnology and Bioengineering* **1990**, 35, (1), 94-98.
41. Cao, L. Q.; van Rantwijk, F.; Sheldon, R. A., Cross-linked enzyme aggregates: A simple and effective method for the immobilization of penicillin acylase. *Organic Letters* **2000**, 2, (10), 1361-1364.
42. Lopez-Gallego, F.; Betancor, L.; Hidalgo, A.; Alonso, N.; Fernandez-Lafuente, R.; Guisan, J. M., Co-aggregation of enzymes and polyethyleneimine: A simple method to prepare stable and immobilized derivatives of glutaryl acylase. *Biomacromolecules* **2005**, 6, (4), 1839-1842.
43. Fazary, A. E.; Ismadji, S.; Ju, Y. H., Biochemical studies on native and cross-linked aggregates of *Aspergillus awamori* feruloyl esterase. *International Journal of Biological Macromolecules* **2009**, 44, (3), 240-248.
44. Romero, O.; Vergara, J.; Fernandez-Lafuente, R.; Guisan, J. M.; Illanes, A.; Wilson, L., Simple Strategy of Reactivation of a Partially Inactivated Penicillin G Acylase Biocatalyst in Organic Solvent and its Impact on the Synthesis of beta-Lactam Antibiotics. *Biotechnology and Bioengineering* **2009**, 103, (3), 472-479.
45. Cabana, H.; Jones, J. P.; Agathos, S. N., Preparation and characterization of cross-linked laccase aggregates and their application to the elimination of endocrine disrupting chemicals. *Journal of Biotechnology* **2007**, 132, (1), 23-31.
46. Aytar, B. S.; Bakir, U., Preparation of cross-linked tyrosinase aggregates. *Process Biochemistry* **2008**, 43, (2), 125-131.
47. Sangeetha, K.; Abraham, T. E., Preparation and characterization of cross-linked enzyme aggregates (CLEA) of subtilisin for controlled release applications. *International Journal of Biological Macromolecules* **2008**, 43, (3), 314-319.

48. Wilson, L.; Illanes, A.; Abian, O.; Pessela, B. C. C.; Fernandez-Lafuente, R.; Guisan, J. M., Co-aggregation of penicillin G acylase and polyionic polymers: An easy methodology to prepare enzyme biocatalysts stable in organic media. *Biomacromolecules* **2004**, 5, (3), 852-857.
49. Scism, A. J.; Bemiller, J. N.; Caskey, A. L., Determination of 2,4-Dihydroxy-1,4(2h)-Benzoxazin-3-One Glucosides in Corn (Zea-Mays-L). *Analytical Biochemistry* **1974**, 58, (1), 1-13.
50. Hennig, M.; Scott, L. G.; Sperling, E.; Bermel, W.; Williamson, J. R., Synthesis of 5-fluoropyrimidine Nucleotides as sensitive NMR probes of RNA structure. *Journal of the American Chemical Society* **2007**, 129, (48), 14911-14921.
51. Gilles, A. M.; Cristea, I.; Palibroda, N.; Hilden, I.; Jensen, K. F.; Sarfati, R. S.; Namane, A.; UghettoMonfrin, J.; Barzu, O., Chemienzymatic synthesis of uridine nucleotides labeled with [N-15] and [C-13]. *Analytical Biochemistry* **1995**, 232, (2), 197-203.
52. Parkin, D. W.; Leung, H. B.; Schramm, V. L., Synthesis of Nucleotides with Specific Radiolabels in Ribose - Primary C-14 and Secondary H-3 Kinetic Isotope Effects on Acid-Catalyzed Glycosidic Bond Hydrolysis of Amp, Damp, and Inosine. *Journal of Biological Chemistry* **1984**, 259, (15), 9411-9417.
53. Bouhss, A.; Sakamoto, H.; Palibroda, N.; Chiriach, M.; Sarfati, R.; Smith, J. M.; Craescu, C. T.; Barzu, O., Enzymatic-Synthesis of Guanine-Nucleotides Labeled with N-15 at the 2-Amino Group of the Purine Ring. *Analytical Biochemistry* **1995**, 225, (1), 18-23.
54. Rawls, J. M., Enzymatic-Synthesis of Orotidine-C-14 5'-Monophosphate and Its Use in Assay of Orotate Phosphoribosyltransferase and Orotidylate Decarboxylase. *Analytical Biochemistry* **1978**, 86, (1), 107-117.
55. Way, J.; Parks, R. E., Enzymatic Synthesis of 5'-phosphate Nucleotides of Purine Analogues. *Journal of Biological Chemistry* **1958**, 231, 467-480.
56. Snetkova, E. V.; Kaminskii, Y. L.; Akulov, G. P.; Nikolaeva, Z. K., Enzymatic-Synthesis of Multiply Tritium-Labeled Nucleosides and Nucleotides from Nitrogen Bases and Ribose. *Khimiya Prirodnikh Soedinenii* **1988**, (2), 258-264.
57. Scott, L. G.; Tolbert, T. J.; Williamson, J. R., Preparation of specifically H-2- and C-13-labeled ribonucleotides. *Rna-Ligand Interactions Pt A* **2000**, 317, 18-38.
58. Schultheisz, H. L.; Szymczyna, B. R.; Scott, L. G.; Williamson, J. R., Pathway engineered enzymatic de novo purine nucleotide synthesis. *Acs Chemical Biology* **2008**, 3, (8), 499-511.
59. Tolbert, T. J.; Williamson, J. R., Preparation of specifically deuterated and C-13-labeled RNA for NMR studies using enzymatic synthesis. *Journal of the American Chemical Society* **1997**, 119, (50), 12100-12108.

60. Tolbert, T. J.; Williamson, J. R., Preparation of specifically deuterated RNA for NMR studies using a combination of chemical and enzymatic synthesis. *Journal of the American Chemical Society* **1996**, 118, (34), 7929-7940.
61. Wu, W. D.; Bergstrom, D. E.; Davisson, V. J., A combination chemical and enzymatic approach for the preparation of azole carboxamide nucleoside triphosphate. *Journal of Organic Chemistry* **2003**, 68, (10), 3860-3865.
62. Da Costa, C. P.; Fedor, M. J.; Scott, L. G., 8-Azaguanine reporter of purine ionization states in structured RNAs. *Journal of the American Chemical Society* **2007**, 129, (11), 3426-3432.
63. Hoffman, D. W.; Holland, J. A., Preparation of C-13 Labeled Ribonucleotides Using Acetate as an Isotope Source. *Nucleic Acids Research* **1995**, 23, (16), 3361-3362.
64. Michnicka, M. J.; Harper, J. W.; King, G. C., Selective Isotopic Enrichment of Synthetic Rna - Application to the Hiv-1 Tar Element. *Biochemistry* **1993**, 32, (2), 395-400.
65. Nikonowicz, E. P.; Sirt, A.; Legault, P.; Jucker, F. M.; Baer, L. M.; Pardi, A., Preparation of C-13 and N-15 Labeled Rnas for Heteronuclear Multidimensional Nmr-Studies. *Nucleic Acids Research* **1992**, 20, (17), 4507-4513.
66. Batey, R. T.; Inada, M.; Kujawinski, E.; Puglisi, J. D.; Williamson, J. R., Preparation of Isotopically Labeled Ribonucleotides for Multidimensional Nmr-Spectroscopy of Rna. *Nucleic Acids Research* **1992**, 20, (17), 4515-4523.
67. Batey, R. T.; Battiste, J. L.; Williamson, J. R., Preparation of isotopically enriched RNAs for heteronuclear NMR. *Nuclear Magnetic Resonance and Nucleic Acids* **1995**, 261, 300-322.
68. Grammatikos, S. I.; Tobien, K.; Noe, G.; Werner, R. G., Monitoring of intracellular ribonucleotide pools is a powerful tool in the development and characterization of mammalian cell culture processes. *Biotechnology and Bioengineering* **1999**, 64, (3), 357-367.
69. Pruvost, A.; Becher, F.; Bardouille, P.; Guerrero, C.; Creminon, C.; Delfraissy, J. F.; Goujard, C.; Grassi, J.; Benech, H., Direct determination of phosphorylated intracellular anabolites of stavudine (d4T) by liquid chromatography tandem mass spectrometry. *Rapid Communications in Mass Spectrometry* **2001**, 15, (16), 1401-1408.
70. Aussenac, J.; Chassagne, D.; Claparols, C.; Charpentier, M.; Duteurtre, B.; Feuillat, M.; Charpentier, C., Purification method for the isolation of monophosphate nucleotides from Champagne wine and their identification by mass spectrometry. *Journal of Chromatography A* **2001**, 907, (1-2), 155-164.
71. Ryll, T.; Wagner, R., Improved Ion-Pair High-Performance Liquid-Chromatographic Method for the Quantification of a Wide Variety of Nucleotides and Sugar Nucleotides in Animal-Cells. *Journal of Chromatography-Biomedical Applications* **1991**, 570, (1), 77-88.



72. Ryll, T.; Wagner, R., Intracellular Ribonucleotide Pools as a Tool for Monitoring the Physiological-State of Invitro Cultivated Mammalian-Cells during Production Processes. *Biotechnology and Bioengineering* **1992**, 40, (8), 934-946.
73. Xing, J. S.; Apedo, A.; Tymiak, A.; Zhao, N., Liquid chromatographic analysis of nucleosides and their mono-, di- and triphosphates using porous graphitic carbon stationary phase coupled with electrospray mass spectrometry. *Rapid Communications in Mass Spectrometry* **2004**, 18, (14), 1599-1606.
74. Dennison, C., *A guide to protein isolation*. 2nd ed.; Kluwer Academic Publishers: Dordrecht ; Boston, 2003; p xvii, 249 p.
75. Hatti-Kaul, R.; Mattiasson, B., *Isolation and purification of proteins*. Marcel Dekker: New York, 2003; p xix, 652 p.
76. Roe, S., *Protein purification techniques : a practical approach*. 2nd ed.; Oxford University Press: New York, 2001; p xvi, 262 p.
77. Himmel, M. E.; Georgiou, G.; American Chemical Society. Division of Biochemical Technology.; American Chemical Society. Meeting, *Biocatalyst design for stability and specificity*. American Chemical Society: Washington, DC, 1993; p xiii, 335 p.
78. Eike, J. H.; Palmer, A. F., Effect of NaBH<sub>4</sub> concentration and reaction time on physical properties of glutaraldehyde-polymerized hemoglobin. *Biotechnology Progress* **2004**, 20, (3), 946-952.
79. Miles, C. A.; Avery, N. C.; Rodin, V. V.; Bailey, A. J., The increase in denaturation temperature following cross-linking of collagen is caused by dehydration of the fibres. *Journal of Molecular Biology* **2005**, 346, (2), 551-556.
80. Magalhaes, J. M. C. S.; Machado, A. A. S. C., Urea potentiometric biosensor based on urease immobilized on chitosan membranes. *Talanta* **1998**, 47, (1), 183-191.
81. Farris, S.; Song, J. H.; Huang, Q. R., Alternative Reaction Mechanism for the Cross-Linking of Gelatin with Glutaraldehyde. *Journal of Agricultural and Food Chemistry* **2010**, 58, (2), 998-1003.
82. Richards, F. M.; Knowles, J. R., Glutaraldehyde as a Protein Cross-Linking Reagent. *Journal of Molecular Biology* **1968**, 37, (1), 231.
83. Shah, S.; Sharma, A.; Gupta, M. N., Preparation of cross-linked enzyme aggregates by using bovine serum albumin as a proteic feeder. *Analytical Biochemistry* **2006**, 351, (2), 207-213.
84. Liu, M. L.; Mao, X. A.; Ye, C. H.; Huang, H.; Nicholson, J. K.; Lindon, J. C., Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy. *Journal of Magnetic Resonance* **1998**, 132, (1), 125-129.

## CHAPTER V

### DISSERTATION SUMMARY AND FUTURE DIRECTION

#### Synopsis

Pharmaceutically active nucleotide analogues are indispensable in protecting against viral pathogens. Preparation of these molecules presents challenges in controlling regio- and stereochemistry, isolation of unstable intermediates, and requires multiple protecting group manipulations. Biosynthesis of these analogues provides solutions to these synthetic issues, but suffers from the stringent substrate requirements of enzymes. Directed evolution presents a possible route to allow these enzymes to accept alternative substrates. The work described in this dissertation provides a significant improvement over current methods in nucleotide analogue synthesis by overcoming the above obstacles encountered with chemical synthesis.

Directed evolution was used to relax the active site specificity of hypoxanthine phosphoribosyl transferase (HPRT) for increased production of the nucleotide analogue, ribavirin monophosphate (RMP). A novel *in vivo* screening method for negative selection of variant HPRT enzymes was introduced, in which biosynthesized RMP inhibited intracellular inosine monophosphate dehydrogenase (IMPDH). Resulting inhibition prevented host cell growth in minimal media, allowing detection of improved mutants. The best mutant, 8B3PRT, showed increased expression levels of a unique enzyme with altered substrate specificity, and consequentially increased RMP production. An analytical method for intracellular nucleotide quantification indicated RMP is twice as concentrated in the cells with 8B3PRT compared to cells with wild type HPRT.

The evolved enzyme, 8B3PRT was assessed as a general nucleotide analogue biocatalyst. A high throughput assay for nucleobase turnover was developed, based on pyrophosphate release, revealing 8B3PRT to have uniquely relaxed substrate specificity. This facilitated the biosynthesis of several nucleotide analogues from the single enzyme, in a one-step reaction. Product was then purified in a single step anion exchange method.

A multi-step pathway for biosynthesis of nucleotide analogues from ribose was engineered by cloning several enzymes into an *E. coli* host. Additional enzymes for cofactor recycling were included to allow catalytic use of ATP, and to relieve ADP byproduct inhibition. Self-immobilization of the five enzymes in a glutaraldehyde cross-linked enzyme aggregate (CLEA) created a biocatalyst with greatly increased stability. Furthermore, the CLEA protein particle was removable from stirred reactions, allowing subsequent reuse over several iterations. Aggregates created from non-purified, crude cell lysate showed nearly equal activity as purified enzymes, suggesting this method is a viable alternative to whole cell biosynthesis.

### **Significance**

Biosynthesis of nucleotides with the engineered, immobilized pathway, described in this dissertation, alleviates the problems associated with traditional chemical synthesis. Stereo-specific and regiospecific addition of the nucleobase onto ribose is superior to unwieldy chemical methods. In the enzymatic pathway, multi-step syntheses are reduced to a single cascade process, without the need for isolation of intermediates. Additional circumvention of protecting group additions and removals greatly reduces the number of synthetic steps needed. Together, these improvements increase product yield and reduce chemical waste production.

Reduction of chemical wastes, generated in pharmaceutical syntheses, reduces time and costs associated with purchase of needed solvents and cost of resulting waste disposal. The

pharmaceutical industry alone generates 25 kg to over 100 kg of chemical waste for every 1 kg of product made.<sup>1</sup> The cumulative environmental impact of this chemical waste is of growing concern, as well as the accompanying geopolitical issues stemming from dependence on petroleum based solvents.<sup>2</sup> Bioengineering enzymatic alternatives, like the CLEA nucleotide pathway presented here, is one potential solution in reducing or eliminating this waste.

The *in vivo* screening method described in this dissertation serves as a proof of principle, with potential for a wider application in directed evolution of other drug synthesizing enzymes. The federal drug administration reports there are 317 marketed drugs that act through enzyme inhibition.<sup>3</sup> It is hypothesized that cloning the target enzyme of a suitable pharmaceutical candidate into *E. coli*, along with the necessary enzymes for biosynthesis, could provide a tailored system for similar *in vivo* screening, or negative selection. Whereas existing selection methods depend on linking cell survival to improved catalysis, this negative selection would use growth cessation due to target inhibition, a much simpler task than engineering product-dependent survival.

Formation of the stable CLEA biocatalyst using crude cell lysate opens the opportunity for application as an alternative to whole cell biosynthesis. A non-purified, CLEA immobilized, lysed cell could be used for chemical biosynthesis in circumstances where toxicity of product, substrate, or an intermediate is problematic, or when transport through the cell membrane is a issue.

For example, biosynthesis of catechol from glucose in *E. coli* has a limited yield of 5% due to product toxicity. The biosynthesis of this compound has been made possible by introduction of *Klebsiella pneumoniae aroA* and *aroY* genes into a 3-dehydroshikimate synthesizing *E. coli*.<sup>4</sup> While this green chemistry approach avoids the use of petroleum based, carcinogenic, volatile solvent benzene, the low yield prevents economically viable industrial use.

A maximum yield of 43% was attained only after re-engineering the pathway to the penultimate product, followed by extraction and chemical decarboxylation.<sup>5</sup> It is conceivable that CLEA immobilization of the original construct, using cell lysate, would create a whole cell pathway, immune to catechol toxicity.

## **Future Directions**

### **Directed Evolution**

#### ***Mutation Methods***

While nucleotide analogue biosynthesis with the evolved biocatalyst, 8B3PRT, is a significant improvement over chemical synthesis, the modest change in catalytic activity over wild-type HPRT leaves room for improvement. Other methods of introducing diversity are available, and could possibly lead to greater catalytic improvement. Of all of the mutants that were created by epPCR screened, that with greatest improvement contained a mutated residue in the active site. Perhaps direct mutation of the active site through CASTing or GSSM, described in Chapter I, could greatly improve the enzyme.

