REENGINEERING A HUMAN-LIKE URICASE FOR THE TREATMENT OF GOUT

A Dissertation
Presented to
The Academic Faculty

By

James Timothy Kratzer

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry and Biochemistry

Georgia Institute of Technology

August 2013

Copyright © 2013 BY JAMES KRATZER

REENGINEERING A HUMAN-LIKE URICASE FOR THE TREATMENT OF GOUT

Approved by:

Dr. Eric A. Gaucher, Advisor
School of Biology
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Loren D. Williams
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Nick V. Hud
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Eric A. Ortlund
Department of Biochemistry
School of Medicine
Emory University

Dr. Wendy L. Kelly
School of Chemistry and Biochemistry
Georgia Institute of Technology

Date Approved: April 22, 2013

If we knew what we were doing it wouldn't be called

research

would it?

Albert Einstein

To grandma Lois...

For as long as I can remember she would send me handwritten letters for holidays and in recent years she'd write, "Hurry up with your gout drug James because it's getting harder and harder to write these notes." That is until her gout flare-ups took away one of her greatest joys - writing personal correspondence. While she suffered from many ailments in her twilight years, her contraindicated gout seemed to be one of the most debilitating.

ACKNOWLEDGEMENTS

I need to first thank my advisor, Dr. Eric Gaucher, for the wonderful opportunity to learn under his mentorship. He has been a great boss and taught me a great deal about science and about life. I have been very fortunate to find an advisor that has provided the environment to really push myself and to work on such an exciting project My lab mates have been both a support base and a source of lessons learned I especially want to thank those individuals who were instrumental in getting this work done. First, Ryan Randall helped me get started in the lab. She was a constant source of support and help, enabling me to keep my sanity. Both our lab postdocs Dr. Megan Cole and Dr. Betül Kacar really supported me and aided in experimental design and data interpretation.

A big thank you also goes out to those individuals that helped conduct this research. First off is Christina Graves who worked for over a year in our lab on the pseudogene expression aspects of the project. She also trained me in cell culture work. A Georgia Tech rotation student Samit Watve, and a summer REU student from the University of Puerto Rico Valerie Cruz also helped a great deal in collecting kinetic data for some of the uricase variants. I also want to thank Dr. Laura O'Farrell for her guidance, training, and hours of laboratory support that were instrumental in generating animal data.

I am ever so grateful for all of the training opportunities and funding support I have received during my graduate studies at Georgia Tech. First and foremost is the Chemistry department for their support throughout the six years I

have been pursing my degree. I thank Dr. Cam Tyson who has helped me navigate the ups-and-downs and red tape of graduate school.

Thank you to my committee members: Dr. Nick Hud, Dr. Wendy Kelly, Dr. Loren Williams, and Dr. Eric Ortlund. For helping me grow as a scientist and always providing such helpful feedback. I also am grateful for all of the feedback that professor Kelly provided on my dissertation and writing. A special thank you to Dr. Andy Bommarius for his mentorship, and words of encouragement over the years.

Professor Ortlund wears multiple hats and he and his lab really helped me learn how to get our FPLC system up off the ground when I first started. As collaborators Professor Ortlund and Dr. Michael Murphy have provided me with countless hours of feedback and support, as well as rising to the challenge of crystalizing the first mammalian uricase structure to soon be published. We have also been very fortunate to be collaborating with a rheumatologist at the University of Colorado, Boulder Dr. Rick Johnson.

I also am grateful to my fellow graduate students for always being so responsive to lend a few grams of various reagents, to provide access to their lab equipment. I am especially indebted to the Kelly Group for the countless hours I spent on their Spec collecting kinetics data. It helps remind me that we are all in the crazy world of research even though we no longer see each other everyday in class anymore. A special thanks to the crew, which I'm one of the last to finish up: Katy, Lisa, JP, Michael, and Grant.

This work has been supported by a number of funding sources. The School of Biology, a Pre-doctoral fellowship from the Center for Drug Design,

Development and Delivery, the Georgia Research Alliance, a Georgia Tech Research Institute TRIBES grant, and NASA's Astrobiology program, and a Technology Innovation Generating Economic Results (TI:GER) training fellowship. The TI:GER program really opened up my ideas to the broader impact of my research and thanks to all of the support and all I learned from my Team members of PharmEvolve: Yogi Patel, David Weingarten, Seth Griffin, and Alan White.

I also want to thank my wife Stacia because she has always been my number #1 cheerleader and she has always stuck by my side. I'll be making it up to her for countless years to come. Finally, I thank my parents, sister Michelle, and my in-laws for being ever so supportive, keeping me out of debt and well fed.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

LIST OF TABLES ix

LIST OF FIGURES x

LIST OF SYMBOLS AND ABBREVIATIONS xiv

LIST OF SEQUENCES xvii

LIST OF VENDORS xviii

CHAPTER 1: INTRODUCTION 1

1.1 Purine catabolism and causes of elevated uric acid 1

1.2 Gout and other medical conditions of elevated uric acid 5

1.3 Unmet medical need in management of gout 10

1.4 Properties of uricase, its reaction, and evolutionary history 11

1.5 Challenges in therapeutic uricases 16

1.6 Scope of this work 18

1.7 References 21

CHAPTER 2: EXPLORING THE HUMAN URICASE PSEUDOGENE 26

2.1 INTRODUCTION 26

2.1.1 The properties of the human uricase pseudogene 27
2.1.2 Potential advantages of uricase inactivation 29
2.1.3 The fate of the uricase pseudogene is therapeutically important 30
2.2 MATERIALS AND METHODS 31

2.2.1 Transcriptional Studies 31
2.2.2 Translational Studies 31
2.3 RESULTS AND DISCUSSION 38

2.3.1 Evidence of *hUox* transcript in human tissue 38
2.3.2 Human cells can express full-length hUox 40

2.4 CONCLUSIONS 45

2.5 REFERENCES 47

CHAPTER 3: INDENTIFICATION OF AN ACTIVE HUMAN-LIKE URICASE 50

3.1 INTRODUCTION 50

3.1.1 Ancestral Sequence Reconstruction (ASR) overview 51
3.1.2 Successful examples of ASR application 55
3.2 MATERIALS AND METHODS 56

3.2.1 Direct reactivation of the human uricase pseudogene 56
3.2.2 Computational inference of ancient uricases 57
3.2.3 Uricase expression and purification 58
3.2.4 Measuring uricase enzymatic activity 61
3.2.5 Uricase storage stability assays 62
3.2.6 Making uricase variants based upon the inferred ancestral sequences 62
3.2.7 Uricase structural modeling and solving mammalian uricase crystal structure. 63
3.3 RESULTS AND DISCUSSION 64

3.3.1 Development of uricase controls 64
3.3.2 Human uricase has accumulated deleterious mutations 66
3.3.3 Resurrecting ancestral uricases 67
3.3.4 The structure of mammalian uricases 73
3.3.5 Properties of chimeric uricases 81
3.4 CONCLUSIONS 83

3.5 REFERENCES 85

CHAPTER 4: OPTIMIZING A HUMAN-LIKE URICASE 89

4.1 INTRODUCTION 89

4.2 MATERIALS AND METHODS 89

4.2.1 Synthesis of branch mutants 89
4.2.2 Protein expression and purification 90
4.2.3 Enzymatic activity assay 90
4.2.4 Quantifying effect of individual branch mutations 90
4.2.5 Humanizing ancestral uricase 91
4.3 RESULTS AND DISCUSSION 91

4.3.1 An19/22-Human chimeras: introducing solvent-accessible human residues into the An19/22 background 91

4.3.2 Exhaustively testing the effects of individual branch mutations within the An19/22 background 96
4.3.3 Human residue that kills the activity of An19/22 98
4.3.4 Properties of "human-like" uricases 101
4.4 CONCLUSIONS 103

4.5 REFERENCES 104

CHAPTER 5: Modifying and testing uricases *in vivo* 105

5.1 INTRODUCTION 105

5.2 MATERIALS AND METHODS 108

5.2.1 PEGylation of "human-like" uricases 108
5.2.2 Pharmacokinetics experiments with unmodified and PEGylated uricases 110
5.3 RESULTS AND DISCUSSION 112

5.3.1 The effects of PEGylation on the uricases 112
5.3.2 Pharmacokinetic data of unmodified and PEGylated uricases 120
5.4 CONCLUSIONS 124

5.5 REFERENCES 126

CHAPTER 6: FUTURE DIRECTIONS AND CONCLUDING REMARKS 129

6.1 FUTURE DIRECTIONS 129

6.1.1 Going after the expressed human "pseudogene" 129
6.1.2 Optimization of PEGylation strategy 129
6.1.3 Testing therapeutic effect 131
6.1.4 Determining the safety of uricases 131
6.1.5 Explore the co-administration of uricase's downstream enzymes 132
6.2 CONCLUDING REMARKS 132

6.3 REFERENCES 135

SUPPORTING MATERIAL 136

PUBLICATIONS 147

REFERENCES 148

VITA 161

LIST OF TABLES

Table 1.1 Several uricase sources have been under clinical development. 18

Table 3.1 The age of and accuracy of inferred uricase sequences.....70

Table 3.2 Summary of modern and ancient uricases. 81

Table 3.3 Using An19/22 as a stable backbone for exploring non-humanizing mutations.....83

Table 4.1 Summary of An19/22-Human chimeras kinetics.....95

Table 4.3 The stable An19/22 and An26 will accommodate select subsets of human residues while retaining its specific enzyme activity. 103

LIST OF FIGURES

Figure 1.1 Purines are heterocyclic nitrogenous compounds.....2

Figure 1.2 Nitrogenous waste is excreted from the body in a form whose complexity is dictated by the host organism's purine metabolism.....4

Figure 1.3 Gout causes intense pain and swelling 6

Figure 1.4. Uricase catalyzes a cofactor-less oxidation reaction 12

Figure 1.5 Uricase is functional as a homotetramer with active sites at the dimer-dimer interfaces.....14

Figure 2.1 The pseudogenization of human uricase is the result of three genetic lesions 28

Figure 2.2 Uricase transcripts are present in human tissue 39

Figure 2.3 Constructs synthesized to test effects of stop codons on hUox expression in humans. 42

Figure 2.4 The rabbit polyclonal antibody against mouse uricase displayed poor selectivity for hUox detection. 43

Figure 2.5 Western blot detection of full-length human uricase expression in 293T cells. 44

Figure 3.1 Ancestral sequence reconstruction allows researchers to travel back in time and explore ancient proteins. 52

Figure 3.2 Phylogenetic tree highlighting distribution and relationship of modern and ancient sequences. 54