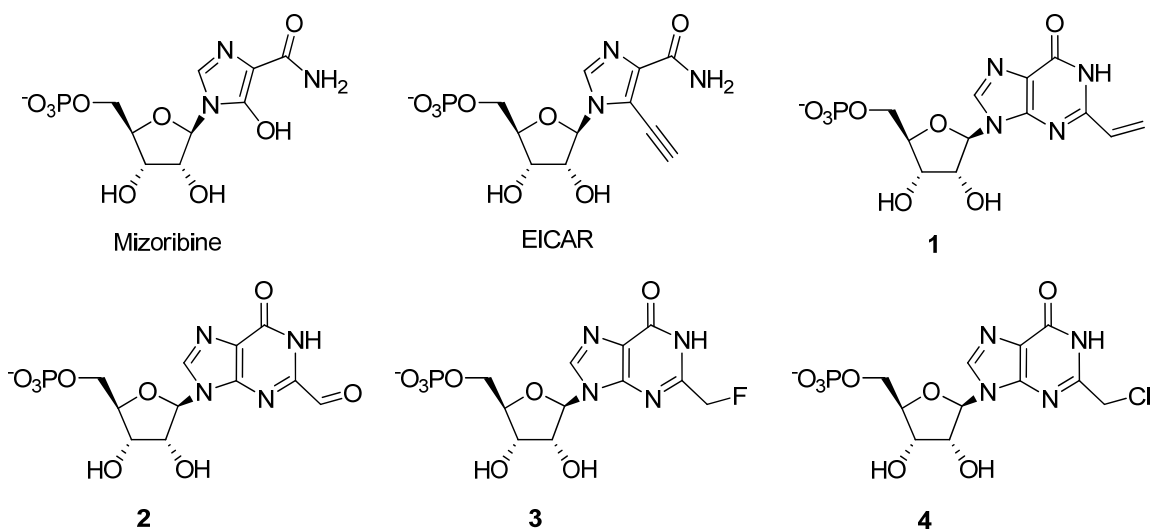
A library of mutants created using a recombination method, such as DNA shuffling, or one of its many variations, could also increase the chance of improved catalysis. As described in Chapter I, DNA shuffling creates multiple crossovers in the target DNA, allowing combination of beneficial mutations. The sum of the catalytic improvement of the mutations could be additive or synergistic, and without cross-over, the combined improvement may be lost. The source of genes to be shuffled could be a mixture of random mutants from mutagenic PCR, and/or other

PRTs such as APRT, OPRT, XGPRT, cloned into vector plasmids. The combination of these variants would almost certainly demonstrate unique substrate specificity.

Recently, phenotypic variants of *E. coli*, known as persisters, have been described. These persisters are a distinct physiological state that constitute approximately 1% of cells in stationary phase.<sup>6</sup> Temporary existence in a stationary phase has allowed these cells to survive antibiotic exposure, due to the pause of cellular metabolism. Since increased production of RMP decreases rate of growth, that high RMP production could induce a pseudo persister-like state. Since antibiotics such as ampicillin effect only dividing cells, we hypothesize that a mutant library could be enriched with RMP producers when cultured in the presence of both TCA and ampicillin. Preliminary experiments performed in our lab support this hypothesis.

### ***Additional Substrates***

Using any of the above described methods for mutation, additional substrates could be used to drive evolution. Chapter III presented several purine base analogues that are accepted by the 8B3PRT enzyme. The nucleotide product of one of these nucleobases, 6-chloroinosine monophosphate, happens to be an IMPDH inhibitor, ( $K_i = 78 \mu\text{M}$ ) acting upon the same active site as RMP.<sup>7</sup> Other IMPDH inhibiting nucleotide analogues could be screened in a similar manner. Possibilities include the monophosphate nucleotides formed from nucleobases of 1- $\beta$ -D-ribofuranosylimidazole-4-carboxamide (EICAR), ( $K_i = 16 \mu\text{M}$ ),<sup>8</sup> or known IMPDH inhibitors Mizoribine, and 2-substituted inosine 5'-monophosphates, as shown in Figure V-1.<sup>9</sup>



**Figure V-1: Nucleotide analogue inhibitors of inosine monophosphate dehydrogenase, adapted from Wang, G. Y. *et al.*<sup>9</sup>**

XGPRT from *E. coli* has been used to synthesize 8-azaguanine monophosphate from the corresponding base, 8-azaguanine, without mutation. Since 8B3PRT showed greatly increased acceptance of guanine over wild type HPRT, it is highly likely that 8-azaguanine would show some basal activity as a starting point for directed evolution. Phosphorylated product is not an IMPDH inhibitor, so a unique screening method would need to be developed. Highly catalytically active cells could be distinguished through fluorescent activated cell sorting (FACS), as accumulation of the fluorescent 8-azaguanine monophosphate product builds. This substrate and accompanying screening method could be used in combination with any of the above proposed mutation methods, including DNA shuffling of 8B3PRT with XGPRT and an epPCR library.

Addition of deoxyribose nucleotide analogues to the established ribose nucleotides would double the scope of this biocatalytic pathway. Ribokinase is reported to accept 2'-deoxy-D-ribose, D-arabinose, D-xylose, and D-fructose in addition to its natural substrate, D-ribose.<sup>10</sup> Unfortunately, while phosphoribosyl pyrophosphate synthetase allowed dATP in the ATP

binding pocket, the enzyme has no activity with deoxyribose-5-phosphate, or any other substitutions of r-5-p.<sup>11</sup> Activity of the final enzyme, HPRT, with deoxy-PRPP is unknown. Evolution of this pathway for creation of deoxyribonucleotide analogues would provide considerably more analogue products, but would require significant genetic shift and creation of novel screening methods. Additionally, synthesis of a deoxy-PRPP substrate for evolution of HPRT could prove to be a troubling impediment.

The ultimate goal for a nucleotide analogue biosynthetic pathway is to create final product from primary metabolites. An organism constructed to make analogues from glucose or ribose, without added nucleobase would reduce production cost and waste, and eliminate the need for a nucleobase solvent. The greatest limitation in the CLEA immobilized pathway is the solubility of the bases. Piecing together an enzymatic pathway for biosynthesis of the nucleobase is feasible, yet daunting task, requiring knowledge of a wide scope of possible enzyme candidates with appropriate flexibility needed for unusual substrates. However, in the case of an end product with an established screening method, such as for RMP, a retrosynthetic approach could allow evolution of preceding steps. For example, a library of asparagine synthetase mutants could be screened for conversion of 1,2,4-triazole-3-carboxylate to 1,2,3-triazole-3-carboxamide. Bioretrosynthesis screening of asparagine synthetase would be performed with co-expressed 8B3PRT using the original *in vivo* IMPDH inhibition screen. Asparagine synthase mutants catalyze the transformation from acid to amide, which is then condensed upon the ribose sugar by 8B3PRT.

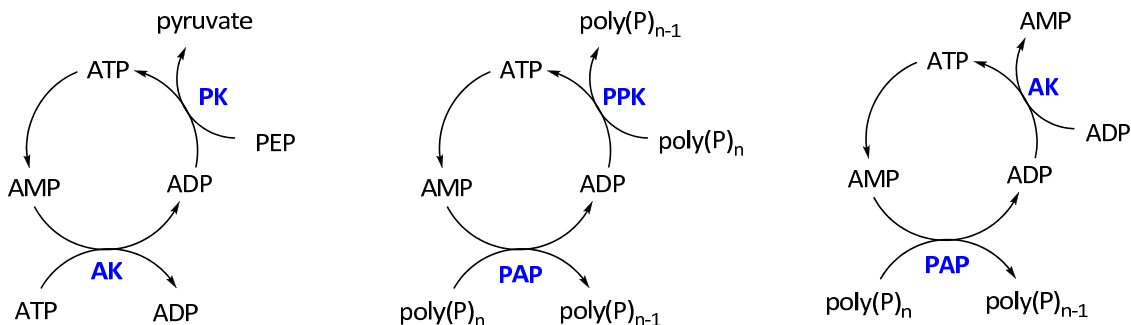
### **CLEA Improvements and Applications**

The greatest limitation of the CLEA particle is the inadequate mechanical properties due to its amorphous nature. CLEAs are well suited to stirred reactions, and easily removed by



centrifugation, but care must be taken to avoid over compression due to excessive centrifugal force. Moreover, attempts to create a continuous flow process with the CLEA pathway described in this dissertation were unsuccessful. Pore size of the filter was either too large, allowing leakage of the aggregate, or so small that fluidic flow was impeded as a result of clogging by the globulous protein. Co-immobilization of cross-linked enzymes and CLEA particles in polyvinyl alcohol (PVA) has proven to increase mechanical stability, as well as increasing turnover in organic solvents.<sup>12, 13</sup> Co-immobilization in PVA is a relatively simple solution to improving the physical properties of CLEAs, possibly allowing larger scale, continuous flow processes.

ATP recycling with pyruvate kinase and adenylate kinase, as described in Chapter IV, is just one of several methods for regeneration of this cofactor.<sup>14-19</sup> In this dissertation, ADP is converted to ATP by pyruvate kinase in combination with phosphoenolpyruvate. Other possibilities include coupling of acetylphosphate with acetate kinase, or polyphosphate with polyphosphate kinase (PPK). AMP was converted to two equivalents of ADP when combined with adenylate kinase (AK) and ATP. Alternatively, polyphosphate:AMP phosphotransferase (PAP) could substitute for AK, using polyphosphate as the phosphoryl donor. Scheme V-1 shows the different possibilities in combining these methods for use in the nucleotide analogue biosynthesis pathway. Polyphosphate biopolymer is significantly less expensive than PEP, the phosphate source used in combination with PK (\$37/kg polyP versus \$213/g PEP). However, PAP has not yet been cloned, only purified from cellular lysates. Also, PPi byproduct from the PRT reaction is an inhibitor of PAP, so an additional enzyme, inorganic pyrophosphatase, would need to be added.



**Scheme V-1: ATP recycling schemes. Left; original method used for nucleotide pathway using adenylate kinase (AK) and pyruvate kinase (PK). Middle; polyphosphate kinase (PPK) and polyphosphate:AMP phosphotransferase (PAP). Right; Adenylate kinase (AK) and PAP coupled regeneration. Adapted in part from Zhao, *et al.*<sup>14</sup>**

As described earlier in this chapter, using the example of catechol biosynthesis, CLEA particles can provide a remedy to problems with product toxicity in the biosynthesis of non-natural products. CLEA immobilization of cell lysates could also be applied to combinatorial biosynthetic strategies that are currently limited to producing non-natural compounds that are non toxic to the host. Examples of such ventures include heterologous expression of plant biosynthetic pathways into *E. coli* for biosynthesis of novel polyketides,<sup>20</sup> or for generation of new antibiotics. Manipulation of genetic pathways for combinatorial biosynthesis of novel aminocoumarin antibiotics, such as Novobiocin, Clorobiocin, and Coumermycin, have been reported, but are limited to various species of slow growing *Streptomyces*. Heterologous expression of these pathways into *E. coli* could greatly increase the product output.

Whole cell biosynthesis is also limited to the transport or permeability of starting materials through the host cell membrane. Charged substrates are severely limited, if not impossible substrates for whole cell synthesis due to transport barriers, and great lengths must be taken to bypass this barrier. Biosynthesis of simvastatin using a whole cell system required adaptation of the pathway to accept a membrane-permeable substrate,  $\alpha$ -dimethylbutyryl-S-

methyl-mercaptopropionate.<sup>21</sup> The original simvastatin producing cells, immobilized as crude lysate CLEAs, would have no issue with substrate permeability.

CLEA particles have also shown considerable stability in organic solvents, compared to non-bound enzyme. It is conceivable that even greater resistance could be achieved through co-immobilization, or entrapment in a gel. Therefore, whole cell lysate CLEAs can provide product in cases where substrate solubility have prevented enzymatic biosynthesis, due to aqueous requirements of the enzymes. Such as the case in the biosynthesis of  $\alpha$ -pinene, which required an aqueous-organic two phase system due to lack of substrate solubility in aqueous solutions.<sup>22</sup> These are just a few examples of limited whole cell biosyntheses of non-natural products that could benefit from CLEA formation. No doubt this same method could improve countless other whole cell systems with similar limitations.

## References

1. Sheldon, R. A., E factors, green chemistry and catalysis: an odyssey. *Chemical Communications* **2008**, (29), 3352-3365.
2. Klare, M. T., The deadly nexus: Oil, terrorism, and America's national security. *Current History* **2002**, 101, (659), 414.
3. Robertson, J. G., Mechanistic basis of enzyme-targeted drugs. *Biochemistry* **2005**, 44, (15), 5561-5571.
4. Draths, K. M.; Frost, J. W., Environmentally Compatible Synthesis of Catechol from D-Glucose. *Journal of the American Chemical Society* **1995**, 117, (9), 2395-2400.
5. Li, W. S.; Xie, D. M.; Frost, J. W., Benzene-free synthesis of catechol: Interfacing microbial and chemical catalysis. *Journal of the American Chemical Society* **2005**, 127, (9), 2874-2882.
6. Shah, D.; Zhang, Z. G.; Khodursky, A.; Kaldalu, N.; Kurg, K.; Lewis, K., Persisters: a distinct physiological state of E-coli. *Bmc Microbiology* **2006**, 6.
7. Antonino, L. C.; Straub, K.; Wu, J. C., Probing the Active-Site of Human Imp Dehydrogenase Using Halogenated Purine Riboside 5'-Monophosphates and Covalent Modification Reagents. *Biochemistry* **1994**, 33, (7), 1760-1765.

8. Wang, W.; Papov, V. V.; Minakawa, N.; Matsuda, A.; Biemann, K.; Hedstrom, L., Inactivation of inosine 5'-monophosphate dehydrogenase by the antiviral agent 5-ethynyl-1-beta-D-ribofuranosylimidazole-4-carboxamide 5'-monophosphate. *Biochemistry* **1996**, 35, (1), 95-101.
9. Wang, G. Y.; Sakthivel, K.; Rajappan, V.; Bruice, T. W.; Tucker, K.; Fagan, P.; Brooks, J. L.; Hurd, T.; Leeds, J. M.; Cook, P. D., Synthesis of azole nucleoside 5'-monophosphate mimics (P1Ms) and their inhibitory properties of IMP dehydrogenases. *Nucleosides Nucleotides & Nucleic Acids* **2004**, 23, (1-2), 317-337.
10. Chuvikovsky, D. V.; Esipov, R. S.; Skoblov, Y. S.; Chupova, L. A.; Muravyova, T. I.; Miroshnikov, A. I.; Lapinjoki, S.; Mikhailopulo, I. A., Ribokinase from E-coli: Expression, purification, and substrate specificity. *Bioorganic & Medicinal Chemistry* **2006**, 14, (18), 6327-6332.
11. Krath, B. N.; Hove-Jensen, B., Class II recombinant phosphoribosyl diphosphate synthase from spinach - Phosphate independence and diphosphoryl donor specificity. *Journal of Biological Chemistry* **2001**, 276, (21), 17851-17856.
12. Groger, H.; Capan, E.; Barthuber, A.; Vorlop, K. D., Asymmetric synthesis of an (R)-cyanohydrin using enzymes entrapped in lens-shaped gels. *Organic Letters* **2001**, 3, (13), 1969-1972.
13. Wilson, L.; Illanes, A.; Pessela, B. C. C.; Abian, O.; Fernandez-Lafuente, R.; Guisan, J. M., Encapsulation of crosslinked penicillin G acylase aggregates in lentikats: Evaluation of a novel biocatalyst in organic media. *Biotechnology and Bioengineering* **2004**, 86, (5), 558-562.
14. Zhao, H. M.; van der Donk, W. A., Regeneration of cofactors for use in biocatalysis. *Current Opinion in Biotechnology* **2003**, 14, (6), 583-589.
15. Gross, A.; Abril, O.; Lewis, J. M.; Geresh, S.; Whitesides, G. M., Practical Synthesis of 5-Phospho-D-Ribosyl Alpha-1-Pyrophosphate (Prpp) - Enzymatic Routes from Ribose 5-Phosphate or Ribose. *Journal of the American Chemical Society* **1983**, 105, (25), 7428-7435.
16. Calhoun, K. A.; Swartz, J. R., Energizing cell-free protein synthesis with glucose metabolism. *Biotechnology and Bioengineering* **2005**, 90, (5), 606-613.
17. Kim, D. M.; Swartz, J. R., Regeneration of adenosine triphosphate from glycolytic intermediates for cell-free protein synthesis. *Biotechnology and Bioengineering* **2001**, 74, (4), 309-316.
18. Nahalka, J.; Patoprsty, V., Enzymatic synthesis of sialylation substrates powered by a novel polyphosphate kinase (PPK3). *Organic & Biomolecular Chemistry* **2009**, 7, (9), 1778-1780.

19. Resnick, S. M.; Zehnder, A. J. B., In vitro ATP regeneration from polyphosphate and AMP by polyphosphate: AMP phosphotransferase and adenylate kinase from *Acinetobacter johnsonii* 210A. *Applied and Environmental Microbiology* **2000**, 66, (5), 2045-2051.
20. Katsuyama, Y.; Funa, N.; Miyahisa, I.; Horinouchi, S., Synthesis of unnatural flavonoids and stilbenes by exploiting the plant biosynthetic pathway in *Escherichia coli*. *Chemistry & Biology* **2007**, 14, (6), 613-621.
21. Xie, X. K.; Tang, Y., Efficient synthesis of simvastatin by use of whole-cell biocatalysis. *Applied and Environmental Microbiology* **2007**, 73, (7), 2054-2060.
22. Schewe, H.; Holtmann, D.; Schrader, J., P450(BM-3)-catalyzed whole-cell biotransformation of alpha-pinene with recombinant *Escherichia coli* in an aqueous-organic two-phase system. *Applied Microbiology and Biotechnology* **2009**, 83, (5), 849-857.

APPENDIX A

PERCENT CONVERSION NMR FROM CHAPTER III

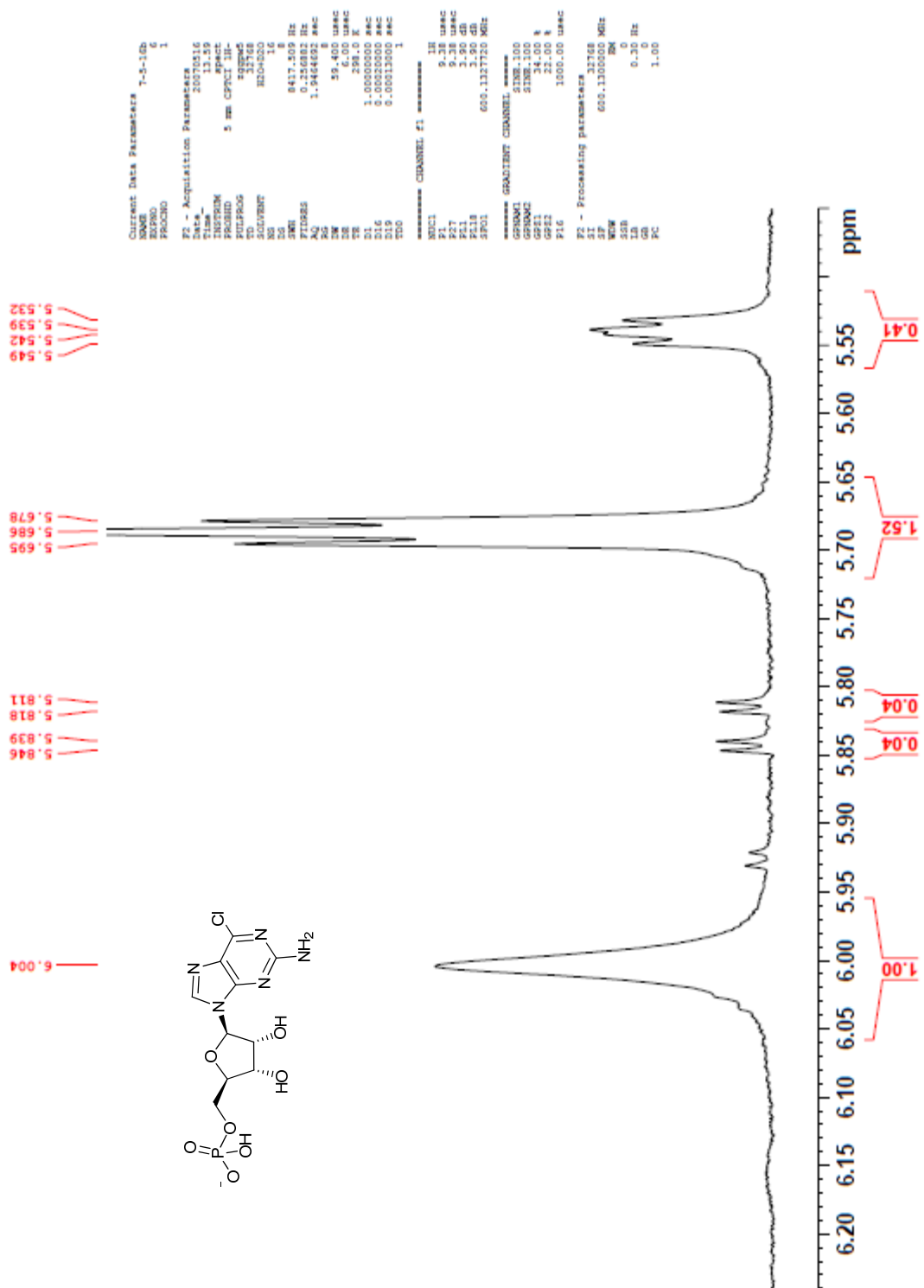


Figure A1: 6-chloroguanosine monophosphate

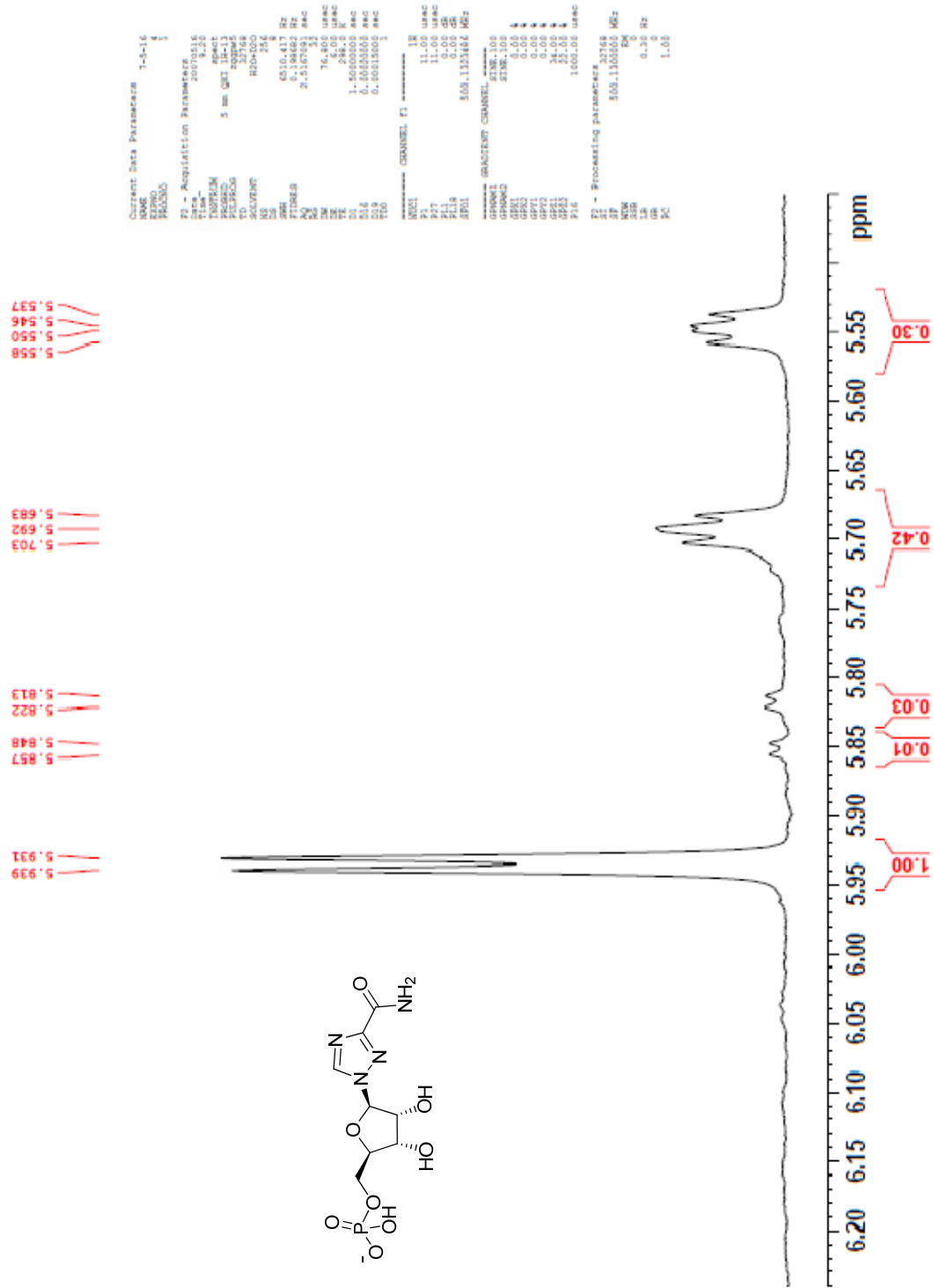


Figure A2: ribavirin monophosphate



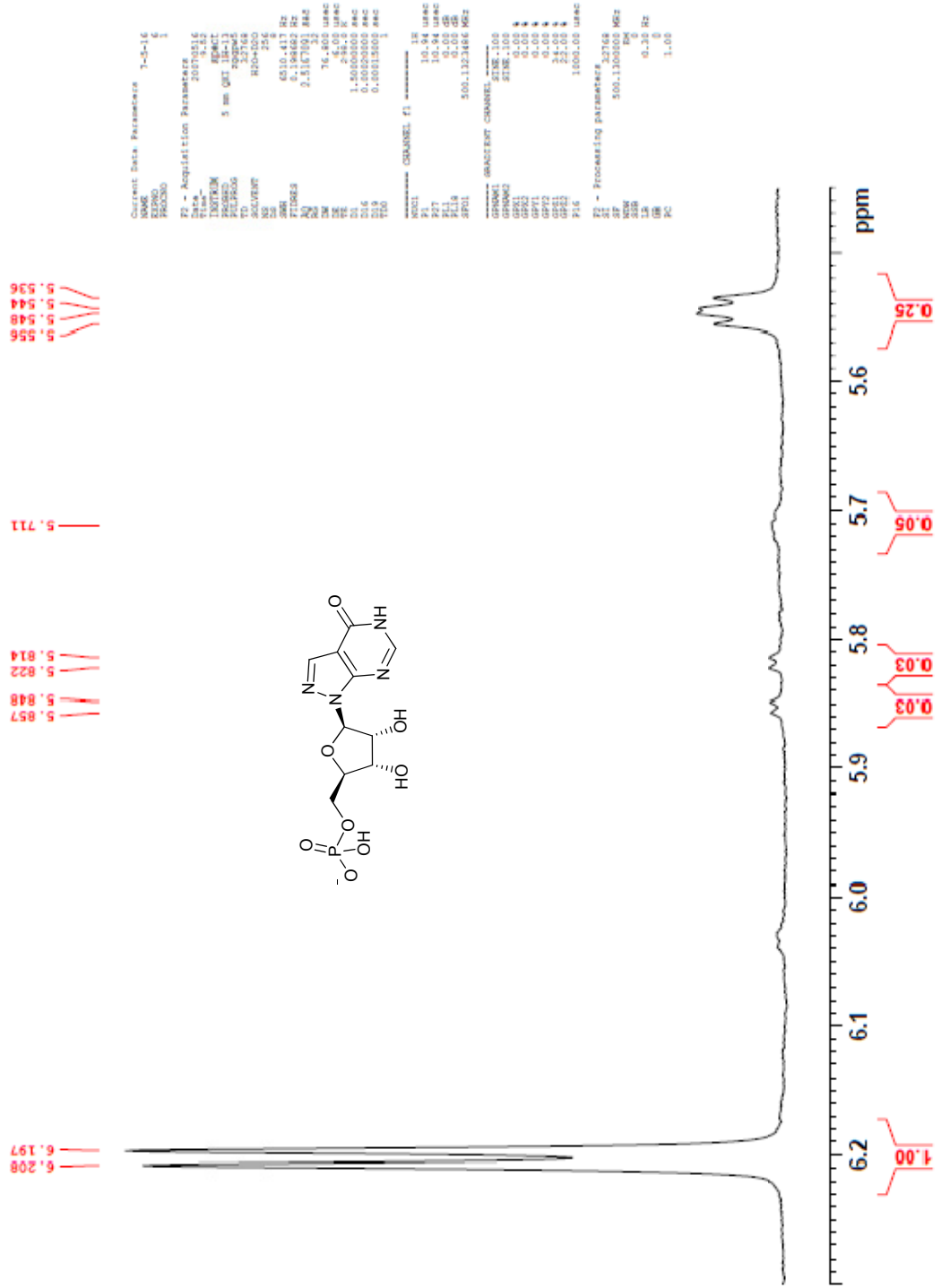


Figure A3: allopurinol monophosphate

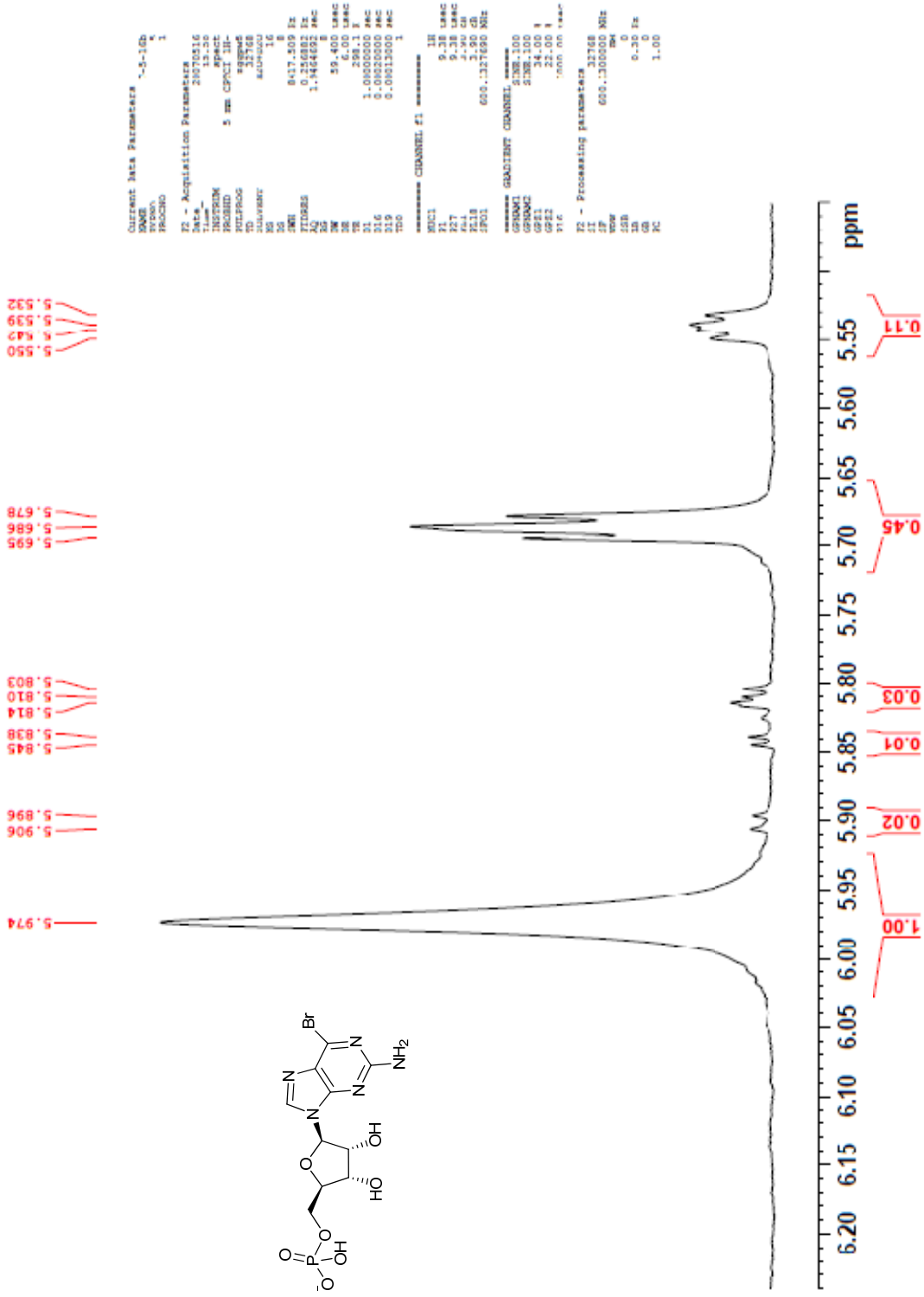


Figure A4: 6-bromoguanosine monophosphate

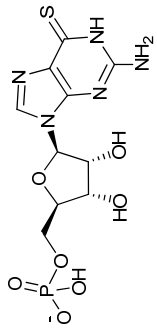
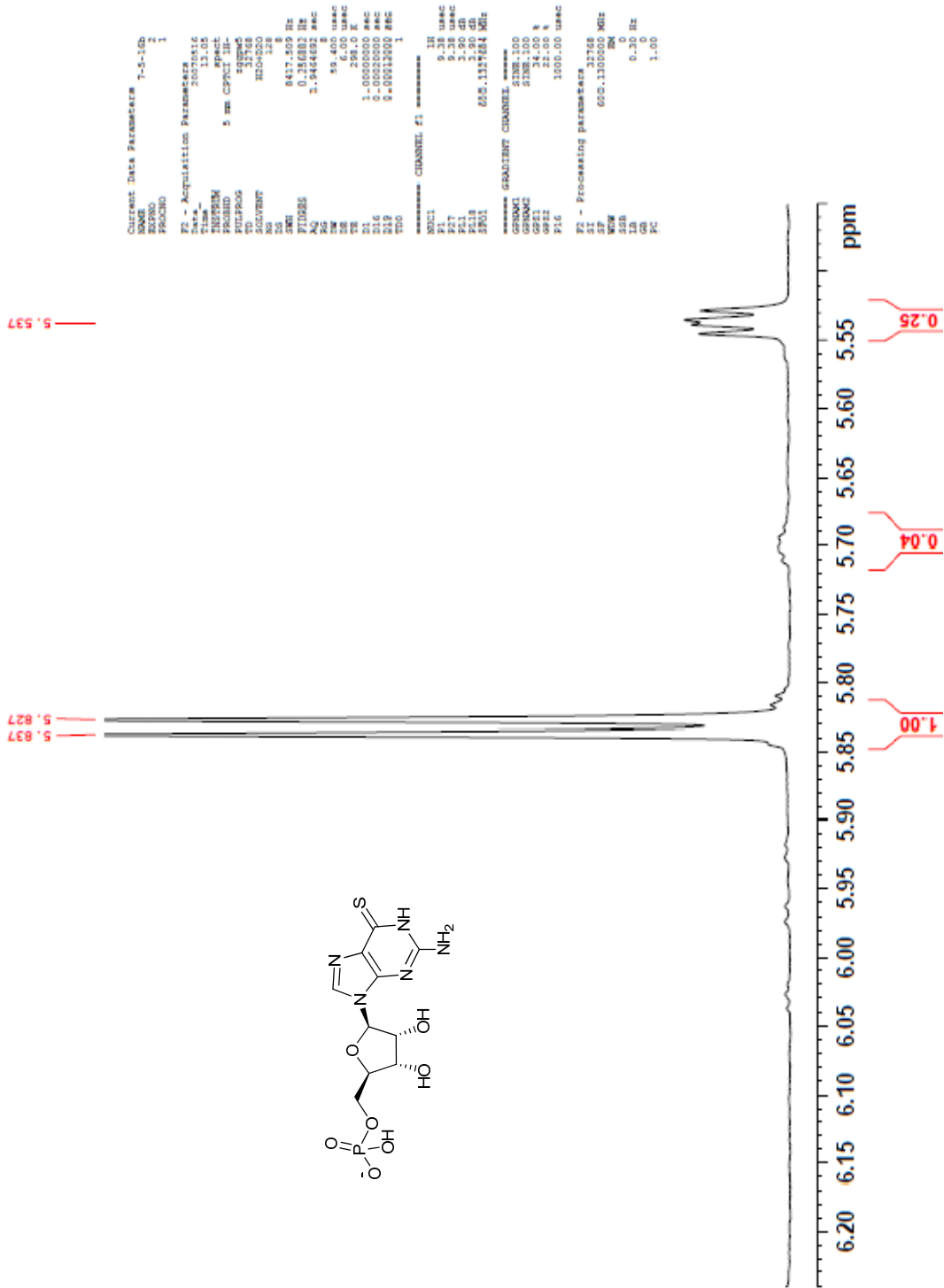