Figure 3.3 Tetrameric uricase is obtained by SEC. 65

Figure 3.4 The activity of uricase proteins was assessed using UV-Vis spectrophotometry. 66
Figure 3.5 Phylogram of uricase genes constructed by the maximum likelihood method 68
Figure 3.6 Phylogenetic tree showing the ancestral sequence space explored. .69
Figure 3.7 Enzymatically active mammalian uricase is trapped within the inclusion body produced during overexpression in *E. coli.* 71
Figure 3.8 More recent ancestors are the least soluble under experimental conditions 72
Figure 3.9 Solved mammalian crystal structre shows structural similarity to published microbial structures 74
Figure 3.10 An19/22 is the common ancestor of placental mammals. 75
Figure 3.11 An19/22 extraction efficiency and storage buffer stability. 76
Figure 3.12 An26 is the common ancestor of primates. 77
Figure 3.13 An 27: The common ancestor of the Old World monkeys and hominoids. 77
Figure 3.14 An30 the common ancestor of apes is highly unstable. 78
Figure 3.15 Soluble activity of purified tetrameric uricase 79
Figure 4.1 Dissecting branch changes: An19/22_19-26 93
Figure 4.2 Dissecting branch changes: An19/22_31-32 94
Figure 4.3 Dissecting branch changes: An19/22_32-H. 95
Figure 4.4 Individual branch mutations in the stable An19/22 showed that single human residues can completely kill activity. 97

mutations. Figure 4.5 Deactivating mutation S232L is only present in the human pseudogene 99
Figure 4.6 Deleterious mutation Y240C arose in the last common ancestor of gorilla, chimpanzee, and human 100
. Figure 4.7 Deleterious mutation F222S arose in the common ancestor of hominoids..... 100
Figure 4.8 The G83E mutation is located at the packing interface of uricase tetramers. 101
Figure 5.1 12 % SDS-PAGE gel can only differentiate between reacted and unreacted protein samples..... 113
Figure 5.2 A 4 % Native PAGE gel run with CAPS-NH_4OH running buffer shows a better separation between PEG-protein species. 115
Figure 5.3 Extent of PEGylation can be modulated by altering the molar excesses of activated PEG relative to the protein's surface lysines available for conjugation. 116
Figure 5.4 The PEGylation reactions reach completion after a single hour. 117
Figure 5.5 PEGylation improves uricase activity at sub-optimal pH. 118
Figure 5.6 Specific activities of crude PEGylation reactions 119
Figure 5.7 Various PEGylation strategies effect on the SEA An19/22 uricase at physiological pH 120
Figure 5.8. In healthy rats unmodified An19/22 displayed enhanced pharmacokinetics over the PBC uricase. 121
Figure 5.9 PEGylating an ancestral uricase improves its circulation time in healthy rats..... 123

Figure S1 DNA multiple sequence alignment generated for ancestral sequence reconstruction 137
Figure S2 A robust phylogeny was generated with MrBayes. 141

LIST OF SYMBOLS AND ABBREVIATIONS

-EXON 3	*hUox* mammalian expression vector missing exon 3 and internal stop-codons are mutated to arginines
3-STOP	*hUox* mammalian expression vector with strong terminal stop-codon
2-STOP	hUox mammalian expression vector with 2 naturally-occurring stop-codons
6xHis	Hexahistidine tag
Å	Ängstroms
A. flavus	*Aspergillus flavus*
A280	Absorbance at 280 nanometers
A293	Absorbance at 293 nanometers
AEX	Anion exchange
An	Ancestor
An19/22	Uricase from common ancestor of placental mammals
An19/22Med	An19/22 uricase construct with a subset of individual branch mutations that had a neutral to mildly deleterious impact on its biomolecular properties
An19/22Plus	An19/22 uricase construct with a subset of individual branch mutations that had a neutral to advantageous impact on its biomolecular properties
An19/22_19-26	An19/22 uricase construct with select surface mutations present in the branch from An19/22 to An26
An19/22_31-32	An19/22 uricase construct with select surface mutations present in the branch from An31 to An32/33
An19/22_32-H	An19/22 uricase construct with select surface mutations present in the branch from An32/33 to the inactive human uricase
An19/22_G.max	An19/22 uricase construct with mutations from the soybean (*Glycine max*) uricase
An19/22_LysI	An19/22 uricase construct with 2 surface lysines
An19/22_LysII	An19/22 uricase construct with 3 surface lysines
An19/22_JPO	An19/22 uricase construct with mutations reported to improve the activity of bacterial uricases
An19/22_USPTO	An19/22 uricase construct with mutations reported to improve the stability of the *Candida utilis* uricase
An26	Uricase from the common ancestor of primates
An26Plus	An26 uricase construct with a single advantageous branch mutation

An27	Uricase from the common ancestor of the old world monkeys and hominoids
An30	Uricase from the common ancestor of greater and lesser apes
An31	Uricase from the common ancestor of the great apes
An32/33	Uricase from the common ancestor of the chimpanzee and human
ARF	Acute renal failure
ASR	Ancestral sequence reconstruction
ATP	Adenosine triphosphate
BB	Bugbuster non-ionic cell lysis detergent
Da	Dalton
dH_2O	Deionized water
DNA	Deoxyribonucleic acid
ESR	Electron spin resonance
g	Gram
GTP	Guanosine 5'-triphosphate
HEK	Human embryonic kidney cells
HIU	5-Hydroxyisourate
HO-PEG	Hydroxy PEG
hUox	Human uricase protein
hUox	Human uricase pseudogene
IMAC	Immobilized metal affinity chromatography
IB	Inclusion body
IU	International unit
IV	Intravenous
KHK	Fructokinase
K_{SA}	Number of surface (i.e. solvent accessible lysine residues) in target protein
LB	Luria broth
MCMCMC	Metropolis coupled Markov chain Monte Carlo
ME	Molar excess
mL	Milliliter
M_n	Average molecular weight (g/mol) of activated PEG determined by gel filtration chromatography
M_r	Molar ratio of activated PEG to target protein
mPEG	Methoxy-PEG
MW	Molecular weight
Mya	Millions of years ago
N	Degree of PEGylation
NAD^+	Nicotianmide adenine dinucleotide
ng	Nanogram
NMR	Nuclear magnetic resonance

NO-STOP	*hUox* mammalian expression vector where internal stop-codons are changed to arginines
OHCU	2-Oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline
P	Mass (mg) of target protein to react in PEGylation reactions
pAb	Polyclonal antibody
PAML	Phylogenetic analysis by maximum likelihood
PBC	Pig-baboon chimeric uricase
PBS	Phosphate buffered saline
PBST	PBS with 0.08 % Tween-20™
PBSTM	PBST with 5 % (w/v) powdered milk
PEG	Polyethylene glycol
PEG-NPC	Nitrophenyl carbonate activated polyethylene glycol
PEG-SG	Succinimidyl glutarate activated polyethylene glycol
PEG-SVA	Succinimidyl valerate activated polyethylene glycol
pI	Isoelectric point
PNP	Purine nucleoside phosphorylase
R5P	Ribose-5'-phosphate
RFU	Relative fluorescent unit
RNA	Ribonucleic acid
SASA	Solvent accessible surface area
SD	Sprague Dawley
SDM	Site directed mutagenesis
SDS-PAGE	Sodium dodecyl sulfate
SEA	Specific enzyme activity
SEC	Size exclusion chromatography
S_{NMR}	Degree of substitution (%) of activated PEG determined by NMR
T-fold	Tunneling fold
TFG	Treatment failure gout
TLS	Tumor lysis syndrome
Trx	Thioredoxin
Uox	Uricase gene
Uox	Uricase enzyme
UV	Ultraviolet
α	Alpha helix
β	Beta sheet

LIST OF SEQUENCES

Sequence S1 Primers used to query cDNA library for hUox.....142

Sequence S2 Protein sequences of functional uricases used for ASR 143

Sequence S3 Inferred ancestral uricase protein sequences. 145

LIST OF VENDORS

Agilent Technologies (Santa Clara, CA)
BD (Franklin Lakes, NJ)
Beckman Coulter (Brea, CA)
Cell Biolabs (San Diego, CA)
Charles River Laboratories (St. Wilmington, MA)
CloneTech (Mountain View, CA)
EMD Millipore (Billerica, MA)
Emerson (Danbury, CT)
Formulatrix (Waltham, MA)
GE Healthcare (Waukesha, WI)
GenScript (Piscataway, NJ)
Integrated DNA Technologies (Coralville, IA)
Invitrogen (Carlsbad, CA)
Laysan Bio Inc (Arab, AL)
OriginLab (Northampton, MA)
PALL (Port Washington, NY)
Pierce (Rockford, IL)
Promega (Madison, WI)
Qiagen (Germantown, MD)
Rigaku (The Woodlands, TX)
Santa Cruz Biotechnology (Dallas, TX)
Sigma-Aldrich (St. Louis, MO)
Starna Cells (Atascadero, CA)
Thomson Reuters (New York, NY)
UVP (Upland, CA)

SUMMARY

We are interested in uricase for two key reasons. The first is to elucidate the evolutionary course of its inactivation in humans (and the great apes). The second is to develop a novel uricase therapy that will uniquely meet the needs of the biomedical community. Perhaps an appropriate subtitle for this dissertation research is: *Uricase: An enzyme whose inactivation may have enhanced our ancestors' fitness but now contributes to human disease*. This research encompasses the emerging field of evolutionary synthetic biology. This interdisciplinary field brings together an evolutionary understanding of how biomolecules change over time and adapts those biomolecules for various utilitarian applications. Considering those changes that have already been tested by nature focuses our efforts in engineering desired biomolecular properties. My main research project surrounds the development of a uricase for the management of gout and uric acid levels.

Uric acid is a natural product of DNA breakdown and is normally excreted from the body. However, at elevated levels, uric acid may form sharp crystals within the joints, resulting in swelling and inflammation characteristic of gouty arthritis. This disease afflicts more than 10 million patients worldwide, and its incidence is on the rise. Most concerning, is the substantial population of treatment-failure gout patients that cannot benefit from or tolerate small molecule treatments. My research in developing a uricase for breaking down uric acid specifically addresses the unmet needs in gout and other uric-acid related diseases.

To make a more "human-like," and therefore safer, uricase we employ the evolutionary biology approach of ancestral sequence reconstruction (ASR) to reengineer a functional human-like uricase. In short, this approach allows us to experimentally "resurrect" ancient proteins and explore functionality with the uricase family, thereby combining functionality across modern day uricases. I have experimentally synthesized, purified, and characterized a number of ancestral uricases.

Our collaborators at Emory University, Dr. Eric Ortlund and Dr. Michael Murphy, have solved the crystal structure of uricase An19/22 (no mammalian uricase crystal structures have yet been published). Through kinetic assays, ancestral uricase An19/22 was identified as being the most catalytically active of those explored, and complete activity was abolished in the more recent ancestral uricases, An30 onward, Excitingly, An19/22 uricase also displayed improved *in vivo* stability when tested head-to-head with Krystexxa®, the FDA-approved uricase, in healthy rats.

CHAPTER 1: INTRODUCTION

This research approaches the challenge of developing a novel protein for the treatment of gout by applying an understanding of how proteins change over time, and utilizing this information to direct engineering efforts towards functional variants. In essence, it is the marriage of the historically disparate studies of evolutionary biology with synthetic biology into a new evolutionary synthetic biology. The new field can be conceptualized by considering two great minds of the 20th century. The prominent geneticist and evolutionary biologist Theodore Dobzhansky stresses the intractability of biology with evolution, "Nothing in Biology makes sense except in the light of evolution [1]." Therefore, we approach biomolecular engineering challenges by considering the evolutionary forces that have driven functional differences among extant or modern-day proteins. The synthetic biology aspect of this work can be conceptualized by the following quote attributed to theoretical physicist Richard Feynman, "What I cannot create, I do not understand ". By bringing an evolutionary prospective to synthetic biology problems nature can guide the rational design of macromolecules for a number of utilitarian applications.

1.1 Purine catabolism and causes of elevated uric acid

Purines are nitrogenous heterocyclic aromatic compounds comprised of a pyrimidine ring bound to an imidazole ring. These biomolecules play vital roles in the body and are synthesized *de novo*, obtained from exogenous sources, or recovered from salvage pathways. Shown are the structures of several purine

nucleosides, where the nucleobase is bound to the sugar ribose at the 9-position (Figure 1.1). Purines perform many biological roles including cellular energy systems in the form of adenosine triphosphate (ATP), to signaling in the form of guanosine triphosphate (GTP). Together adenosine monophosphate (AMP) and guanosine monophosphate (GMP) comprise roughly half of the genetic storage molecules: deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) [1]. Inosine monophosphate (IMP) is commonly found in transfer RNA (tRNA), and is essential for translation of wobble base pairs [2]. The fourth purine nucleoside shown is xanthosine monophosphate (XMP), which is an intermediate in purine metabolism, and can be formed from IMP by IMP dehydrogenase (Figure 1.2).

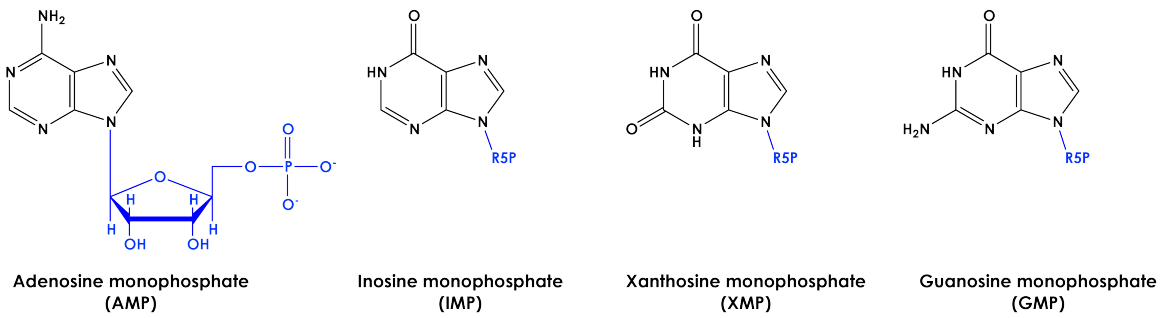


Figure 1.1 Purines are heterocyclic nitrogenous compounds. The four purine nucleosides are adenosine monophosphate (AMP), inosine monophosphate (IMP), xanthosine monophosphate (XMP), and guanosine monophosphate (GMP). Ribose-5'-phosphate (R5P) is shown in blue in each structure.

The breakdown of purines is a multistep process by which organisms remove nitrogenous waste. AMP enters this pathway as a waste product of ATP depletion or nucleic acid breakdown. GMP is the other purine found in DNA and RNA. AMP is converted to IMP by AMP deaminase and IMP is converted to XMP by IMP dehydrogenase. The 5'-phosphate is removed from each purine nucleoside's ribose by a nucleotidase. Adenosine is converted to inosine by the action of adenosine deaminase. The ribose sugar is removed from purine bases

by purine nucleoside phosphorylase (PRP): when inosine is the substrate hypoxanthine is the product of this reaction, and when guanosine is the substrate the reaction produces guanine. Xanthine is the direct product of PRP action upon xanthosine. The subsequent oxidation by xanthine oxidase yields uric acid (UA). The first acid dissociation constant of UA is 5.4; therefore, at physiological pH, UA exists as the singly charged monoanion urate [3]. Urate is only modestly soluble at neutral pH, and therefore uses minimal water as the nitrogenous waste excreted by humans, hominoid primates, birds, reptiles, and terrestrial insects [4]. Uric acid's insolubility at lower pH can explain why patients that suffer from urate stones often have urine that is more acidic, as low as pH 4.6, than their healthy peers [5].

In most monkeys, and in carnivorous flies, urate is converted by three enzymes to yield *S*-allantoin, which is an order of magnitude more soluble than urate[6]. It was originally thought that uricase catalyzed the direct conversion of uric acid to allantoin. However, in living systems only the *S*-enantiomer of allantoin is found. This stereoselectivity implies that the reaction is enzyme catalyzed [7]. To identify the enzymes responsible for the stereospecific conversion of uric acid to *S*-allantoin, researchers performed a comparative phylogenetic analysis on gene families of previously unknown function that were exclusively present in organisms with a functional uricase and absent in organisms lacking a functional uricase [8].

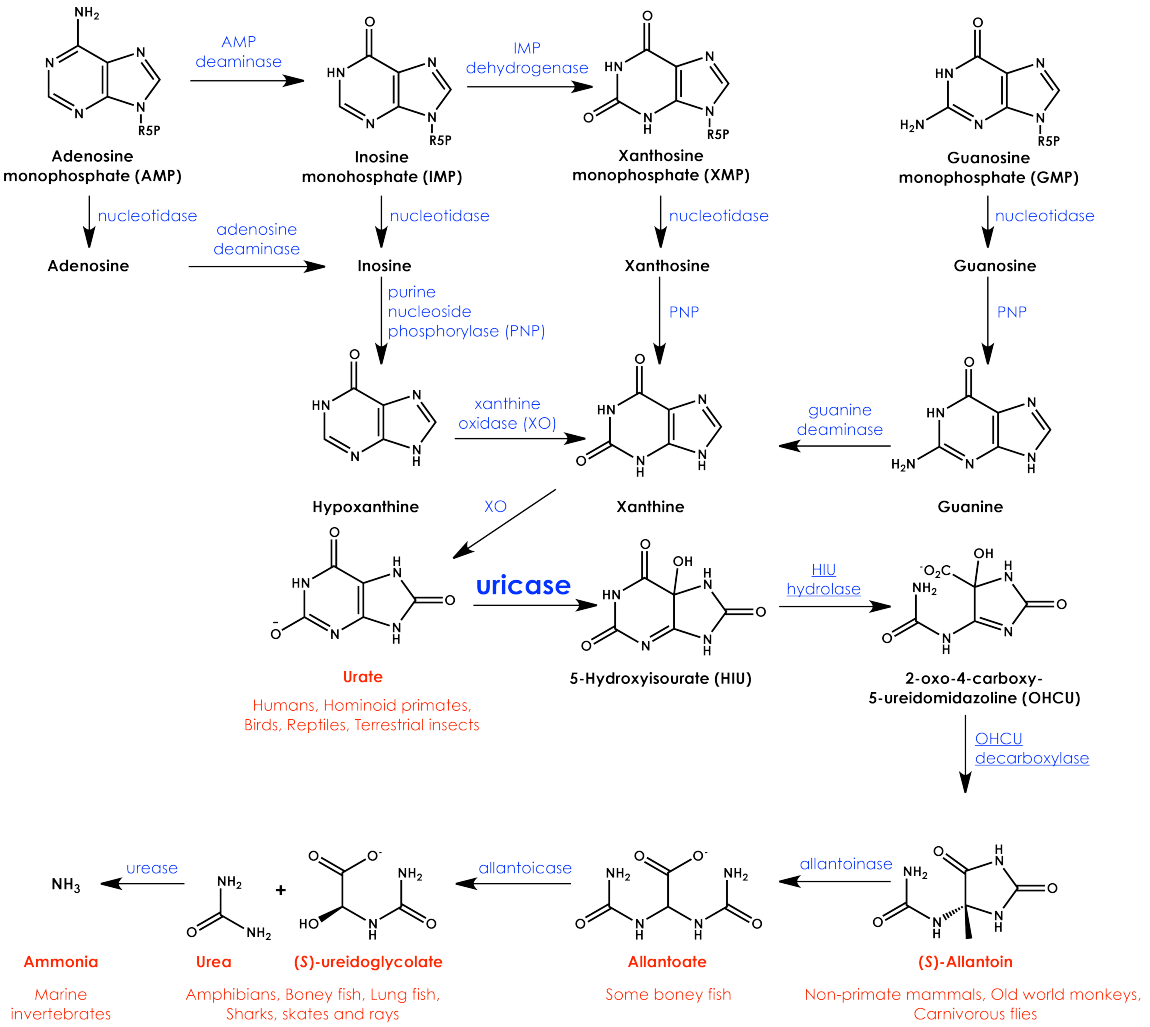


Figure 1.2 Nitrogenous waste is excreted from the body in a form whose complexity is dictated by the host organism's purine metabolism. The enzyme responsible for each metabolic step is shown in blue and each product is shown in bold. The form by which nitrogenous waste is excreted from the body in certain taxa is shown in red. Purines enter this breakdown pathway in four forms (Ribose-5-phosphate is abbreviated as R5P). The first is as adenosine monophosphate (AMP), which is generated by either nucleic acid breakdown or ATP depletion. AMP can be deaminated to form inosine monophosphate (IMP). The remaining two purines are xanthosine monophosphate (XMP), which is generated by the oxidation of IMP, and guanosine monophosphate (GMP), which is liberated by nucleic acid breakdown. The first step involves the removal of phosphate from each of the monophosphates by a nucleotidase to yield the nucleoside, and adenosine is deaminated to yield inosine. The next common enzymatic step is through the action of purine nucleoside phosphorylase (PNP). Ribose is removed to yield the free purine base: Hypoxanthine from inosine which is converted to xanthine by xanthine oxidase (XO), xanthine from xanthosine, and guanine from guanosine. Next, hypoxanthine is converted to xanthine by an oxidase. Guanosine is converted to xanthine by a deaminase. Xanthine is oxidized by XO to yield urate, the terminal purine waste product in humans. The

(Figure 1.2 continued)
enzyme uricase oxidizes urate to produce 5-hydroxyisoruate (HIU). Two subsequent enzymatic steps - first by HIU hydrolase and then OHCU decarboxylase yield (S)-allantoin, which is the purine waste excreted by non-primate mammals. Next, allantoinase breaks down (*S*)-allantoin to yield allantoate, which is excreted by some boney fish. In amphibians, the remainder of boney fishes, sharks, rays and skates allantoincase produces urea and (S)-ureidoglycolate. The least chemically complex nitrogenous waste is ammonia generated by the breakdown of urea by urease in marine invertebrates.

To determine the function of these genes, researchers overexpressed the mouse homologs *MuraH* and *MuraD*, and their enzymatic products were characterized by spectroscopic and NMR analyses. It was confirmed that *MuraH* was a HIU hydrolase and *MuraD* is an OHCU decarboxylase, respectively [8]. First, urate is converted to 5-hydroxyisourate (HIU) by uricase (discussed in greater detail in section 1.4). HIU is then converted by a hydrolase to OHCU, which is subsequently decarboxylated to form *S*-allantoin. In organisms that have access to unlimited water, purines are broken down into even simpler molecules. *S*-allantoin is converted to allantoate by allantoinase, the terminal waste product of some boney fish. In amphibians, most boney fish, dipnoans, and elasmobranchs, allantoate is broken down to urea and *S*-ureidoglycolate by allantoinase. The simplest nitrogenous waste is ammonia, which is generated through the action of urease upon urea in marine invertebrates.