Figure A5: thioguanosine monophosphate

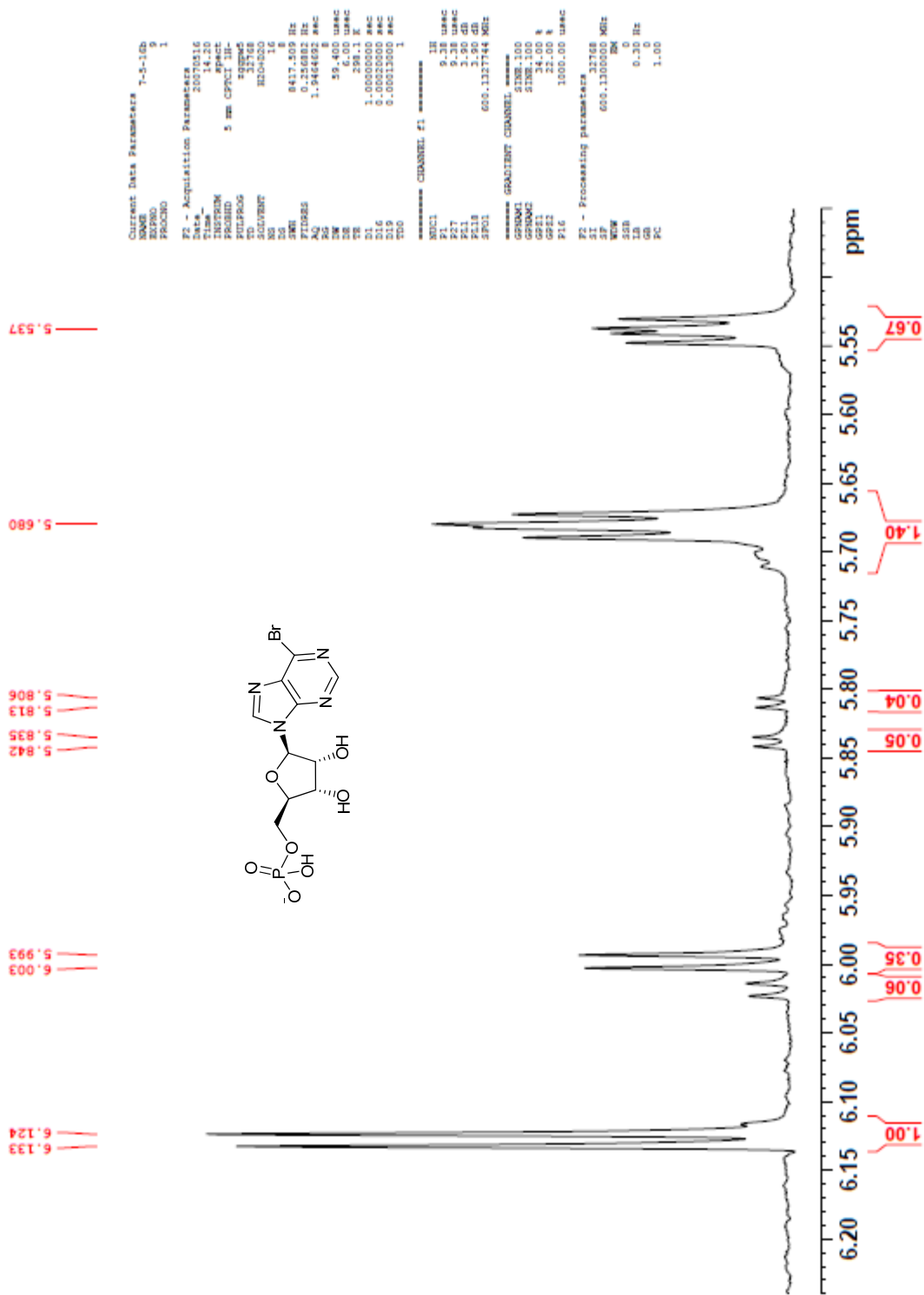


Figure A6: 6-bromopurine monophosphate

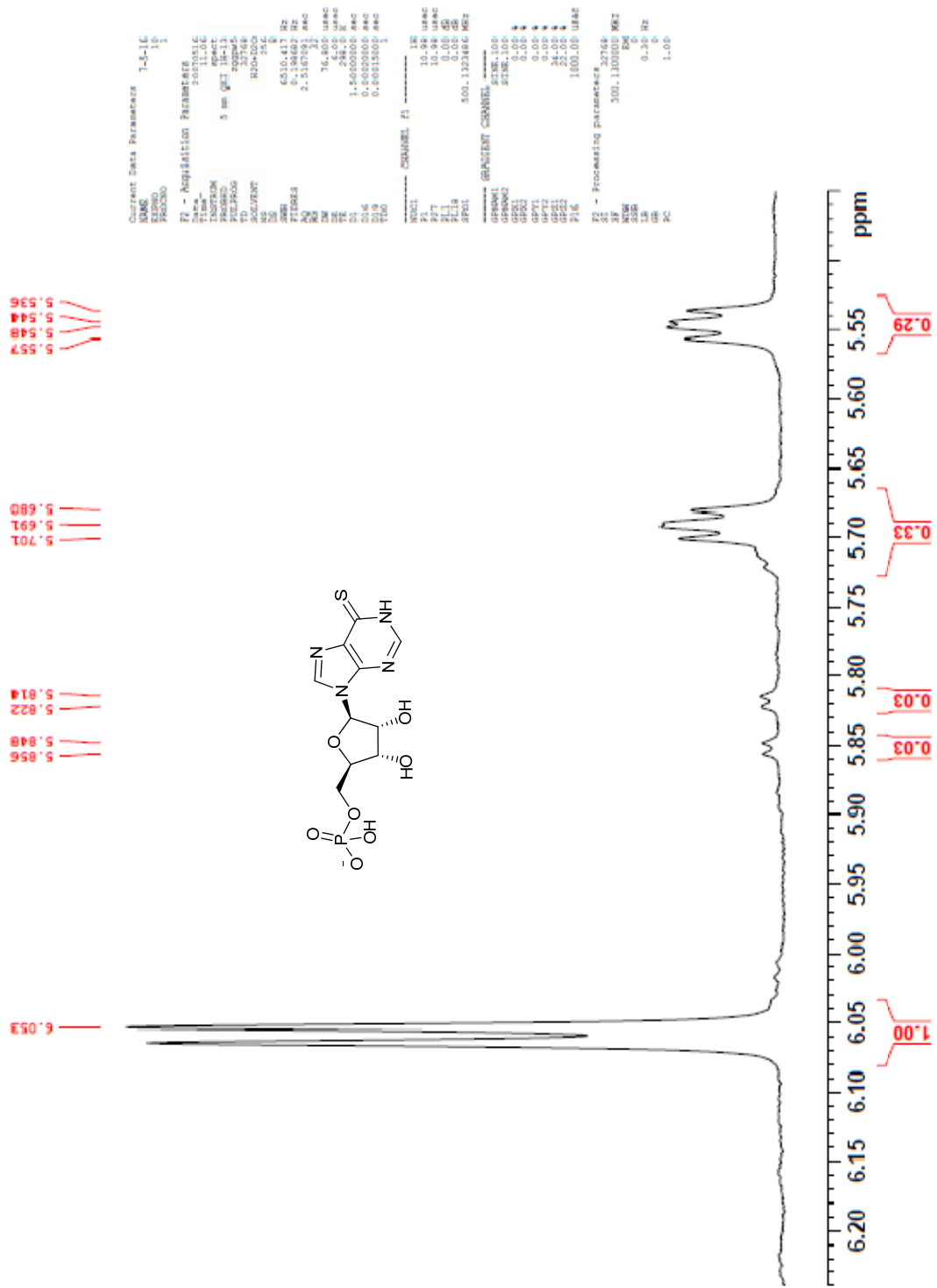


Figure A7: mercaptopurine monophosphate

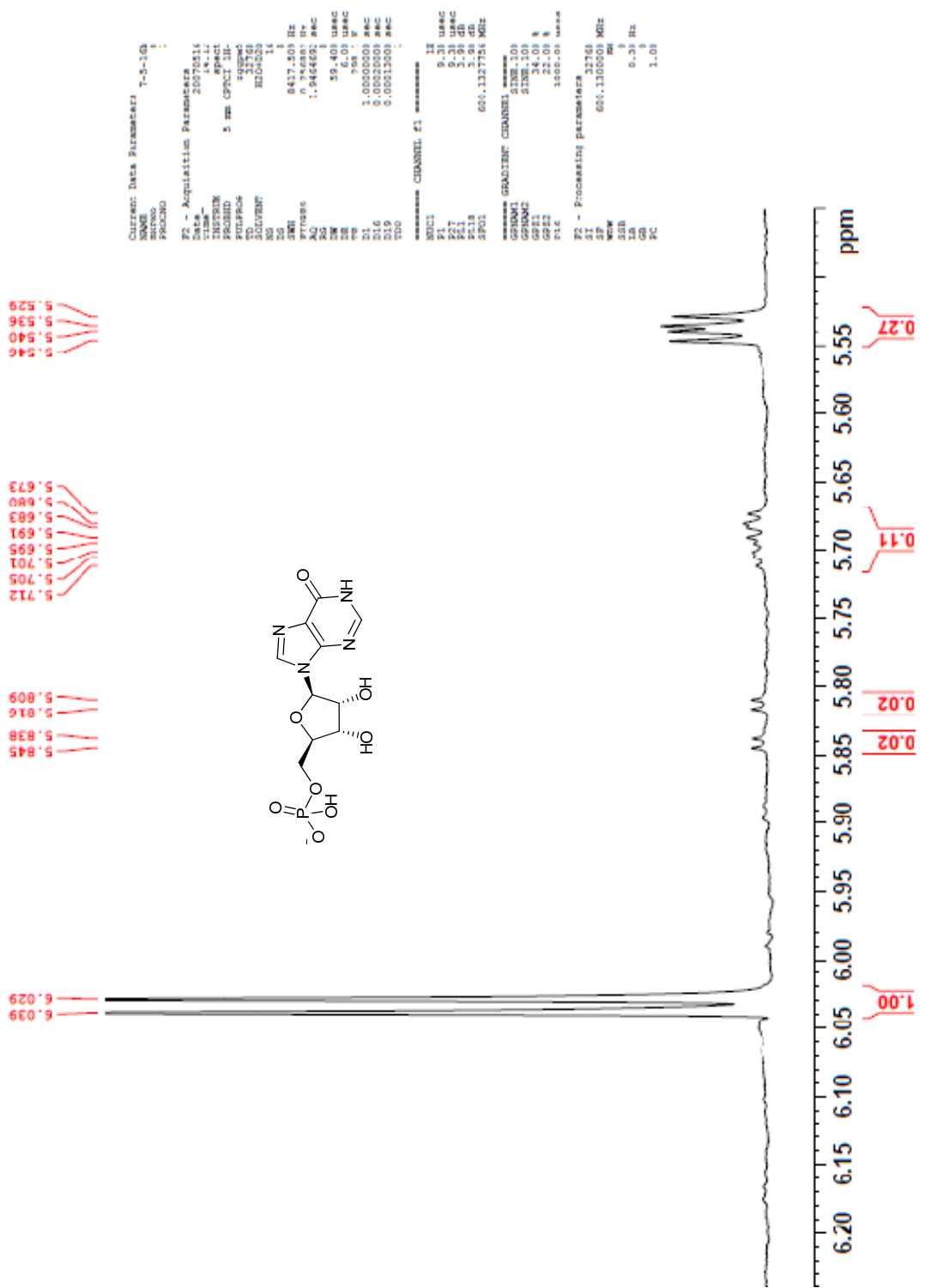


Figure A8: inosine monophosphate

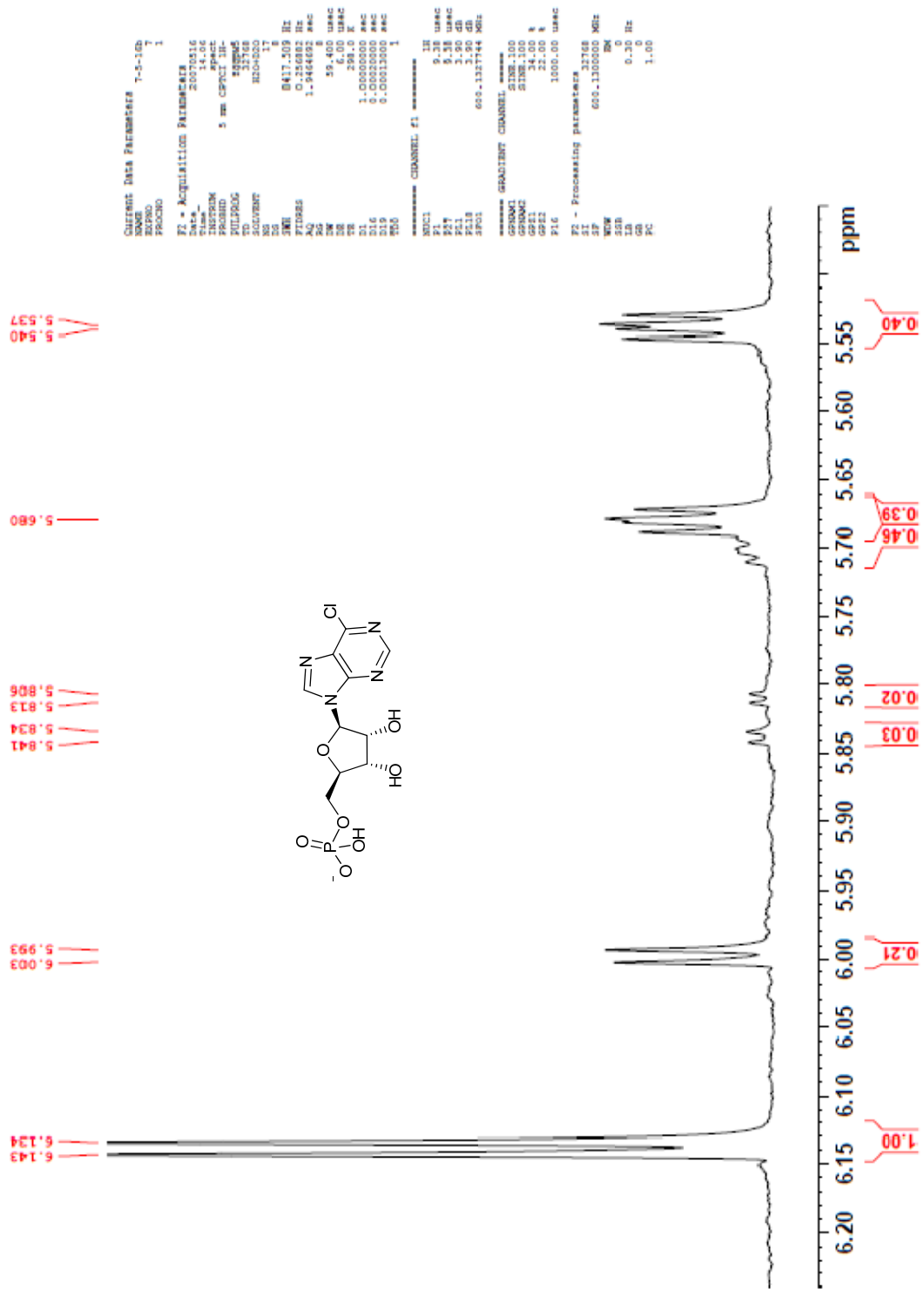


Figure A9: 6-chloropurine monophosphate

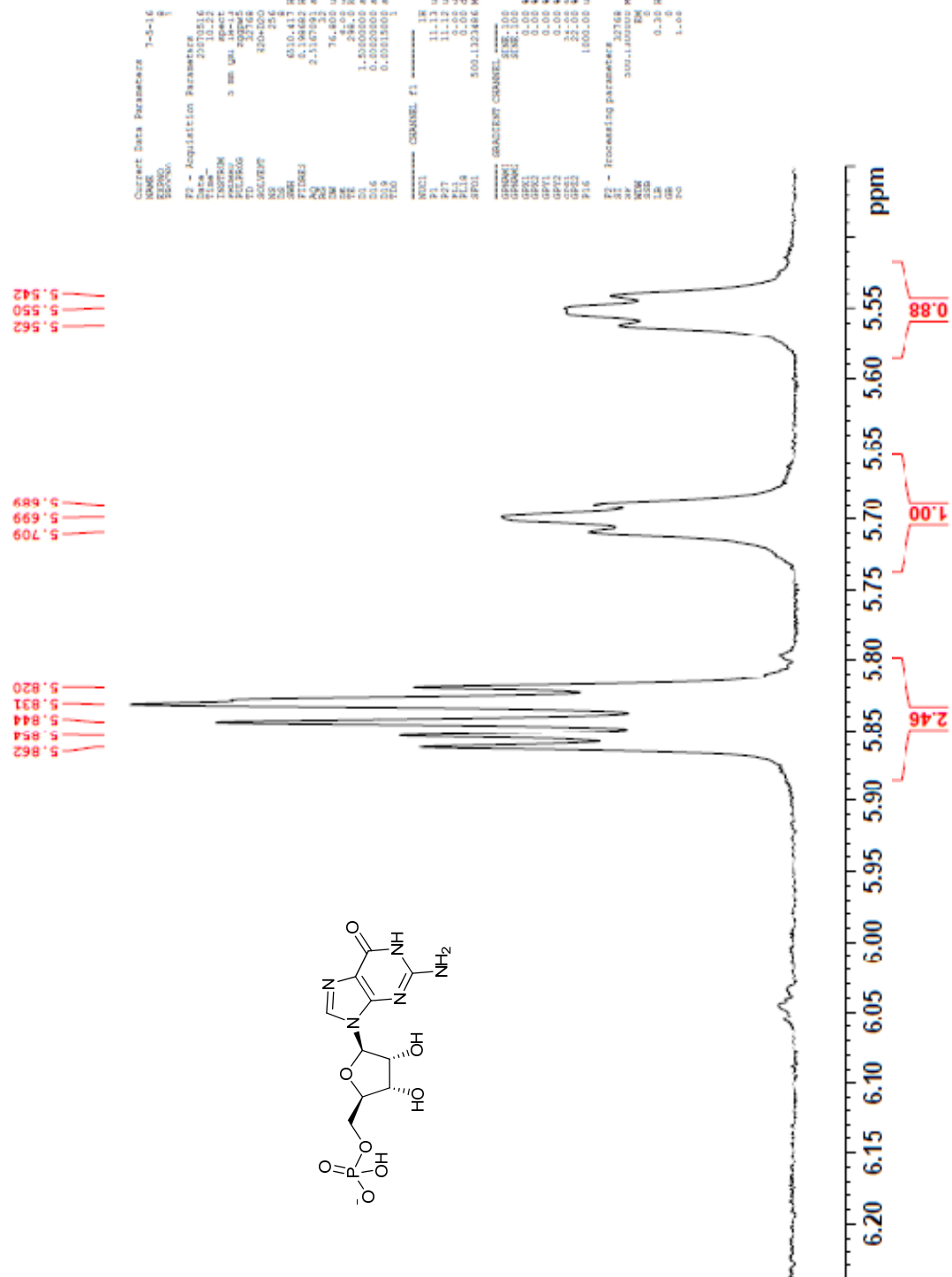


Figure A10: guanosine monophosphate



APPENDIX B

PURIFIED PRODUCT NMR FROM CHAPTER III