1.2 Gout and other medical conditions of elevated uric acid

Gout is a painful form of inflammatory arthritis that is caused when uric acid, a natural by-product of DNA breakdown, reaches levels that exceed what the body can excrete. The excess uric acid forms crystals of monosodium urate within the joints and results in swelling and inflammation (Figure 1.3)

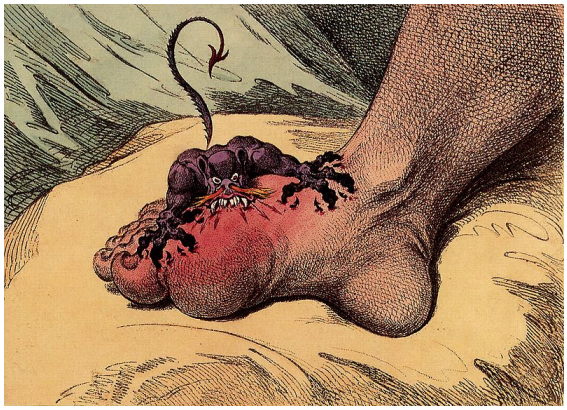


Figure 1.3 Gout causes intense pain and swelling. This caricature was produced in 1779 by James Gilray and its depiction of gout afflicting the big toe as a biting, fire-breathing, demon rings true to patients' descriptions as the worst pain they have ever experienced.

The study of gout is as old as humans' study of the body and the history of medicine. The first written description of gout is believed to be over 4,000 years old from the Egyptian civilization. The great philosopher Socrates was reported to have referred to the affliction as the "unwalkable disease". In 200 AD, the Roman physician Galen described the disfiguring swelling, or tophi, caused by the deposition of monosodium urate monohydrate crystals in the avascular tissue. It was not until the late 17th century that the invention of microscopy allowed Leeuwenhoek to observe monosodium urate crystals in gouty tissue. By the mid 19th century in *The Nature and Treatment of Gout and Rheumatic Gout,* Sir Alfred Baring Garrod's asserts that the deposition of urate is the cause and not the effect of gouty inflammation [9]. In 1960, George Hitchings and Gertrude Ellion developed allopurinol, which is a small molecule inhibitor of xanthine oxidase – they were awarded the Nobel Prize in 1988 in recognition for this work. Allopurinol served as the standard-of-care for gout management until a boon in gout management research over the last decade [10]. To address the specific needs of those that do not benefit from conventional treatments, research has

really picked up in the management of gout. Febuxostat, is a derivative of allopurinol that was approved in 2009, and became the first FDA approved gout medication in 40 years [11]. In recent years, attention has turned to enzymatic treatments, such as the recent FDA approval of Krystexxa® in 2010 (discussed later).

The saturation point of urate in biological fluids is 6.8 mg/dL and, clinically, patients above this level are described as hyperuricemic. While patients that manifest gouty arthritis are always hyperuricemic, the fact that a patient has elevated urate levels is not enough to predict that a gouty flare-up will occur. A normative aging study found that within a 5-year period only 22 % of men with high urate levels, greater than 9 mg/dL, ever developed gout [12]. These epidemiological studies show that uric acid levels alone are not the whole story. There are other local factors that contribute to the development of gout, including trauma or irritation, reduced temperature, and prior joint disease [13]. The locations of the body most commonly afflicted with gouty arthritis can be explained by these various factors. For example, the joint of the big toe and that of the elbow are common sites of gouty arthritis because the former is a site of mechanical stress and lower temperature, while a patient leaning on the elbow joint can mechanically irritate the later. Patients suffering from osteoarthritis in their fingers often must also deal with the added pain of gout in those same joints. It is believed that the release of urate crystals into the joint space is responsible for an acute inflammatory reaction – that is why urate-lowering, treatments are often accompanied by an acute gout attack. The synovial lining cells appear to phagocytize monosodium urate crystals when they are released

into the joint space. Next, an inflammasome complex is formed and releases IL-1β along with other cytokines and proinflammatory molecules – these chemotaxic agents bring an influx of neutrophils into the joint [14]. Interestingly, acute gout attacks can resolve spontaneously – this phenomenon is credited to the dissolution of the crystals or their sequestration within the tissue and some proteins that are exuded into the joint space with the attack. Apolipoprotein B is an example of a protein that can coat crystals and reduce their inflammatory properties [15].

There are several trends associated to the increasing prevalence of patients showing elevated levels of uric acid. These include a shift in eating habits from developed countries with foods high in purines and fructose. Furthermore, 90 % of uric acid that is filtered by the kidneys is reabsorbed instead of being excreted. This suggests that the human body is not treating the presence of uric acid as toxic, but rather as something that is advantageous to retain in the biological fluid [56]. The consumption of certain foods and beverages can lead to an increase in purines being broken down to uric acid. This ties into the historical moniker of gout as the "rich-man" disease, because foods such as organ meats, certain seafood and alcoholic beverages (especially beer) have high purine content [16]. Another underappreciated source of uric acid may be from the consumption of foods that are high in fructose (especially high fructose corn syrup and sucrose which are such a large part of our diets) whose metabolism was hypothesized to result in increases in uric acid concentrations [17].

Natural cell turnover also leads to release of DNA (roughly half of which is purine DNA) that will be broken down. An extreme case of this outpouring of "waste DNA" is experienced by cancer patients in a condition known as tumor lysis syndrome (TLS) [18]. This complication is of greatest concern in malignancies such as leukemias where there is a large tumor burden. When patients undergo aggressive chemotherapy, the large outpouring of tumor DNA can lead to immediate and dramatic increases in uric acid concentrations above its solubility limit, which, if prophylactic measures are not employed, will lead to acute renal failure (ARF) and necessitate dialysis before anticancer therapy can be resumed [19]. Furthermore, the healthy level of uric acid is maintained through the filtration by the kidneys and any impairment of kidney function, in turn, diminishes the body's ability to excrete UA [20]. Lastly, recent studies have begun to explore the genetic component that seems to predispose individuals to gout [21].

1.3 Unmet medical need in management of gout

Conventional treatment in the management of gout has approached the issue using two distinct avenues. The first approach is to promote the excretion of the poorly soluble urate with a uricosuric agent. The only FDA approved uricosuric is probenecid, which requires twice-a-day dosing for optimal therapeutic effect thus often resulting in patient compliance challenges [22]. Probenicid's therapeutic action is completely negated when glomular filtration rates drop below 50 mL/min – especially in older gout patients whose kidney function is often compromised [23]. Furthermore, there is a risk of central nervous system toxicity at higher doses of probenecid [24]. The second avenue of treatment is to block the generation of urate. The standard-of-care for gout management for the last 50 years has relied on the small molecule drug that inhibits xanthine oxidase, allopurinol, and therefore prevents XO from generating uric acid. However, it does not serve all patients adequately. Approximately, 20 % of patients treated with allopurinol report side-effects, and it is estimated that as many as 5 % of all patients must discontinue taking allopurinol due to the severity of experienced side-effects [25]. Allopurinol must be taken routinely to maintain the necessary drug concentration to elicit a therapeutic effect. Patient non-compliance in following the dosing regimen has been reported to be as high as 44 % within a managed care study cohort [26].

It is estimated that there are more than 10 million gout suffers world-wide, and the number of gout patients is growing at an annual rate of approximately 4 % [27]. Of the five million gout suffers in the U.S alone, more than 100,000 patients are described as having treatment failure gout (TFG) since they do not meet

therapeutic endpoints with conventional treatment [28]. A clinical metric for those patients with TFG is that they are unable to maintain serum urate levels below 6 mg/dL [29]. Despite a long clinical study of gout and the debilitating nature of the disease, current medications for gout management are unable to provide adequate treatment for all suffers What is the driving force behind those patients who are categorized as suffering from TFG? The causes for this lack of therapeutic response range from allergy, intolerance (2 % of patients prescribed allopurinol), inadequate response, and even patient non-compliance [30]. The following quote from the *American Journal of Managed Care* summarizes the medical need for the better management of hyperuricemia and gout:

> *"Hyperuricemia, the predisposing condition for gout, is intricately linked with the metabolic syndrome (hypertension, glucose intolerance, dyslipidemia, truncal obesity, increased risk of cardiovascular disease), and there is mounting evidence that hyperuricemia itself may be an independent risk factor for cardiovascular disease. Unfortunately, gout is frequently mismanaged, resulting in unnecessary morbidity and even mortality [31]."*

1.4 Properties of uricase, its reaction, and evolutionary history

Treatment with uricase breaks down the urate found in gouty joints and disfiguring tophi [25]. For many years, the uricase reaction was described as the directed conversion of xanthine to allantoin, and the biomedical literature continues to use this oversimplification [32-38]. Through the application of stop-flow kinetics experiments, reveal that the true product of uricase catalyzed oxidation of uric acid is 5-hydroxyisoruate (HIU) [39]. 5-HIU has poor stability under physiological conditions, so it is no wonder that until faster techniques were applied, it went undetected [40, 41]. Due to its *in vitro* instability, HIU spontaneously decomposes to 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline

(OHCU) that, in turn, decomposes to racemic allantoin on the time scale of several hours [7]. Uricase performs the oxidation of uric acid to 5-hydroxyisorate with the generation of hydrogen peroxide. One concern that was explored with using uricase to treat gout was the potential to raise plasma hydrogen peroxide concentrations to a lethal level (Figure 1.4).

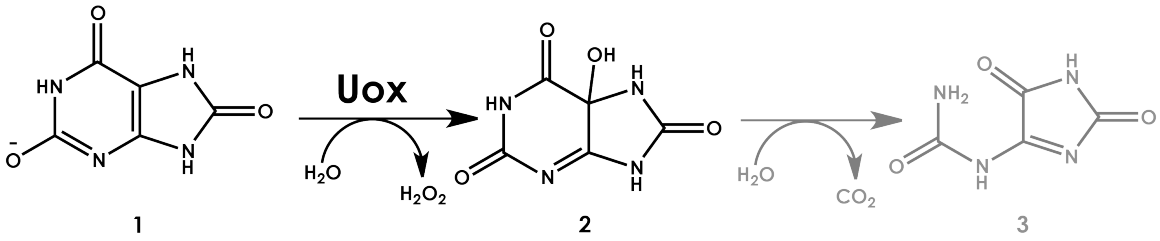


Figure 1.4 Uricase catalyzes a cofactor-less oxidation reaction. The uricase (UOX) reaction converts urate (1) to 5-hydroxisoruate (2) with the release of hydrogen peroxide. 5-HIU is further degraded into allantoin (3) either spontaneously due to its instability under physiological conditions, or enzymatically to the racemically pure S-allantoin enantiomer [42].

Batelli and Stern published one of the earliest systematic studies of uricase in 1909 where they cataloged the tissue localization of uricase in domesticated animals including the horse, dog and rabbit. They found that uricase was ubiquitously detected in the liver, frequently found in the kidneys, and largely absent in other tissues (the lung, spleen, pancreas, brain and muscle) [43]. Specifically within these tissues, uricase is found in microbodies known as peroxisomes. Encapsulated in the peroxisomal single membrane are over fifty metabolic enzymes, and the organelle's name derives from the generation of hydrogen peroxide as a byproduct of the oxidation of specific organic substrates (i.e. urate) by molecular oxygen [44]. Depending upon the host organism, uricase can be found as a soluble tetrameric protein or as paracrystalline array within the peroxisome that is visible by electron microscopy. This diverse solubility is an intrinsic property of the uricases primary amino acid sequence.

Uricase is a unique enzyme that is able to catalyze the conversion of uric acid using gaseous molecular oxygen to 5-hydroxyisoruate in the absence of either a cofactor or a transition metal. The seminal work of Peter Tipton and coworkers with the soybean nodule uricase began parsing out the reaction mechanism employed by this oxidase. Namely that the urate monoanion was first activated via deprotonation via a general base system comprised of two residues conserved in all known uricase enzymes - Thr 57 and Lys 10 (numbering based upon *Aspergillus* Uox) [39]. The exact mechanism by which this enzyme (and other oxygenases like it) are able to bypass the Wigner spin rule which prevents the direct reaction of gaseous triplet oxygen with singlet ground-state organic substrates (in this case urate monoanion) is not well understood. Whereas in other oxygenases, molecular oxygen is excited to the reactive singlet state using a metal/or other organic cofactor. Gabison and coworkers studied the uricase reaction via X-ray diffraction, electron spin resonance spectroscopy (electron paramagnetic resonance spectroscopy), and quantum mechanics calculations to elucidate how uricase is able to overcome this quantum rule [42]. Their work has built on Tipton's work, as well as other structural studies to extend Tipton's proposed catalytic diad to a conserved triad of Thr-Lys-His plus a number of water molecules whose positioning acts as a "push-pull" proton transfer system.

In the literature, only microbial uricase crystal structures have been solved due likely to the poor solubility of mammalian uricase (Figure 1.5) [45]. These crystal structures revealed that uricase is a homotetrameric protein that forms a perfect tunnel of 50 Ångstroms (Å) that channels the substrate to the four active

sites located at the dimer-dimer interfaces. This defining structural feature is known as the tunneling-fold (T-fold) and is formed by an antiparallel beta-sheet comprised of 4 beta strands (β) with 2 antiparallel alpha helices (α) sandwiched between the sets of beta strands (ββαaβ). This fold is responsible for how four identical monomers can come together in a head-to-tail fashion to form a perfect tunnel at which each dimer interface has an active site for 4 active sites per tetramer.

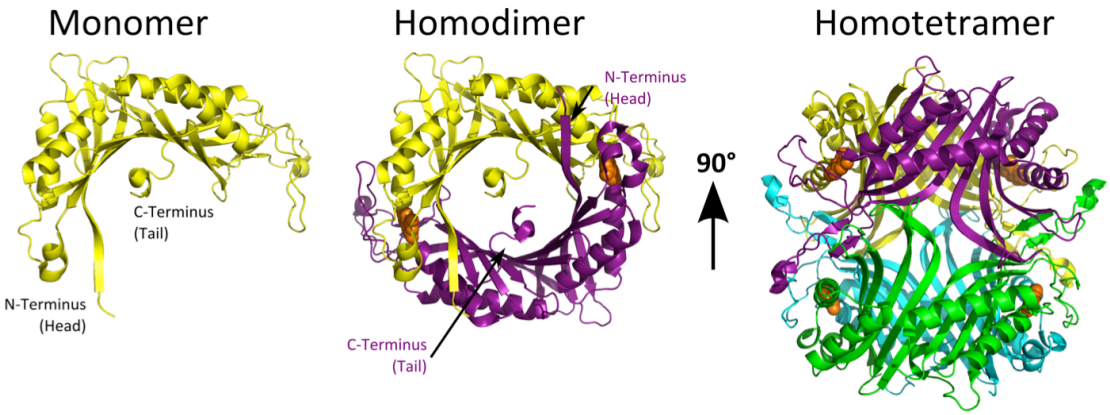


Figure 1.5 Uricase is functional as a homotetramer with active sites at the dimer-dimer interfaces The 3-dimensional representation of the crystal structure of the microbial uricase solved from *Arthrobacter globiformus* (PDB: 2YZB) is shown. At left, an initial monomer is shown in yellow and the multimeric protein is first formed by the coming together of another uricase monomer in a head-to-tail fashion (shown in purple) to form a homodimer (above middle). At far right is the active homotetramer where the homodimer orientation has been rotated 90° along the z-axis and depicts the assembly of a second homodimers (in cyan and green). There are four active sites in the active homotetramer at each dimer-dimer interface where the substrate urate (as orange spheres) binds.

The uricase gene (*Uox*) is a well-conserved ancient gene present in eukaryotic, bacterial and archaeal organisms [46]. While uricase is found to be functional in the liver of most vertebrates, it is lacking functionality in hominoids (i.e. humans). It is immunologically undetectable in some New World monkeys including the macaque and the wooly monkey but is present in most other monkeys [47-49]. The absence of a functional uricase is attributed to genetic

lesions that arose independently as an 18-bp deletion in exon 2 within the branch leading to the Lesser Apes (siamang and gibbon) and as two nonsense mutations in exon 2 and exon 3 in the branch leading to the great apes (chimpanzee, gorilla, and human) some 15 – 20 Ma [50]. These inactivation events and their potential evolutionary roles are discussed further in chapter 2.

1.5 Challenges in therapeutic uricases

Given that humans do not possess a functional uricase enzyme, therapeutic development has been focused on repurposing other functional uricases from nature for use in humans (Table 1.1). The first uricases available to patients was from the filamentous fungus *Aspergillus flavus (A. flavus)* and was either isolated from the microbe (Uricozyme™) or recombinantly expressed (Rasburicase). These fungal uricases are approved as prophylactics to prevent tumor lysis syndrome (TLS) in pediatric patients, but cannot be used to treat gout because they are too immunoreactive in human patients [51]. Fungal uricases from the *A. flavus* and the yeast, *Candida utilis*, have a greater affinity for and catalytic activity on urate at physiological pH compared to vertebrate uricases [52]. Although active in breaking down uric acid, uricases of fungal origin are highly antigenic. In the patient population studied for FDA approval, 15 % to 24 % of patients either were excluded from the study due to a medical history of allergic reactions or developed antibodies to uricase [53]. Repeated administration of the uricase results in allergic reactions, anaphylaxis, and sometimes death [53].

The most recent advancement in therapeutic uricases came with the approval of the first mammalian uricase for the treatment of TFG. This uricase is marketed under the name Krystexxa© and is a pig-baboon chimeric (PBC) uricase. The rationale behind this chimer is that pigs possess one of the most catalytically active uricase characterized in mammals to date; however, its development as a sole therapeutic was abandoned due to high immunogenicity. The researchers combined the amino-terminus portion of the

porcine uricase (residues 1-225) with the carboxyl-terminus of the baboon uricase (residues 226-304) [54]. The result was a chimeric protein with an increased amino acid sequence identity to human uricase (hUox) and an increase in its specific enzyme activity (SEA) of 20 %. PBC uricase is also covalently modified (via PEGylation) to further mask its foreign protein sequence (described in more detail in chapter 5).

Krystexxa®, the PBC uricase, is administered by the intravenous (IV) route due to enhanced bioavailability, efficacy, and tolerability compared to subcutaneous administrations. While an improvement over other uricases explored to date, Krystexxa® has shortcomings, in both its safety and efficacy. In phase III clinical studies, 18 % of patients discontinued treatment in response to serious adverse events. Furthermore, less than 50 % of patients met defined endpoints in lowering blood serum uric acid levels and resolution of gouty tophi [55, 56]. Despite these concerning statistics, the FDA approved Krystexxa® in 2010.

Table 1.1 Several uricase sources have been under clinical development. This table summarizes the major uricases that have been studied in recent years. Prior to recombinant technology, uricase was first isolated from the tissue of pigs or isolated from fungal sources. In both instances, the isolated uricase was not well tolerated and its clinical development was abandoned. The fungal uricase from *Aspergillus flavus* is recombinantly expressed and is only FDA approved for pediatric cancer patients to prevent acute renal failure associated with tumor lysis syndrome. The surface modification of microbial uricases has ben explored to harness their high catalytic activity while masking their intrinsic antigenicity. These modifications are performed by masking the surface with long-chain polymers of PEG. Modification with PEG units of 5 kDa of the yeast *Candida utilis* and the bacteria *Arthrobacter protofomiae* was clinically abandoned, while modification of the *C. utilis* with larger 20 kDa PEGs has progressed to late clinical studies. The first FDA approved mammalian uricase protein pegloticase, marketed under as Krystexxa®, is a chimer of the pig and baboon uricase protein sequences.

Name	Preparation	Sequence Source	Status
Hog liver uricase	Isolated from pig livers	Mammalian, Non-recombinant	Clinical development abandoned
Uricozyme®	Non-recombinant,	*Aspergillus flavus*	No longer manufactured
Elitek®	Recombinant, Non-PEGylated,	Fungal, *Aspergillus flavus*	FDA-approved for tumor lysis syndrome in pediatric patients
Uricase-PEG5	Non-recombinant, PEGylated with MW 5 kDa PEG groups	*Candida utilis* or *Arthrobacter protofomiae*	No longer in clinical development
Uricase-PEG20®	Recombinant, PEGylated	*Candida utilis*	Late clinical studies
Krystexxa®	Recombinant, PEGylated	Mammalian, Pig-Baboon chimer	FDA approved for treatment failure gout

1.6 Scope of this work

The research described within this dissertation is the efforts to date towards both understanding the evolution of the uricase protein family, and applying this knowledge to engineer a safer enzyme for the management of gout. Within this first chapter the stage for the conducted research is set through a discussion of the pathology of gout, the shortcomings of the current standard-of-care and

approved uricase therapies – namely their narrow therapeutic windows due to high antigenicity. In the second chapter, the fate of the human pseudogene is explored first by confirming the presence of uricase transcript in human embryonic tissues and then experimentally confirming that recombinant full-length uricases can be translated in human cells.

Chapter 3 begins the research efforts to develop the a human-like uricase by first attempting to directly reactivate the human uricase, and subsequently applying the evolutionary biology approach of ancestral sequence reconstruction on the uricase protein family. Ancestral proteins are experimentally resurrected and enzymatically characterized to identify active human-like uricases that are suitable for further optimization. The most active uricase is An19/22, and it differs from the human uricase protein sequence by 22 amino acids. The effects of these 22 human residues within An19/22 are explored in Chapter 4. The individual effects are utilized to further humanize the An19/22 while minimizing the cost in both enzymatic activity and stability.