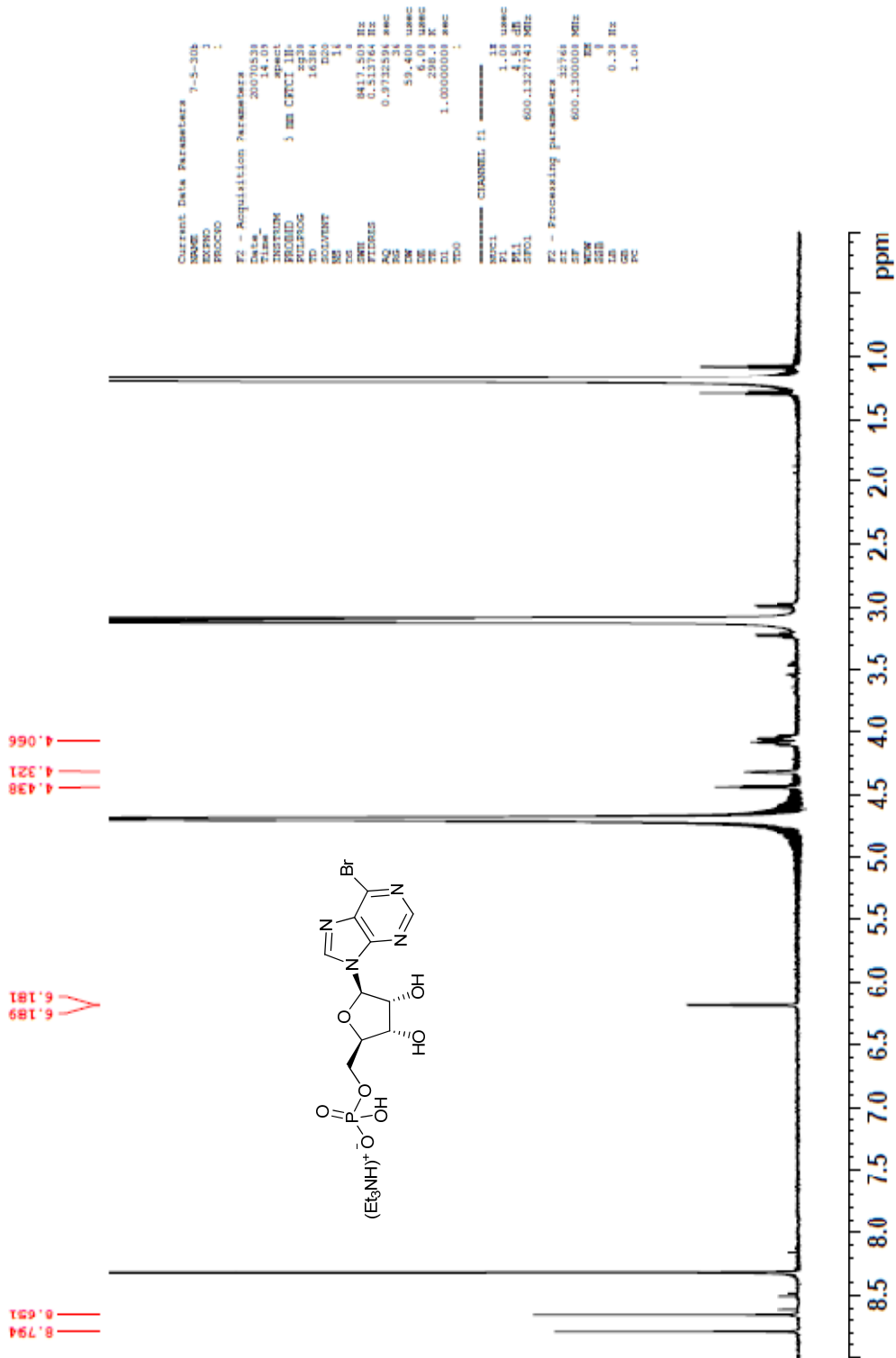


Figure B1: <sup>1</sup>H-NMR; 6-bromopurine monophosphate

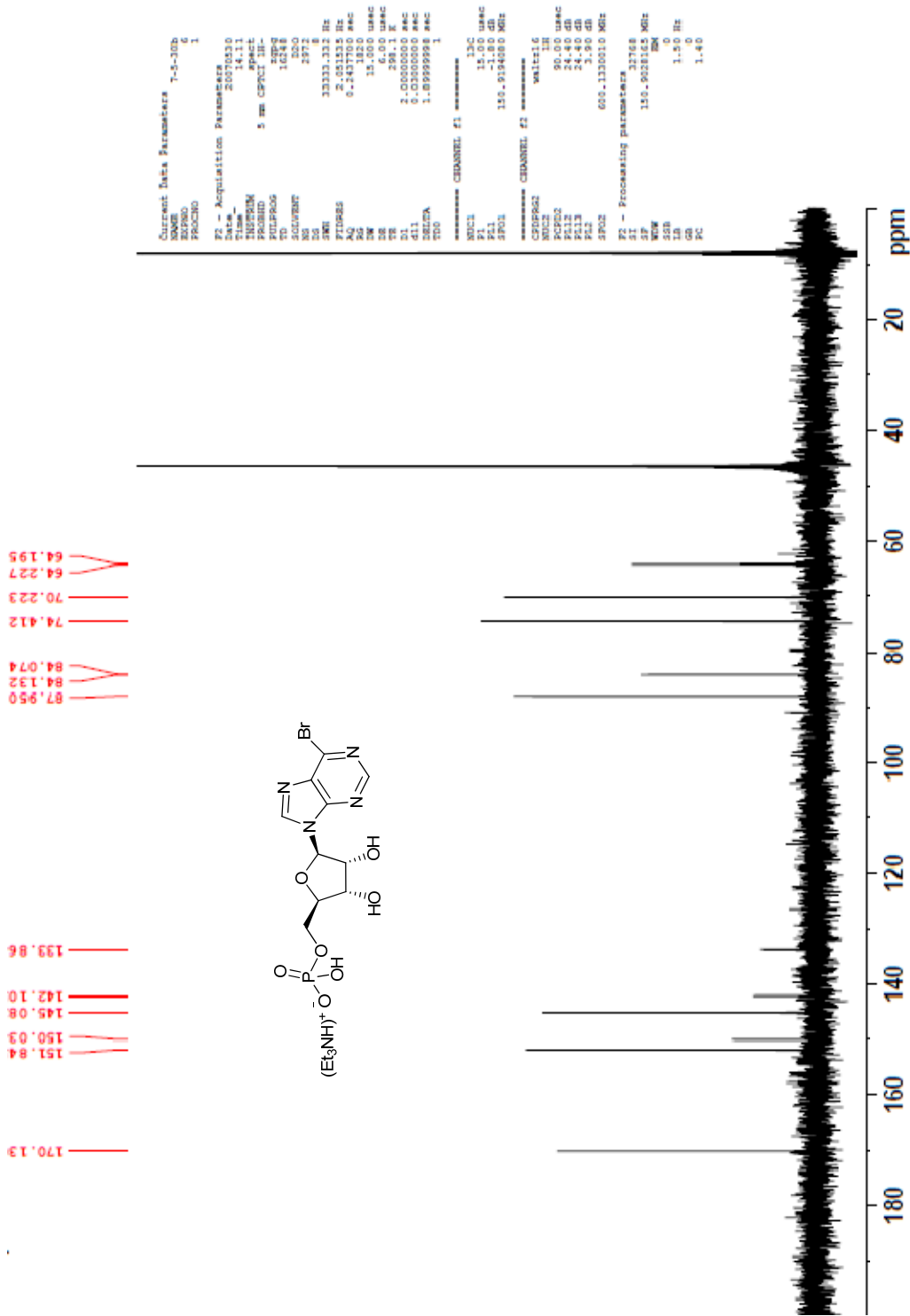


Figure B2: <sup>13</sup>C-NMR; 6-bromopurine monophosphate

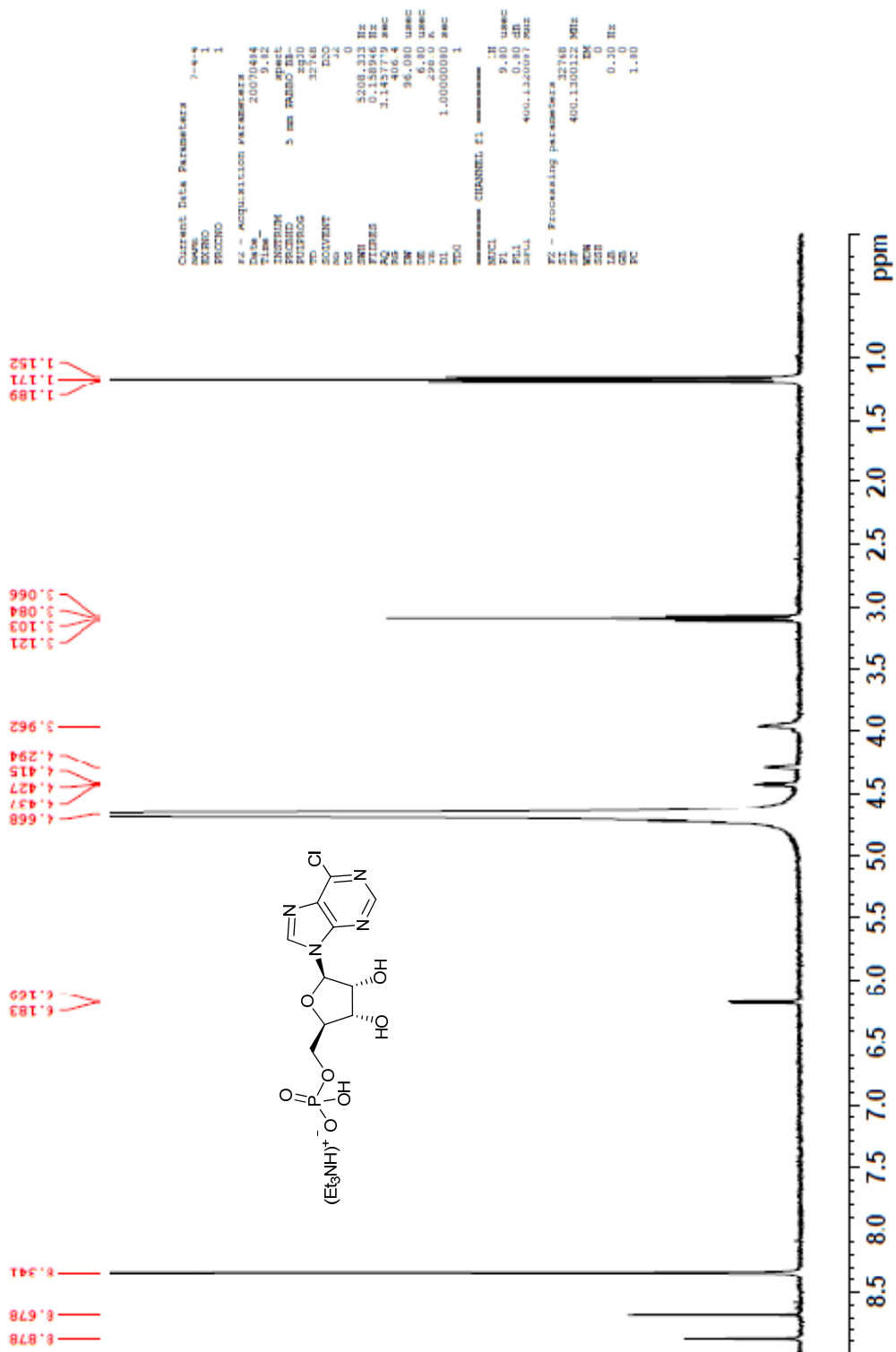


Figure B3: <sup>1</sup>H-NMR; 6-chloropurine monophosphate

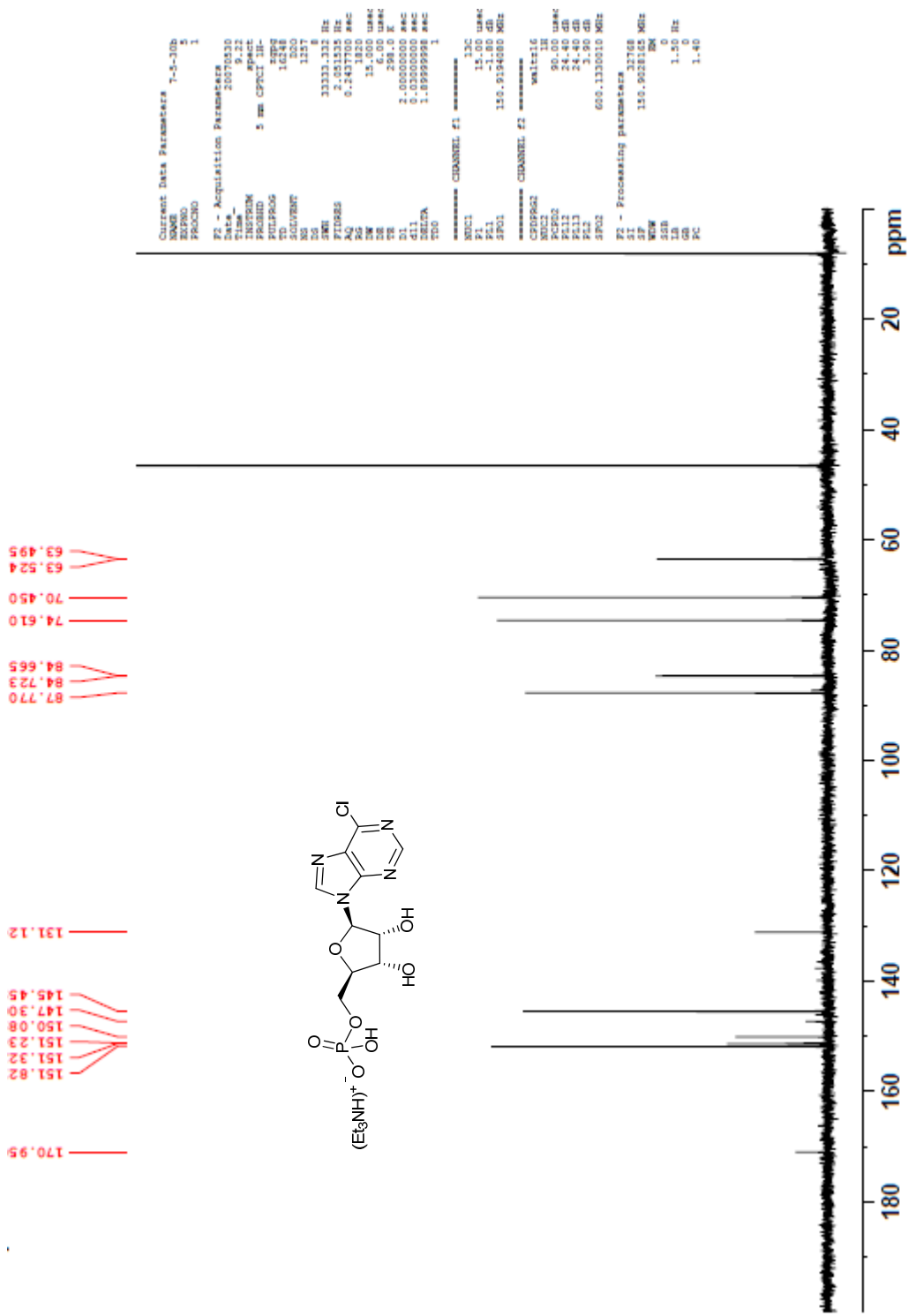


Figure B4: <sup>13</sup>C-NMR; 6-chloropurine monophosphate

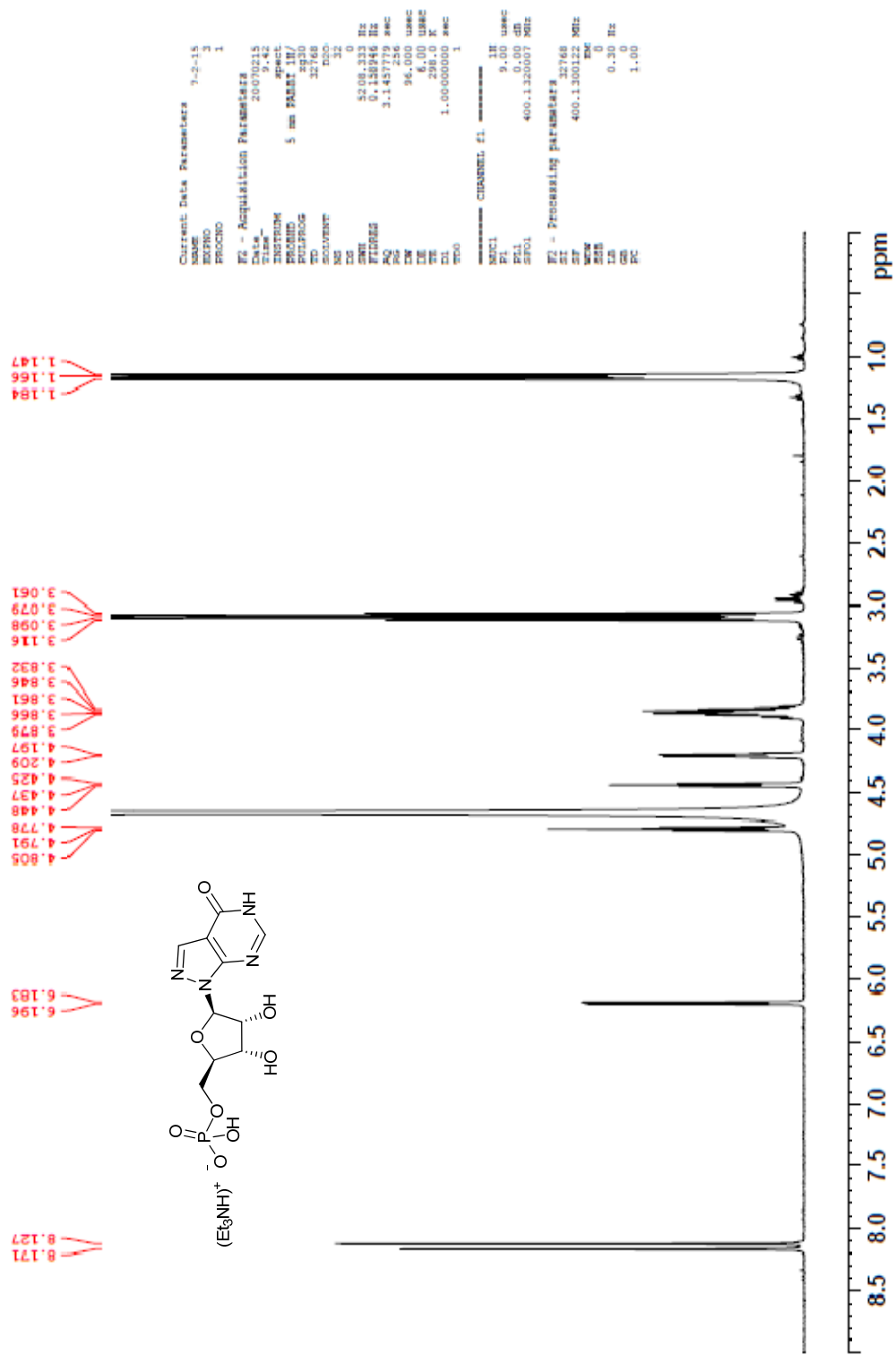


Figure B5: <sup>1</sup>H-NMR; allopurinol monophosphate