In Chapter 5, the covalent modification of uricases that displayed promising *in vitro* kinetic and stability profiles are covalently modified by activated polyethylene glycols (PEGs). The effect of the employed PEGylation strategy upon An19/22 and PBC (the protein component of the only FDA approved uricase for the treatment of gout) are measured. In addition, unmodified- and PEG- An19/22 and PBC uricase are injected into healthy rats. In these head-to-head pharmacokinetics studies, An19/22 displays an enhanced *in vivo* stability compared to PBC uricase. The final chapter summarizes the research efforts to date to "reengineer a human-like uricase for the treatment of

gout". Specifically, there is a focus upon the work remaining on the path towards human clinical trials.

1.7 References

1. Dobzhansky, T., *Biology, molecular and organismic* Am. Zool., 1964. **4**: p. 443-452.

2. Elias, Y. and R.H. Huang, *Biochemical and structural studies of A-to-I editing by tRNA:A34 deaminases at the wobble position of transfer RNA.* Biochemistry, 2005. **44**(36): p. 12057-12065.

3. Keebaugh, A.C. and J.W. Thomas, *The evolutionary fate of the genes encoding the purine catabolic enzymes in hominoids, birds, and reptiles*, in *Mol. Biol. Evol.* 2010: United States. p. 1359-1369.

4. Hayashi, S., S. Fujiwara, and T. Noguchi, *Evolution of urate-degrading enzymes in animal peroxisomes.* Cell Biochem. Biophys., 2000. **32**: p. 123-129.

5. Iwata, H., S. Nishio, M. Yokoyama, A. Matsumoto, and M. Takeuchi, *Solubility of uric acid and supersaturation of monosodium urate: why is uric acid so highly soluble in urine?* J. Urol., 1989. **142**(4): p. 1095-1098.

6. Gabison, L., M. Chiadmi, N. Colloc'h, B. Castro, M. El Hajji, and T. Prange, *Recapture of S -allantoin, the product of the two-step degradation of uric acid, by urate oxidase.* FEBS Lett., 2006. **580**(8): p. 2087-2091.

7. Tipton, P.A., *Urate to allantoin, specifically (S)-allantoin.* Nat. Chem. Biol., 2006. **2**(3): p. 124-125.

8. Ramazzina, I., C. Folli, A. Secchi, R. Berni, and R. Percudani, *Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes.* Nat. Chem. Biol., 2006. **2**(3): p. 144-148.

9. Nuki, G. and P.A. Simkin, *A concise history of gout and hyperuricemia and their treatment.* Arthritis Res. Ther., 2006. **8**: p. 5.

10. Burns, C.M. and R.L. Wortmann, *Gout therapeutics: new drugs for an old disease.* Lancet, 2011. **377**(9760): p. 165-177.

11. Adams, J.U., *New relief for gout.* Nat. Biotechnol., 2009. **27**(4): p. 309-311.

12. Campion, E.W., R.J. Glynn, and L.O. DeLabry, *Asymptomatic hyperuricemia. Risks and consequences in the normative aging study.* Am. J. Med., 1987. **82**(3): p. 421-426.

13. Teng, G.G., R. Nair, and K.G. Saag, *Pathophysiology, clinical presentation and treatment of gout.* Drugs, 2006. **66**(12): p. 1547-1563.

14. Schumacher, H.R., Jr., *The pathogenesis of gout.* Clevel. Clin. J. Med., 2008. **75 Suppl 5**: p. S2-4.

15. Cardona, F., F.J. Tinahones, E. Collantes, A. Escudero, E. Garcia-Fuentes, and F.J. Soriguer, *The elevated prevalence of apolipoprotein E2 in patients with gout is associated with reduced renal excretion of urates.* Rheumatology (Oxf.), 2003. **42**(3): p. 468-472.

16. de Oliveira, E.P. and R.C. Burini, *High plasma uric acid concentration: causes and consequences.* Diabetol. Metab. Syndr., 2012. **4**: p. 12.

17. Johnson, R.J., L.G. Sanchez-Lozada, and T. Nakagawa, *The effect of fructose on renal biology and disease.* J. Am. Soc. Nephrol., 2010. **21**(12): p. 2036-2039.

18. Tiu, R.V., S.E. Mountantonakis, A.J. Dunbar, and M.J. Schreiber, *Tumor lysis syndrome.* Semin. Thromb. Hemost., 2007. **33**(4): p. 397-407.

19. Davidson, M.B., S. Thakkar, J.K. Hix, N.D. Bhandarkar, A. Wong, and M.J. Schreiber, *Pathophysiology, clinical consequences, and treatment of tumor lysis syndrome.* Am. J. Med., 2004. **116**(8): p. 546-554.

20. Juraschek, S.P., L.C. Kovell, E.R. Miller, 3rd, and A.C. Gelber, *Association of kidney disease with prevalent gout in the United States in 1988-1994 and 2007-2010.* Semin. Arthritis Rheum., 2013.

21. Kottgen, A., E. Albrecht, A. Teumer, V. Vitart, J. Krumsiek, C. Hundertmark, G. Pistis, D. Ruggiero, C.M. O'Seaghdha, T. Haller, et al., *Genome-wide association analyses identify 18 new loci associated with serum urate concentrations.* Nat. Genet., 2013. **45**(2): p. 145-154.

22. Ali, S. and E.V. Lally, *Treatment failure gout.* Med. Health R. I., 2009. **92**(11): p. 369-371.

23. Edwards, N.L., *Treatment-failure gout: a moving target.* Arthritis Rheum., 2008. **58**(9): p. 2587-2590.

24. Bardin, T., *Current management of gout in patients unresponsive or allergic to allopurinol.* Joint Bone Spine, 2004. **71**(6): p. 481-485.

25. Stamp, L.K., J.L. O'Donnell, and P.T. Chapman, *Emerging therapies in the long-term management of hyperuricaemia and gout.* Intern. Med. J., 2007. **37**(4): p. 258-266.

26. Riedel, A.A., M. Nelson, N. Joseph-Ridge, K. Wallace, P. MacDonald, and M. Becker, *Compliance with allopurinol therapy among managed care enrollees with gout: a retrospective analysis of administrative claims.* J. Rheumatol., 2004. **31**(8): p. 1575-1581.

27. Terkeltaub, R., *Gout. Novel therapies for treatment of gout and hyperuricemia.* Arthritis Res. Ther., 2009. **11**(4): p. 236.

28. Fels, E. and J.S. Sundy, *Refractory gout: what is it and what to do about it?* Curr. Opin. Rheumatol., 2008. **20**(2): p. 198-202.

29. Sundy, J.S. and M.S. Hershfield, *Uricase and other novel agents for the management of patients with treatment-failure gout.* Curr. Rheumatol. Rep., 2007. **9**(3): p. 258-264.

30. Perry, M.E. and R. Madhok, *Treatment failure gout: failure to treat?* Rheumatology (Oxf.), 2010. **49**(12): p. 2233-2234.

31. Luk, A.J. and P.A. Simkin, *Epidemiology of hyperuricemia and gout.* Am. J. Manag. Care, 2005. **11**(15 Suppl): p. S435-442; quiz S465-438.

32. Przylecki, S.J., *Uricase and its action.* Biochem. J., 1930. **24**(1): p. 81-81.

33. Keilin, D. and E.F. Hartree, *Uricase, amino acid oxidase, and xanthine oxidase.* Proc. R. Soc. B-Biol. Sci., 1936. **119**(813): p. 114-140.

34. Bentley, R. and A. Neuberger, *The mechanism of the action of uricase.* Biochem. J., 1952. **52**(5): p. 694-699.

35. Hruban, Z. and H. Swift, *Uricase - localization in hepatic microbodies.* Science, 1964. **146**(364): p. 1316-&.

36. Pitts, O.M. and D.G. Priest, *Uricse reaction intermediate - mechanism of borate and hydroxide ion catalysis.* Biochemistry, 1973. **12**(7): p. 1358-1363.

37. Salleh, A.B. and W.M. Ledingham, *Some kinetic-studies on immobilized uricase.* Int. J. Biochem., 1981. **13**(10): p. 1113-1118.

38. Suzuki, H. and D.P.S. Verma, *Soybean nodule-specific uricase (Nodulin-35) is expressed and assembled into a functional tetrmeric holoenzyme in Eschericia-coli.* Plant Physiol., 1991. **95**(2): p. 384-389.

39. Imhoff, R.D., N.P. Power, M.J. Borrok, and P.A. Tipton, *General base catalysis in the urate oxidase reaction: evidence for a novel Thr-Lys catalytic diad.* Biochemistry, 2003. **42**(14): p. 4094-4100.

40. Modric, N., A.E. Derome, S.J.H. Ashcroft, and M. Poje, *Tracing and identification of uricase reaction intermediates - A direct C-13-NMR isotope-labeling evidence*. Tetrahedron Lett., 1992. **33**(44): p. 6691-6694.

41. Kahn, K. and P.A. Tipton, *Kinetic mechanism and cofactor content of soybean root nodule urate oxidase*. Biochemistry, 1997. **36**(16): p. 4731-4738.

42. Gabison, L., C. Chopard, N. Colloc'h, F. Peyrot, B. Castro, M. El Hajji, M. Altarsha, G. Monard, M. Chiadmi, and T. Prange, *X-ray, ESR, and quantum mechanics studies unravel a spin well in the cofactor-less urate oxidase*. Proteins, 2011. **79**(6): p. 1964-1976.

43. Mahler, H.R., G. Hubscher, and R. Baum, *Studies on uricase. I. Preparation, purification, and properties of a cuproprotein.* J. Biol. Chem., 1955. **216**(2): p. 625-641.

44. Voet, D. and J.G. Voet, *Biochemistry*. 3rd ed. 2004, New York: Wiley. xvii, 1223 p.

45. Colloc'h, N., *Crystal Structure of the protein drug urate oxidase-inhibitor complex at 2.05 A Resolution.* Nat. Struct. Biol., 1997. **4**(11): p. 947-952.

46. Oda, M., Y. Satta, O. Takenaka, and N. Takahata, *Loss of urate oxidase activity in hominoids and its evolutionary implications.* Mol. Biol. Evol., 2002. **19**(5): p. 640-653.

47. Varelaechavarria, A., R.M. Deocaluna, and H.A. Barrerasaldana, *Uricase protein sequences - conserved during vertebrate evolution but absent in humans.* Faseb J., 1988. **2**(15): p. 3092-3096.

48. Logan, D.C., D.E. Wilson, C.M. Flowers, P.J. Sparks, and F.H. Tyler, *Uric acid catabolism in the woolly monkey.* Metabolism., 1976. **25**(5): p. 517-522.

49. Usuda, N., M.K. Reddy, T. Hashimoto, M.S. Rao, and J.K. Reddy, *Tissue specificity and species differences in the distribution of urate oxidase in peroxisomes.* Lab. Invest., 1988. **58**(1): p. 100-111.

50. Johnson, R.J., E.A. Gaucher, Y.Y. Sautin, G.N. Henderson, A.J. Angerhofer, and S.A. Benner, *The planetary biology of ascorbate and uric acid and their relationship with the epidemic of obesity and cardiovascular disease*. Med. Hypotheses, 2008. **71**(1): p. 22-31.

51. Mayne, N., S. Keady, and M. Thacker, *Rasburicase in the prevention and treatment of tumour lysis syndrome*. Intensive Crit. Care Nurs., 2008. **24**(1): p. 59-62.