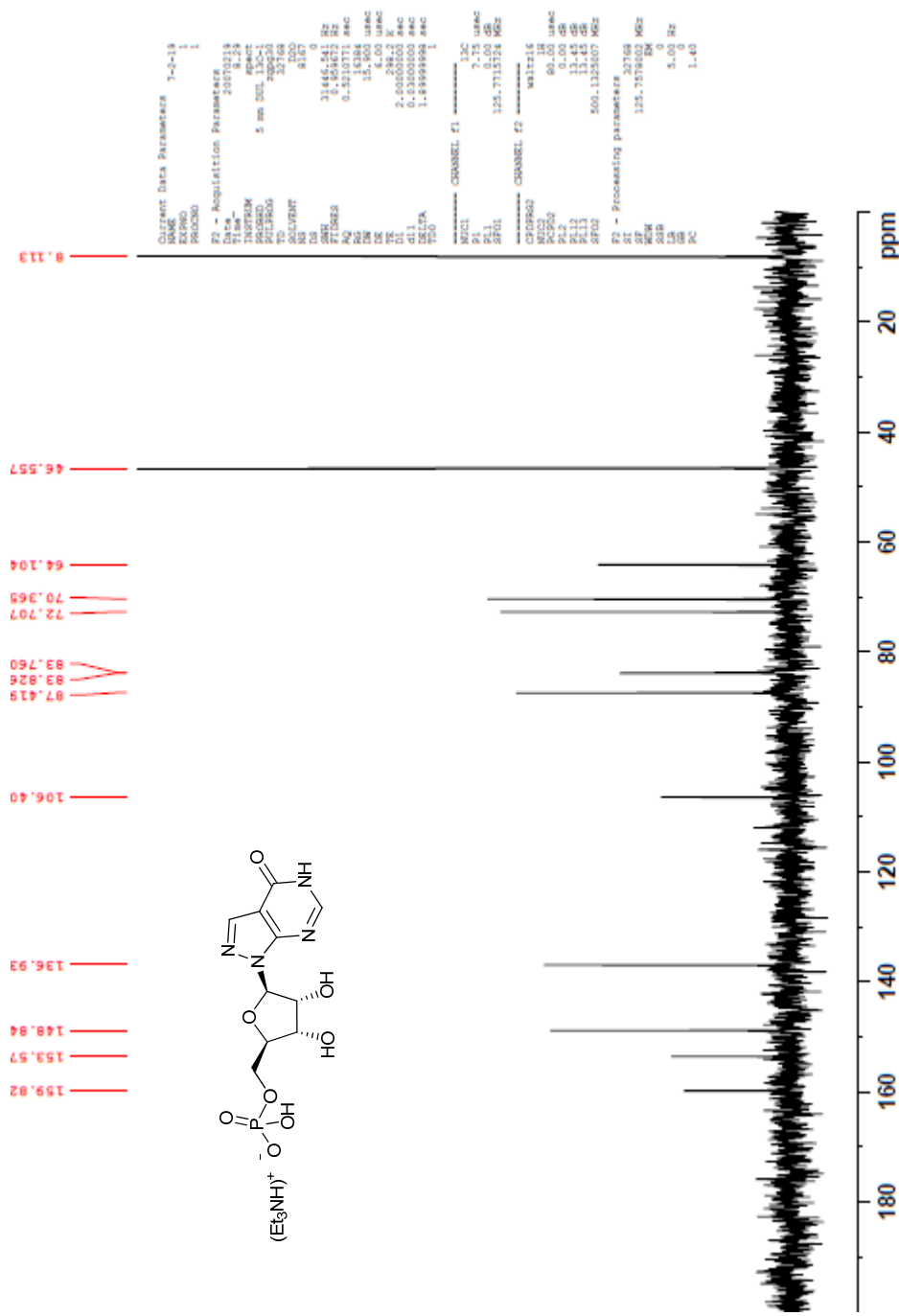


Figure B6: <sup>13</sup>C-NMR; allopurine monophosphate

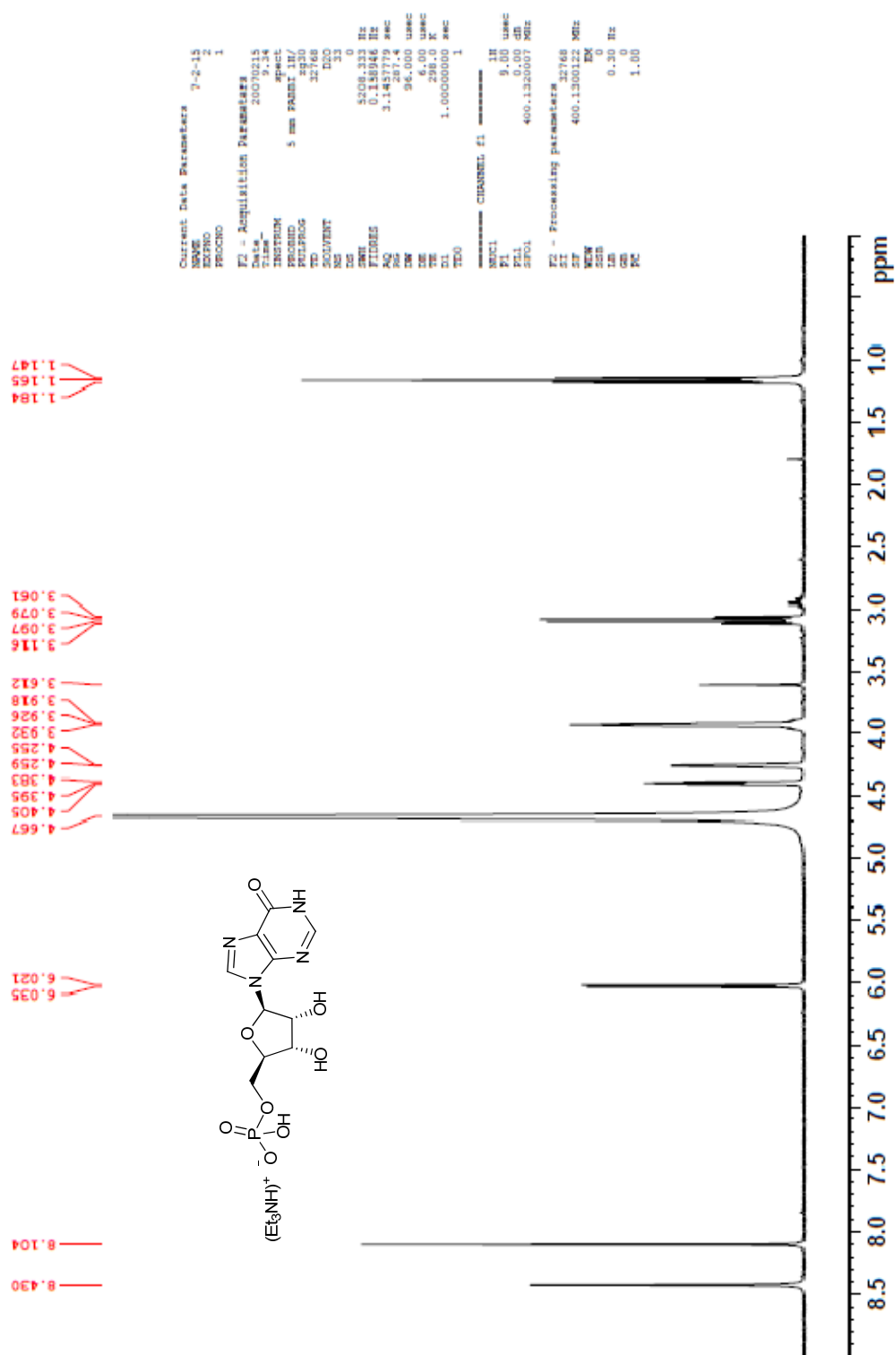


Figure B7: <sup>1</sup>H-NMR; inosine monophosphate



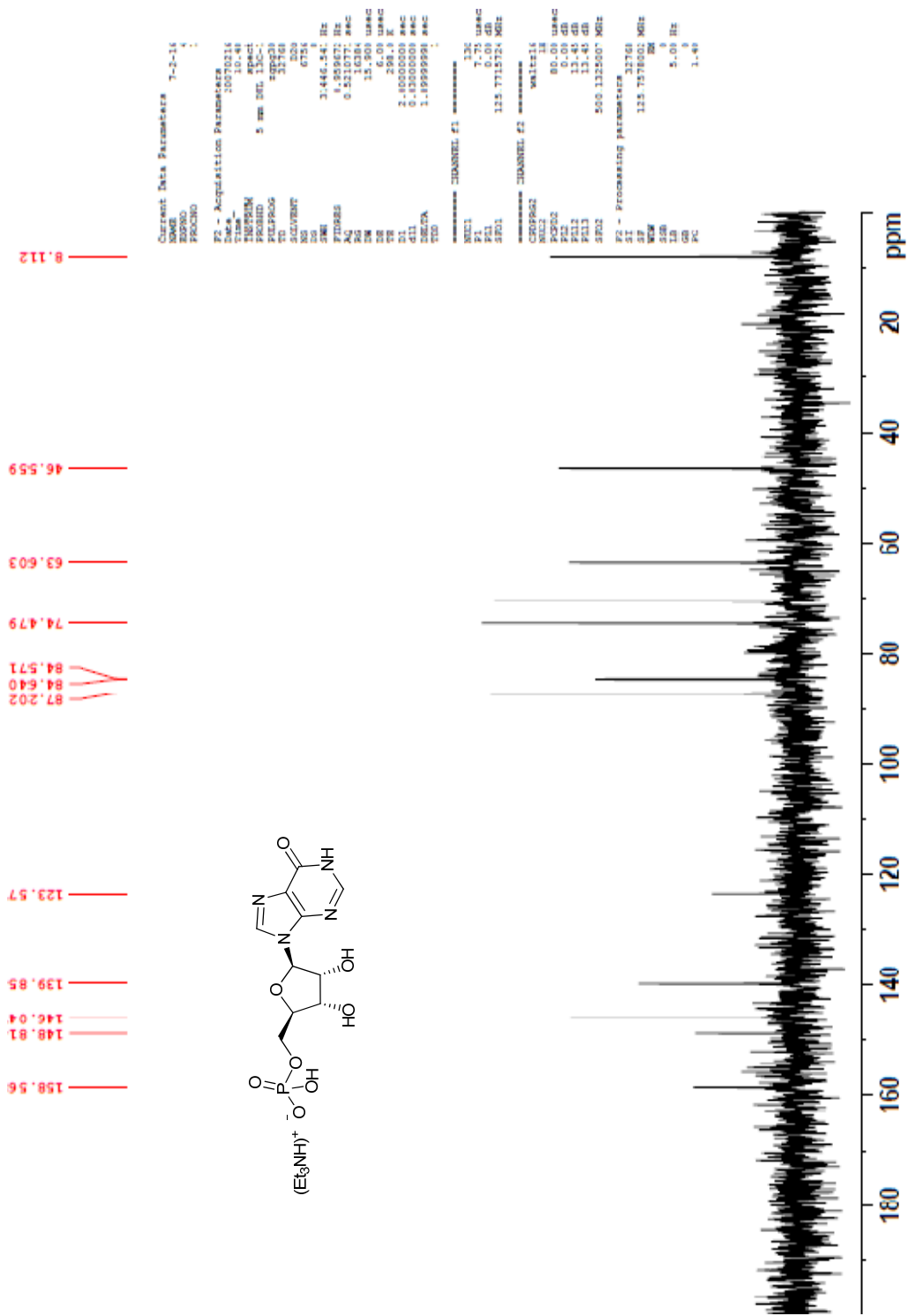


Figure B8: <sup>13</sup>C-NMR; inosine monophosphate

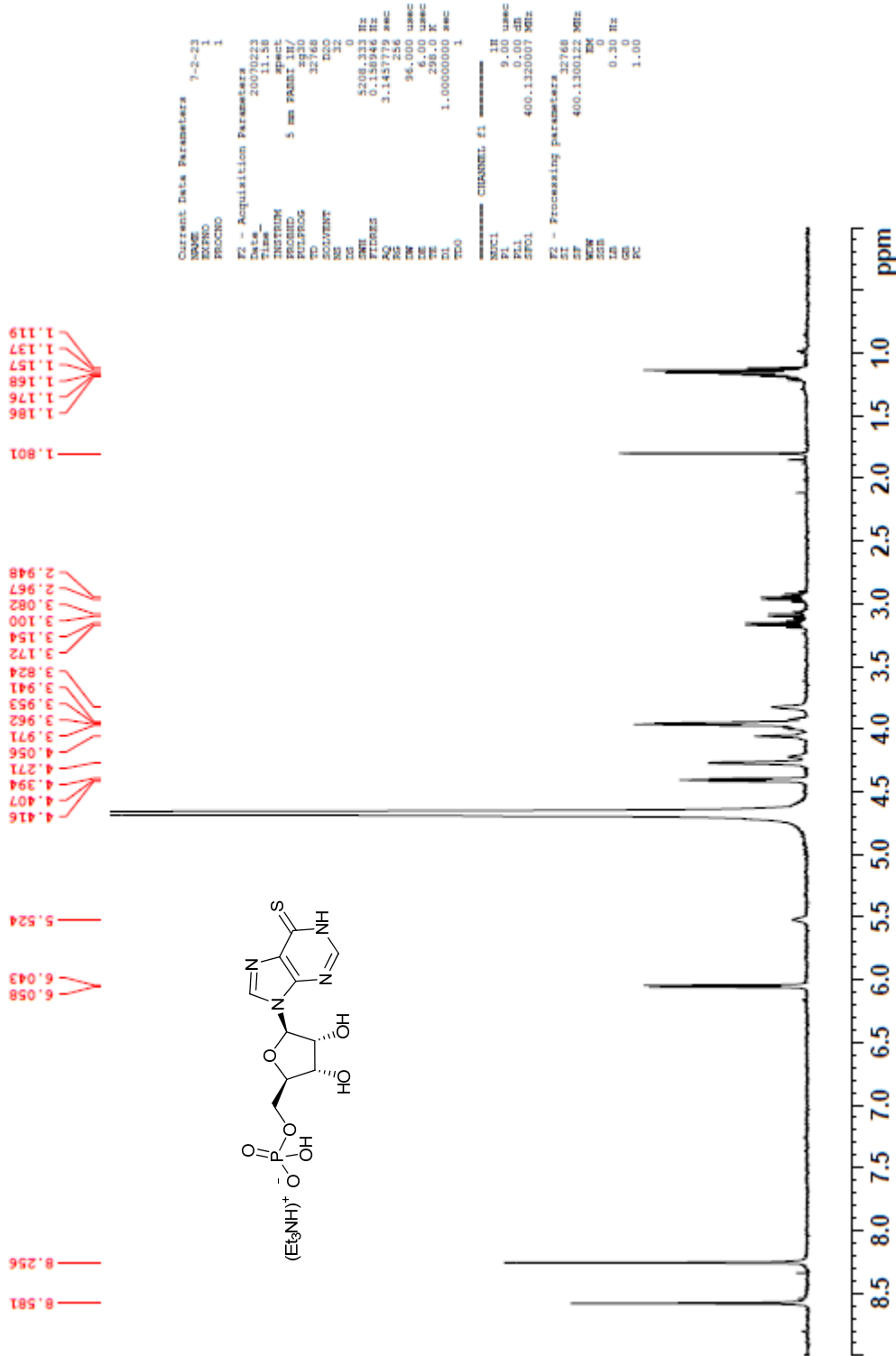


Figure B9: <sup>1</sup>H-NMR; mercaptopurine monophosphate

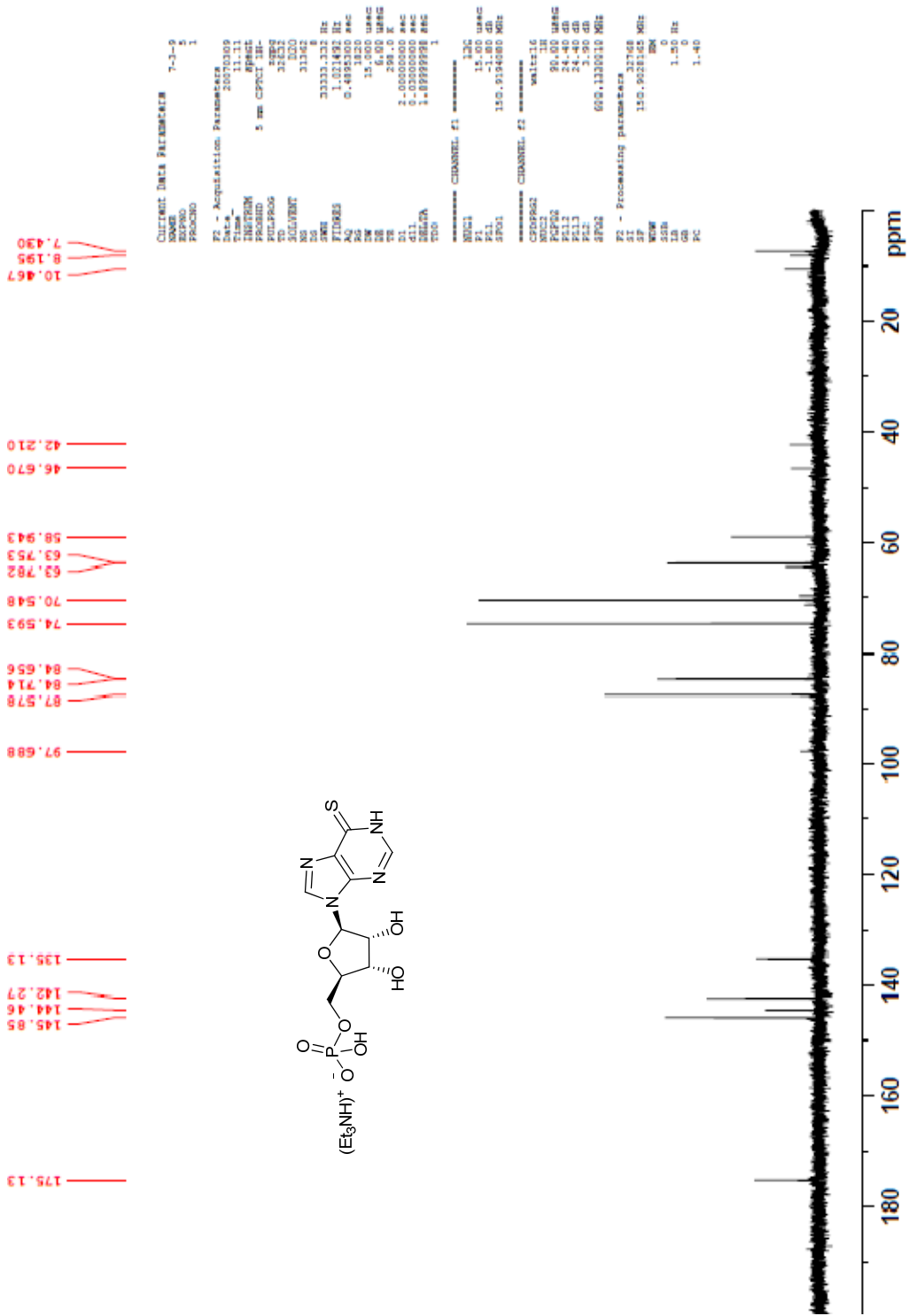