52. Bomalaski, J.S., F.W. Holtsberg, C.M. Ensor, and M.A. Clark, *Uricase formulated with polyethylene glycol (uricase-PEG 20): Biochemical rationale and preclinical studies.* J. Rheumatol., 2002. **29**(9): p. 1942-1949.

53. Goldman, S.C., J.S. Holcenberg, J.Z. Finklestein, R. Hutchinson, S. Kreissman, F.L. Johnson, C. Tou, E. Harvey, E. Morris, and M.S. Cairo, *A randomized comparison between rasburicase and allopurinol in children with lymphoma or leukemia at high risk for tumor lysis.* Blood, 2001. **97**(10): p. 2998-3003.

54. Hershfield, M. and S.J. Kelly. Urate oxidase. 7056713 USPTO application. 2006.

55. Sundy, J.S., H.S. Baraf, R.A. Yood, N.L. Edwards, S.R. Gutierrez-Urena, E.L. Treadwell, J. Vazquez-Mellado, W.B. White, P.E. Lipsky, Z. Horowitz, et al., *Efficacy and tolerability of pegloticase for the treatment of chronic gout in patients refractory to conventional treatment: two randomized controlled trials.* JAMA, 2011. **306**(7): p. 711-720.

56. Lyseng-Williamson, K.A., *Pegloticase: in treatment-refractory chronic gout.* Drugs, 2011. **71**(16): p. 2179-2192.

CHAPTER 2: EXPLORING THE HUMAN URICASE PSEUDOGENE

From the following publications in preparation:

Kratzer, J. T., Lanaspa, M. G., Johnson, R. J. & Gaucher, E. A. Management of intracellular triglyceride levels induced by uric acid using ancient uricase enzymes.

Kratzer, J.T., Murphy, M. N., Ortlund, E. A. & Gaucher, E. A. Evolutionary history of modern and ancient mammalian uricases.

2.1 INTRODUCTION

Pseudogenes are classically defined as DNA sequences derived from functional orthologs that have accumulated deleterious mutations preventing them from functioning within a transcriptional or translational context to yield functional RNAs or proteins [57, 58]. These genes have become non-essential and, in the absence of functional constraints, they freely accumulate disabling mutations including stop codons, repetitive elements, or frame shifts [59]. Until recently, pseudogenes had been referred to as "genomic fossils" and dismissed as "junk DNA." However, current research is presenting a picture of pseudogenes that are anything but "junk." These pseudogenes fill many roles in the cell such as silencing their parent genes and regulating cancer genes [59]. Furthermore, the long held belief that pseudogenes are not transcribed is being questioned through the use of tiling microarrays, which identified 20 % of all known pseudogenes on human chromosome 22 [60].

From an evolutionary biology perspective, studying a pseudogene will shed light upon the forces that led to uricase inactivation. In the development of

a safe and effective uricase therapeutic, an understanding of the mutations responsible for inactivating uricase will be valuable. Most therapeutic proteins either are human proteins or humanized proteins. In humans, the gene encoding uricase is a pseudogene and is non-functional [61]. This chapter details our inquiry into understanding what happens to the uricase pseudogene at the transcriptional and translational levels. Embarking on this path of inquiry will aid us in developing a human-like uricase for the treatment of gout.

2.1.1 The properties of the human uricase pseudogene

The inactivation of uricase in the hominoid lineage is attributed to 4 genetic lesions (Figure 2.1). The earliest lesion is unique to the gibbon uricase and is the result of a 13-bp frame-shift deletion in codon 72 [61]. The remaining 3 mutations are all present in the human uricase pseudogene (*hUox*). The first is a nonsense mutation at codon 33 within exon 2 which is estimated to have occurred approximately 13 million years ago (Mya), and is shared between humans and the great apes [46]. The remaining two inactivating events occurred some 10 Mya: another nonsense mutation at codon 187 in exon 5 and a splice site mutation in intron 2 (located between exon 2 and exon 3) [61, 62]. When *hUox* sequence is aligned with functional mammalian uricases the positions encoding the two nonsense mutations are conserved arginines (with the CGA codon) in orthologous sequences. Both of the premature stop codons are TGA, and it thus seems likely that this TGA → CGA mutation is the result of spontaneous deamination of cytosine to yield thiamine [63].

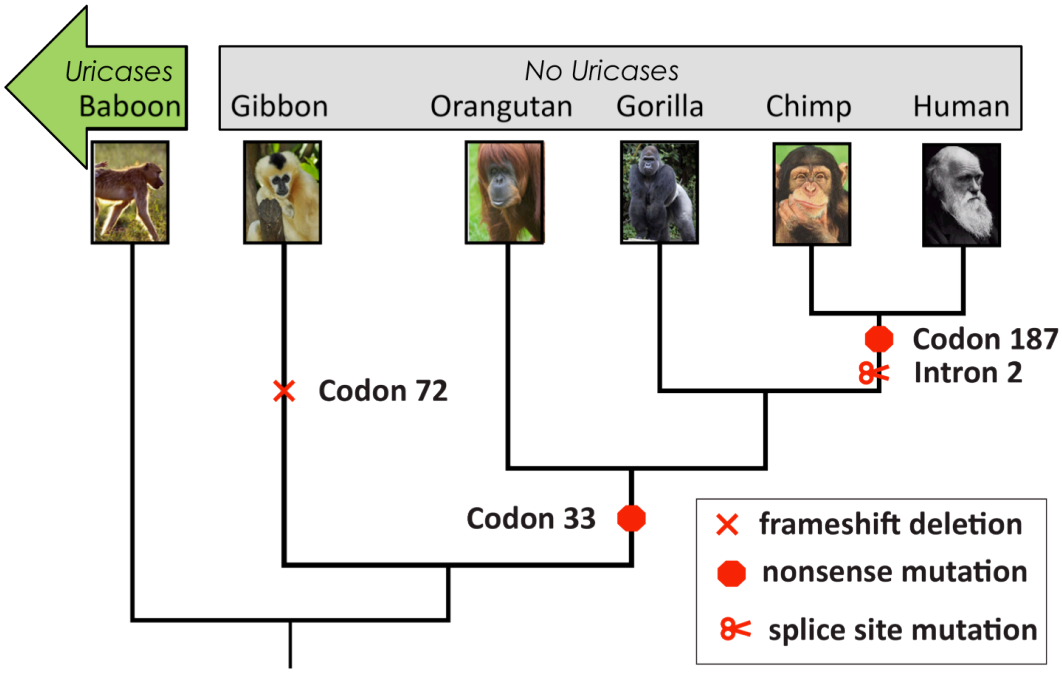


Figure 2.1 The pseudogenization of human uricase is the result of three genetic lesions. Shown is a phylogenetic tree that represents the evolutionary relationship among the uricase gene in hominoids. The baboon uricase is catalytically active and is boxed in green. In contrast, the inactive uricases of the hominoid lineage (gibbon, orangutan, gorilla, chimp, and human) are boxed in gray. The 13 base pair frame-shift deletion is unique to the gibbon sequence, and is independent of the human uricase pseudogenization. The earliest inactivating mutation found is in orangutans and is a nonsense mutation at codon 33. An additional nonsense mutation is also found in the chimp sequence at codon 187. The final deactivating mutation is a splice site mutation also found in the chimp sequence located in intron 2 (between exons 2 and exon 3).

What survival advantage was conferred by uricase's inactivation in our ancestors during the Miocene between 5 – 23 Mya? Some hypotheses that are put forth approach this question from the perspective of "planetary biology." This term captures the interdisciplinary threads of evidence from the fields of paleontology, geology, chemistry and molecular biology that come together to weave hypotheses about early life [64]. Since time-travel is not possible, these multifaceted inquiries provide us with a plausible window into the past.

2.1.2 Potential advantages of uricase inactivation

Uricase is a conserved enzyme present in many microbes and most fish, amphibians, and mammalian species, but is notably absent among humans. The absence of a functional uricase in humans predisposes our species to possessing uric acid levels that are between 10 to 50 times greater than those found in other mammals – which may have several adaptive advantages – with the consequence of gout arising in relatively modern time [29, 65]. A few of these hypotheses are now briefly highlighted.

The end products of purine catabolism reflect the primary nitrogenous waste excreted in animals: uric acid (e.g. humans), urea (e.g. fishes), and ammonia (e.g. crustaceans) [66]. One hypothesis is rooted in the water economy of the host organism. As introduced in Chapter 1, the first exit point of purine catabolism is urate in terrestrial animals where water is scarce. [67]. In organisms such as fish where water is abundant, the highly toxic ammonia can be rapidly removed. However, when water is less abundant, nitrogen is stored as a less toxic breakdown intermediate.

Another hypothesized advantage to elevated uric acid is its ability to protect the body from injury. For example, urate is an extracellular antioxidant that is able to react with a range of oxidizing agents including hydrogen peroxide, hydroxyl radical, peroxynitrite, and nitric oxide [68-70]. Its ability to scavenge free radicals may protect humans from cancer and other life-shortening disorders [71]. Furthermore, urate has been implicated at playing a role in regulating blood pressure in animal models - inhibiting uricase in rats results

in a decrease in endothelial nitric oxide that, in turn, stimulates the renin angiotensin system and leads to a rise in blood pressure. [72].

There is also a potentially advantageous energy role for uricase inactivation, and commensurate urate elevation. Dr. Rick Johnson's group has recently published evidence explaining the previously suspected tie between hyperuricemia and metabolic disease [73]. Their work in human liver cells identified the loop by which uric acid and fructose are metabolically intertwined: uric acid activates the transcription factor ChREBP, which in turn stimulates fructokinase (KHK), the first enzyme in the metabolism of fructose. KHK phosphorylates fructose leading to transient ATP depletion, and the adenine nucleotide which is shuttled through the purine metabolic pathway to uric acid [74].

2.1.3 The fate of the uricase pseudogene is therapeutically important

The most dramatic translational consequence of the nonsense mutations in *hUox* is that only a fragment of the uricase protein is produced due to translational stalling at the first stop codon within exon 2. However, it has been established that the sequence-context by which stop codons are positioned dictates their strength, and that stop codon read-through is a fairly common occurrence [75].

Determining the fate of the *hUox* will direct our efforts to make a safer uricase protein. If indeed the non-functioning *hUox* gene product is ever naturally translated *in vivo* it could still be recognized by the immune system as self. A human uricase protein may therefore not elicit the severe, potentially lethal allergic reaction that occurs with other uricases.

2.2 MATERIALS AND METHODS

2.2.1 Transcriptional Studies[1]

To search for transcripts of the human uricase pseudogene, fetal liver cDNA (ClonTech) was used as the template for a series of polymerase chain reactions (PCR) using standard PCR reagents (Promega). Custom forward (F1-F8) and reverse (R1-R2) primers were designed against the published sequence of the human uricase gene and synthesized by Integrated DNA Technologies (IDT) [76]. For the sequences of these *hUox* querying primers see Sequence S1 in the supplemental information. All PCR reactions were performed in a total volume of 50 μL consisting of: 1 μL of template cDNA (1 ng), 2 μL 10 μM each of forward and reverse primers, 0.25 μL GoTaq™ polymerase enzyme (1.25 units), 10 μL 5X GoTaq™ buffer, 1 μL deoxyribonuclotide mix (10 mM), bring to volume with dH_2O. Products were examined via 1.5 – 2 % (w/v) agarose gel electrophoresis, at 100 V for 30 - 45 minutes. Bands were imaged using EpiChem Darkroom system (UVP). The bands of produced products were then excised using a sterile razor blade and purified using a gel extraction kit (Qiagen).