Figure B10: <sup>13</sup>C-NMR; mercaptopurine monophosphate

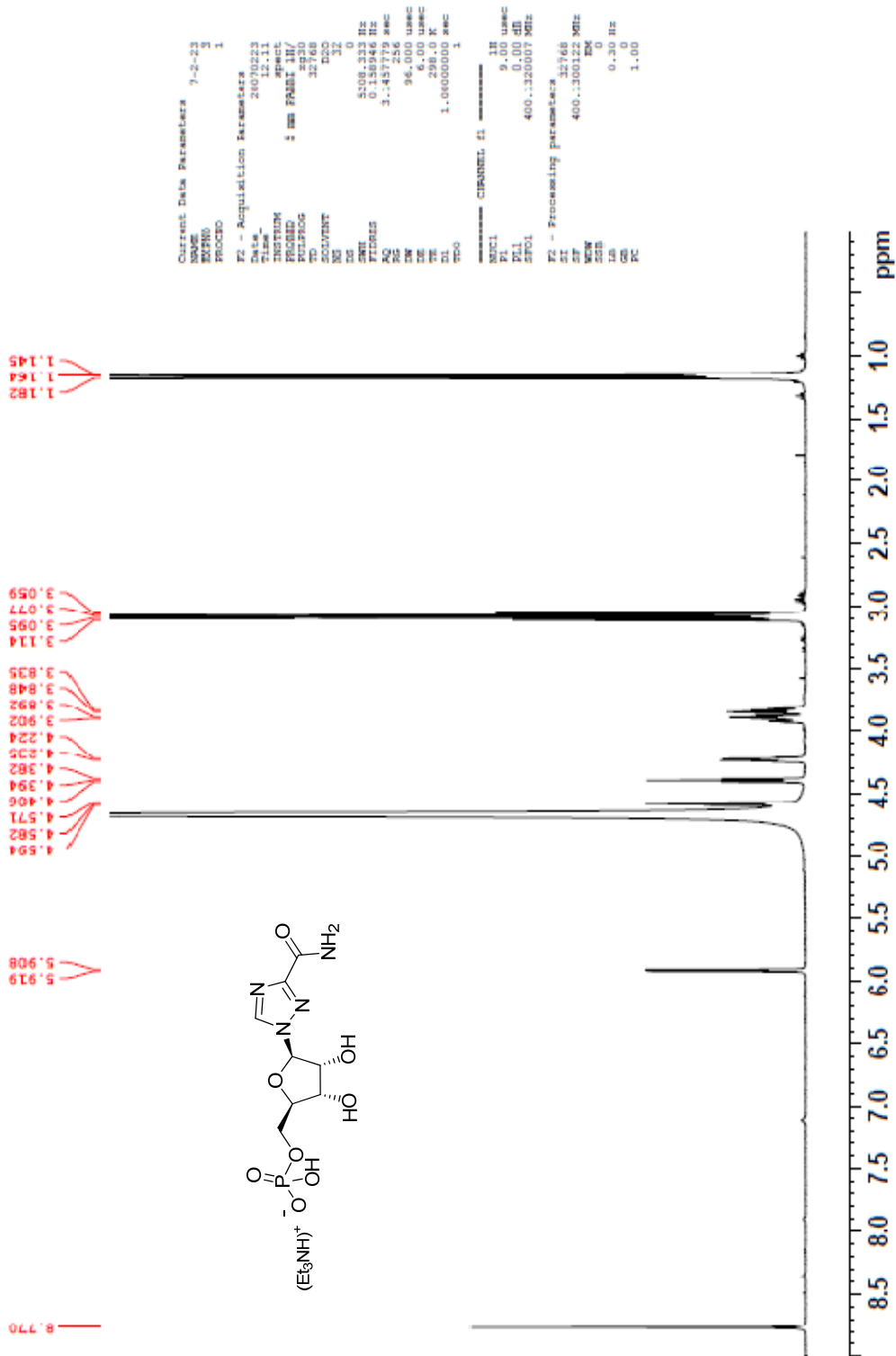


Figure B11: <sup>1</sup>H-NMR; ribavirin monophosphate

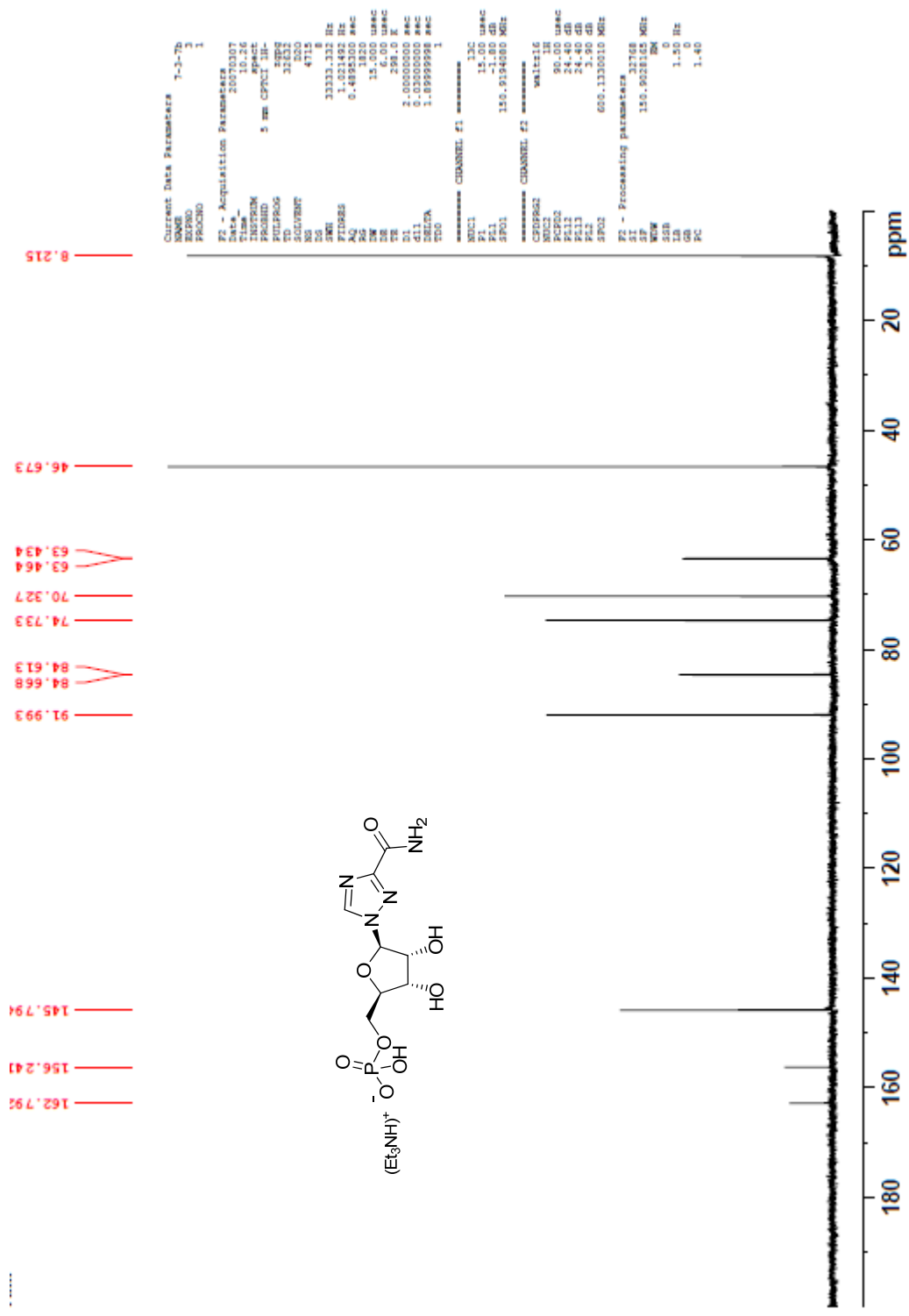


Figure B12: <sup>13</sup>C-NMR; ribavirin monophosphate

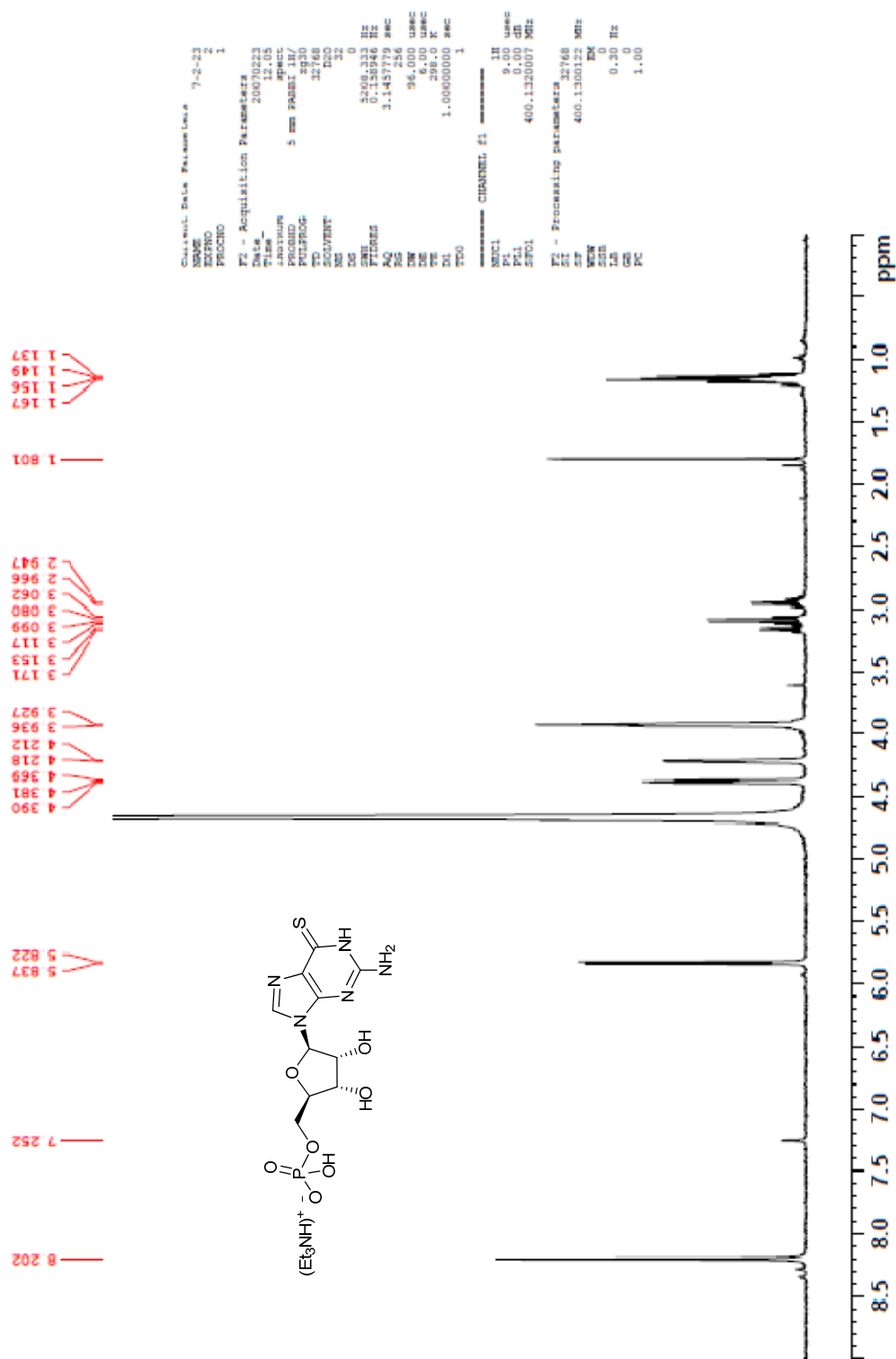


Figure B13: <sup>1</sup>H-NMR; thioguanine monophosphate

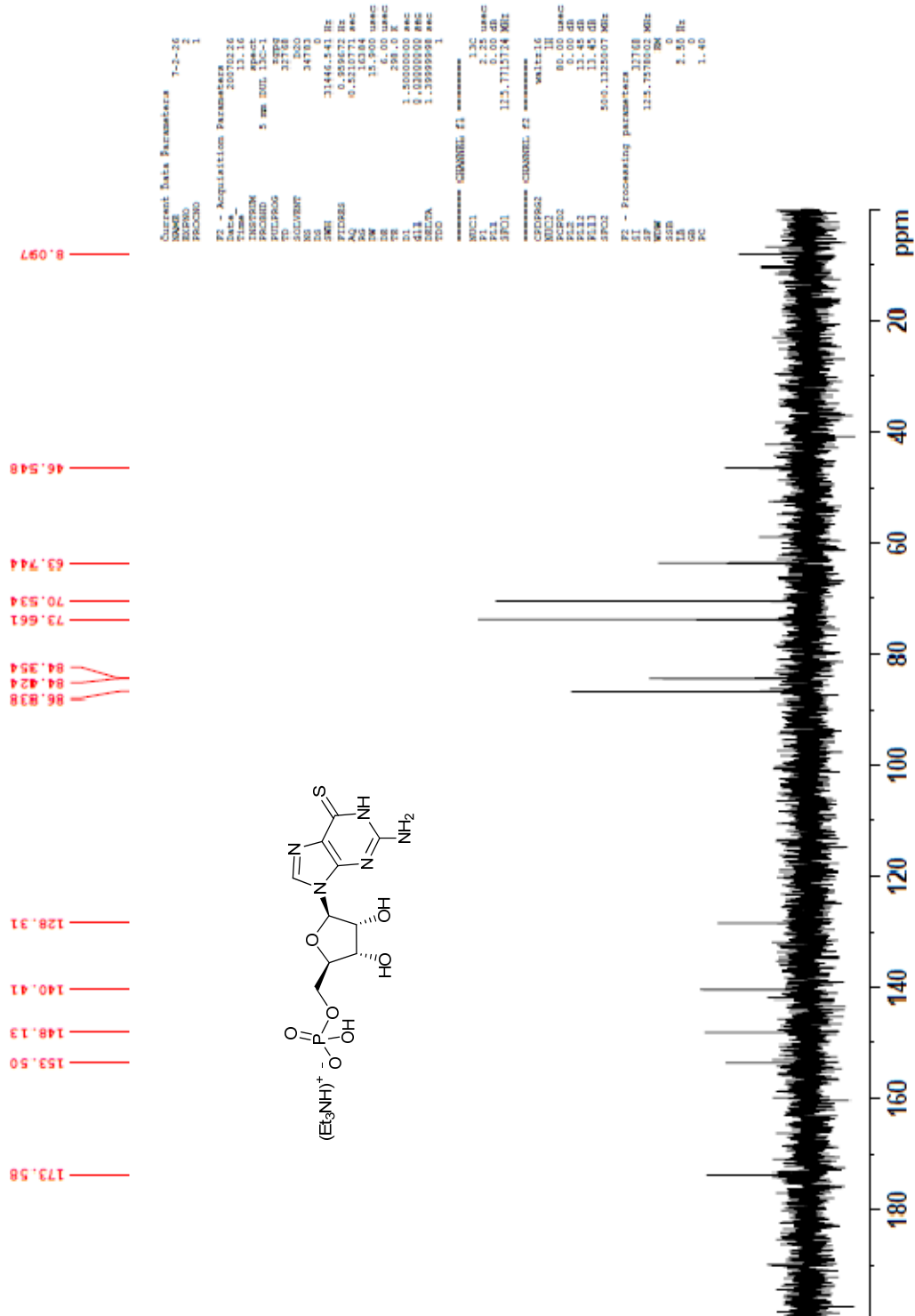


Figure B14: <sup>13</sup>C-NMR; thioguanine monophosphate

APPENDIX C

PURIFIED PRODUCT NMR FROM CHAPTER IV



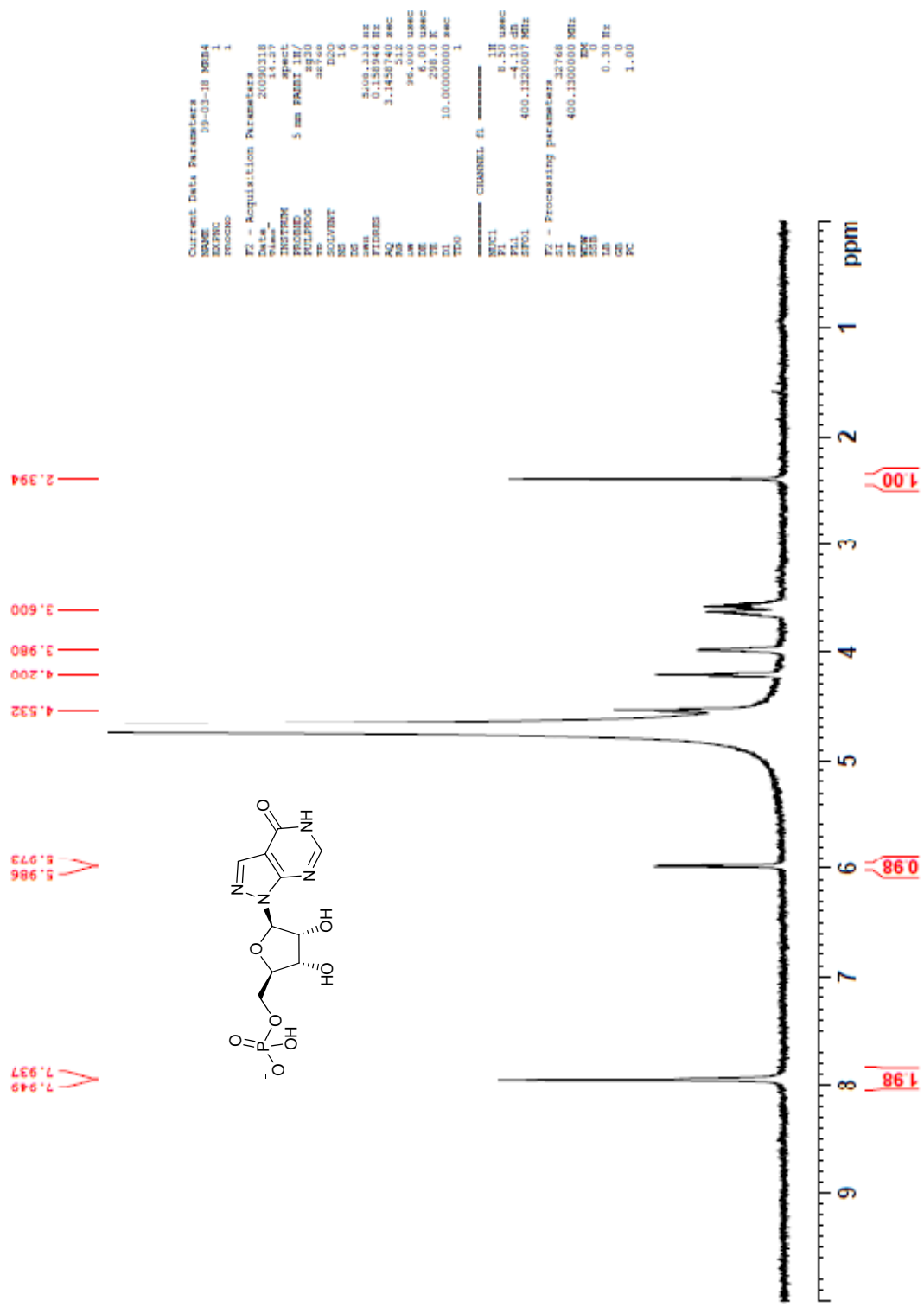


Figure C1: <sup>1</sup>H-NMR; allopurine monophosphate

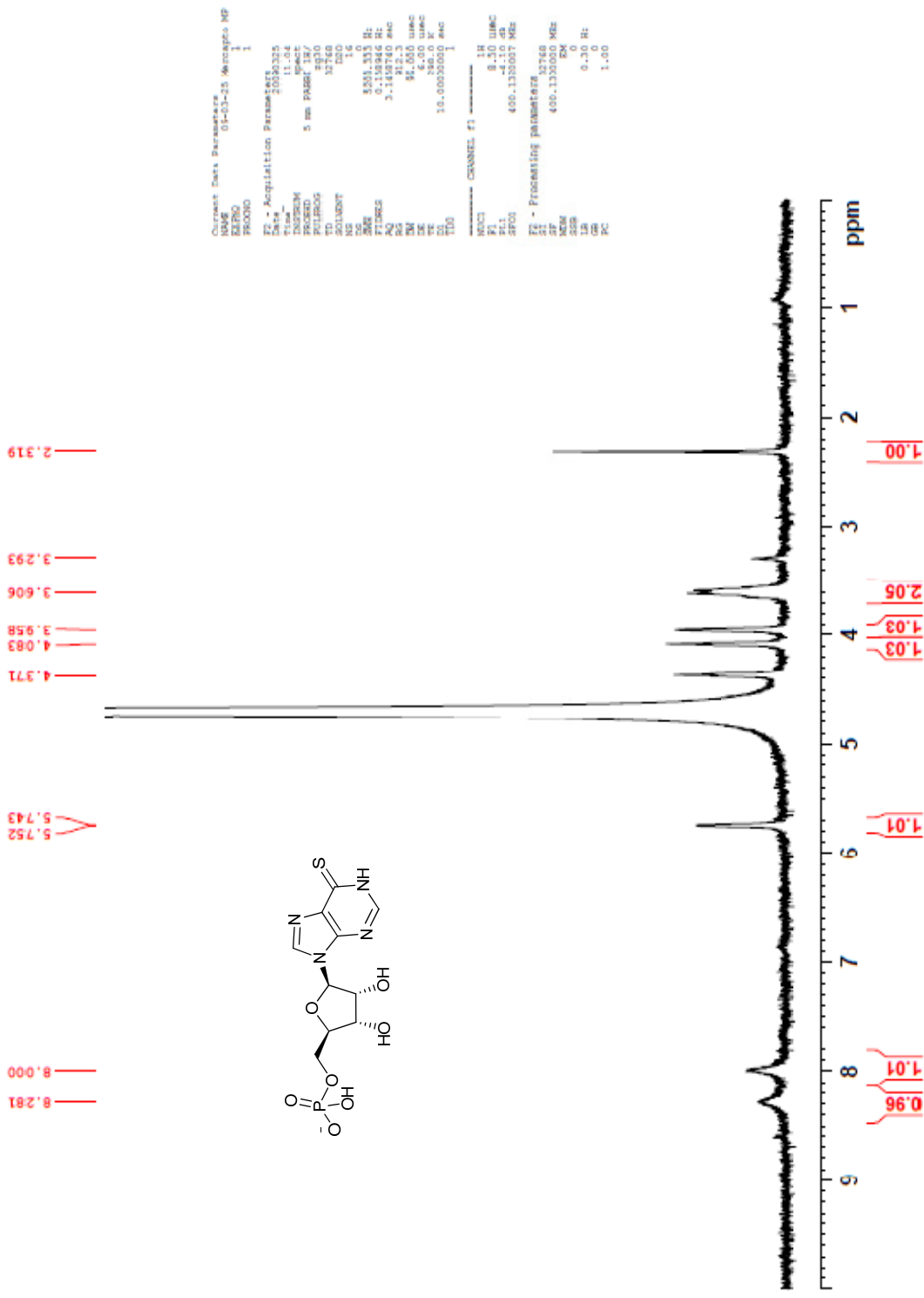


Figure C2: <sup>1</sup>H-NMR; mercaptopurine monophosphate

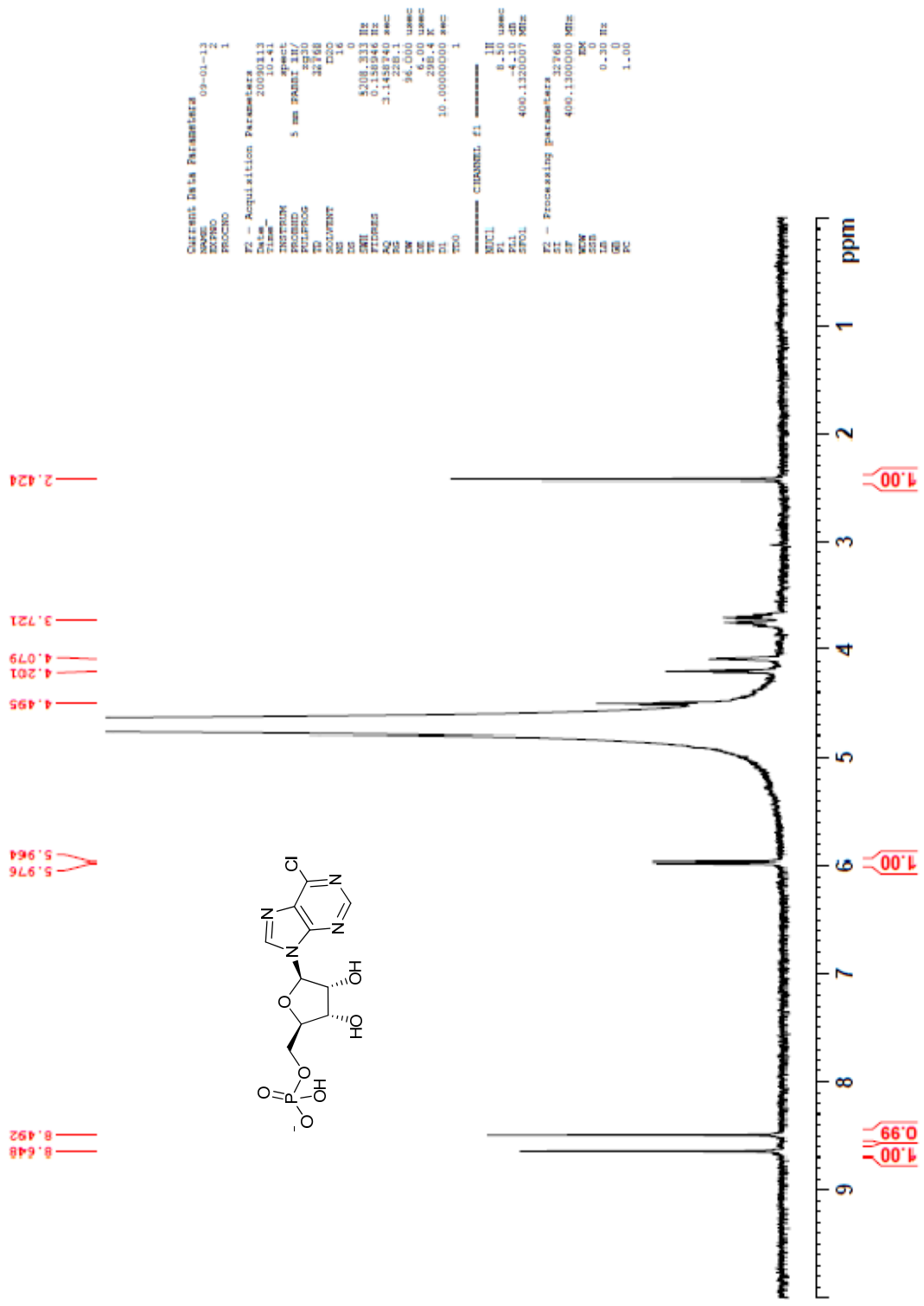


Figure C3: <sup>1</sup>H-NMR; 6-chloropurine monophosphate

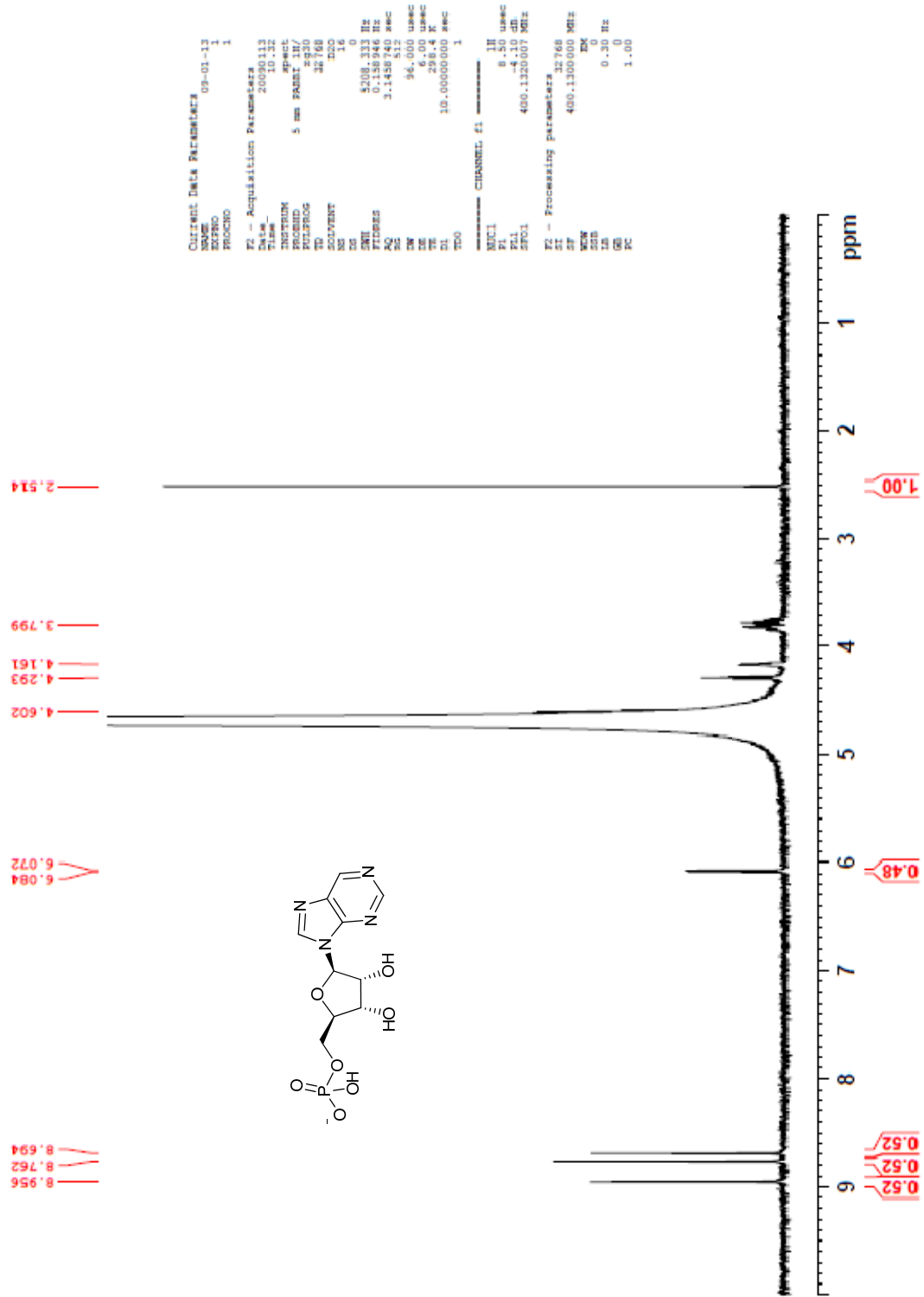


Figure C4: <sup>1</sup>H-NMR; purine monophosphate

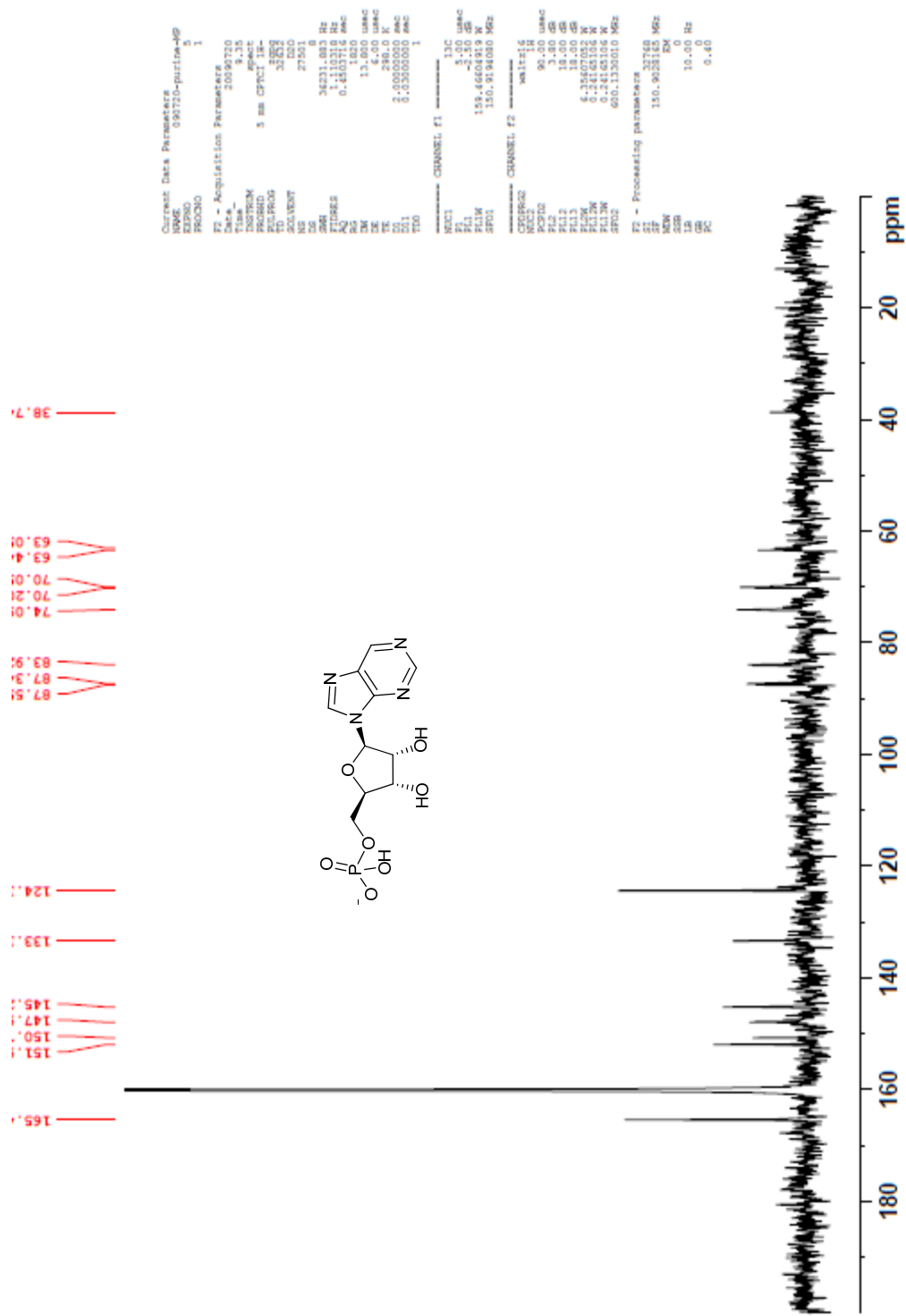


Figure C5: <sup>13</sup>C-NMR; purine monophosphate