2.2.2 Translational Studies

2.2.2.1 Design of hUox expression constructs

Four different *Uox* constructs were synthesized to explore the translation of the human uricase sequence in human embryonic kidney cells (HEK). These constructs were designed to determine the extent to which the number and strength of premature stop codons affect uricase translation in mammalian host

[1] The transcriptional studies were performed by Christina Graves (a Georgia Institute of Technology undergraduate researcher)

cells. Each uricase gene construct contains a hexahistidine tag (6xHis), upstream of the start codon to assist in the purification using immobilized metal affinity chromatography (IMAC). Synthesis of *hUox* with 2 internal stop codons, and *hUox* with 3 stop codons, including the 2 internal stop codons naturally present in the human uricase pseudogene, along with a strong TGA-A stop codon at the carboxyl terminus.

2.2.2.2 The first two hUox expression constructs: two or three stop-codons

The gene corresponding to full-length uricase, according to Wu *et al.* [61], was synthesized with the two nonsense mutations (TGA) at codons 33 and 87. A 6xHis tag was added to the N-terminus using multiple rounds of PCR, and the required cloning sites for the p3XFLAG-CMV expression vector (Sigma) at the 5' *Eco*RI site and at the 3' *Bam*HI were engineered using the following primers:

hU-F1: CATCACCACCATCACATGGCCCACTACCATAAC

hU-F2: GCGAATTCATGGCACACCATCACCACCATCAC

hU-R1: CAGGGATCCCAGTCTTGAAGACAACTTC

hU-R1_TerminalSTOP: CAGGGATCCTCACAGTCT**TGAA**GACAACTTC

The forward primers hU-F1 and hU-F2 added the 6xHis tag (underlined) and the EcoRI restriction site (blue) respectively. Whereas, the reverse primers, hUR1 and hU-R1_TerminalStop, were used in PCR reactions to introduce the BamHI (red) restriction site and strong C-terminal stop (shown in bold), respectively. Using these restriction sites, the *hUox* construct was sub-cloned into P3XFLAG-CMV. The hUox containing the two internal stop-codons (found in the pseudogene) is referred to as the "2-STOP" construct, and the *hUox* construct with the strong terminal stop is called "3-STOP."

2.2.2.3 The third hUox expression construct: Removing internal stop codons

A series of site-directed mutagenesis (SDM) reactions were performed using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies). The following sets of primers were used to mutate the two stop codons to arginines (the codons to be mutated by these primers are underlined, and the point mutations are shown in bold):

Converting codon 33:

33toRfor: 5'- GTTCTCCATATTCAG**C**GAGATGGAAAATATCACAGC -3'

33toRrev: 5'- GCTGTGATATTT TCCATCTC**G**CTGAATATGGAGAAC -3'

Converting codon 187:

187toRfor: 5'- GGTGAAGGAC**C**GATGCTTTGCCACCCAAG -3'

187toRrev: 5'- CTTGGGTGGCAAAGCATC**G**GTCCTTCACC -3'

Both reactions were performed in parallel, and colonies that grew after transformation were sequenced. Successful conversion of two nonsense mutations was confirmed by sequencing. The *hUox* construct with in which both of the internal stops found in the pseudogene are converted to arginines is called the "NO-STOP" construct.

2.2.2.4 The fourth hUox expression construct: Removing exon 3 from hUox

This construct models the scenario where the splice site mutation results in the cellular machinery skipping over exon 3 entirely by using a downstream splice site acceptor as a construct. This gene includes the uricase construct where the two stop codons have been mutated to the arginines, with the entire sequence encoding exon 3 is missing was synthesized (Epoch). The following set of primers were used to generate the "-EXON 3" *hUox* construct:

Eco-F1: CATCACCACCATCACGCTCATTATCACAATAATTACAAG

hU-F2: GCGAATTCATGGCACACCATCACCACCATCAC

Eco-R1: CAGGGATCCCAGGCGGCTGCTC

To this gene, the following set of primers were used to introduce an N-terminal 6xHis tag (underlined) and the *Eco*RI (blue) and *Bam*HI (red) restriction sites for sub-cloning into the p3XFLAG mammalian expression system.

2.2.2.5 HEK-293T culture maintenance and transfection

A cell stock was obtained from Invitrogen and used to seed cultures. They were passaged a minimum of 5 times to maintain a healthy cell density. A large scale preparation of DNA for transfections was prepared by first transforming *hUox* construct-containing p3XFLAG CMV14 vector (CMV14) from Sigma-Aldrich vector into the Nova Blue (NB) *E. coli* K-12 cloning strain (EMD Millipore). Cells were plated on LB agar plates with carbenicillin (CMP) and grown overnight at 37 °C. A single colony was used to seed a 100 mL culture, which after a high cell-density was obtained was spun down and plasmid DNA was obtained using a Maxi Prep Kit (Qiagen). In addition to CMV14, the p3XFLAG CMV13 (Sigma-Aldrich) that adds a LLS leader sequence, for secretion of recombinantly expressed protein into the cellular media, was also tested.

A suspension cell line of Human Embryonic Kidney cells (HEK-293T) was transiently transfected with hUox-containing expression vectors using the FreeStyle 293 Expression system (Invitrogen). The mammalian hUox expression workflow began with determining the number of genes to be transformed and passaging cells a minimum of 5 times to recover from being rescued from a stock, thus ensuring both a proper cell viability and density for transfection.

Based upon the number of variants to transfect and desired yield, the amount of DNA and the number of cells can be determined from the FreeStyle kit's manual (Invitrogen). The appropriate volume of cells was grown so that each reaction would be performed at a cell density of 1 x 10^6 cells/mL. Cell density was initially determined with a V*i*-Cell™ Cell Viability Analyzer (Beckman Coulter) until the doubling times of laboratory stocks was well established. Generally for the data presented, 150 µg of DNA was transfected into 75 mL of cells at the appropriate cell density. To perform a transfection, the calculated volume of cell culture was spun down at 100 x *g* and at 4 °C. The supernatant was carefully decanted and the cell pellet was resuspended in 1 mL of fresh room temperature (RT) Freestyle media (Invitrogen). In a separate tube, a predetermined amount of plasmid DNA was brought up to 5 mL with the Opti-MEM reduced serum media (Invitrogen), and was mixed gently. In another 15 mL conical tube, the appropriate concentration of 293fectin, a cationic-lipid formulation to carry the DNA to be transfected (Invitrogen), was brought up to volume also with Opti-MEM. The contents of the two tubes were mixed together and incubated for 20-30 min after which they were added to fresh media in clean 250 mL erlenmeyer flasks and allowed to incubate for 3 days in a shaking incubator at 150 rpm and 5 % CO_2.

2.2.2.6 Expression and purification of hUox constructs in human cells

Seventy-two hours after transfection, cells were collected by centrifugation and protein was purified either under native conditions when the secreting CMV13 vector was used, or under denaturing conditions when the CMV14 expression vector was used. Cell pellets were either lysed using the native

extraction buffer, or denaturing cell extraction buffer (Invitrogen). Regardless of the buffer used, a protease inhibitor cocktail (Sigma-Aldrich) was added to the extraction buffer as a 1:10 dilution along with PMSF at a final concentration of 1 mM. The cell and lysis buffer mixture was clarified by centrifugation and the supernatant was saved for purification.

The clarified supernatants were then applied to a pre-packed 1 mL HisTrap™ HP column (GE Healthcare) on a AKTA™ UPC 10 fast protein liquid chromatography (FPLC) system. The buffer system used was 20 mM sodium phosphate buffer with 0.5 M NaCl, pH 7.4 with 20 mM imidazole in the binding buffer, and 500 mM imidazole in the elution buffer. A 20-column volume (CV) gradient from 20 mM to 500 mM imidazole was used, and the purified samples were eluted at approximately 250 mM imidazole. The same buffer system was utilized for the denaturing purification with the addition of 8 M urea to both the binding and the elution buffers.

2.2.2.7 Protein detection by western blotting

Following purification, samples were loaded onto two separate sodium dodecyl sulfate polyacrylamide (SDS-PAGE) gels (12 % resolving and 4 % stacking) and run at 150 V for 1 hour at room temperature on a Mini-Protean gel apparatus (Bio-Rad). The purity of the preparation was assessed by Commassie Brilliant Blue staining. Protein samples were transferred to a polyvinylidine fluoride (PVDF) membrane for western blotting using the semi-wet transfer method. A Trans Blot western transfer apparatus (Bio-Rad) was used with Bjerrum and Shafer-Nielsen transfer buffer (5.8 g Tris-Base, 2.9 g Glycine, 3.75 mL 10 % SDS, 200 mL MeOH, and brought to 1 L with dH_2O). Transfers were run at 15 V for 1 hour and

transfer was monitored by following the migration of the Kaleidoscope™ (Bio-Rad) pre-stained protein standards from the acrylamide gel to the membrane.

The Western blot buffers used in membrane blocking, rinsing, and antibody incubation are all phosphate buffered saline (PBS) solutions comprised of: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4 with 0.08 % Tween-20™ added (PBST). In the membrane blocking and antibody incubation steps, 5 % (w/v) powered milk is added to make PBSTM. Each incubation step was performed with gentle agitation on a rocking platform to ensure complete coverage of the membrane. Following transfer a five-minute blocking step, to prevent nonspecific antibody binding, was performed by incubating the membrane at room temperature in PBST. After the blocking step, the membrane was washed 3 times with equal volumes of fresh PBST for 5 min each. For protein detection, one of the following antibodies were used at 0.5 mg/mL in PGS: an anti-HIS mouse monoclonal antibody, an anti-FLAG (DYKDDDDK) mouse monoclonal antibody origin (GenScript), or an anti-uricase (Santa Cruz Biotechnology). Primary antibody incubation was performed overnight at 4 °C at a concentration of 0.5 mg/mL in PBSTM. Following primary incubation the membrane was then washed 3-4 times with PBST for 5 min each. A secondary antibody incubation was performed at room temperature for 1 hour with a Ab-HRP conjugate to enhance the chemilluminescent signal. After washing the membrane a final time, it was then incubated for 5 min with super Pico western luminol solution (Pierce) and then imaged by either film or the a ChemiDoc™ XRS+ Imaging System (Bio-Rad). The parallel gel was stained by coomassie and was used to confirm the size of the illuminated bands.

2.3 RESULTS AND DISCUSSION

2.3.1 Evidence of *hUox* transcript in human tissue

The cell's transcription and translation machinery can deal with the human uricases three inactivating mutations in several possible ways. In the first scenario (Figure 2.2, I), the splicesome bypasses the aberrant splice site acceptor and uses the proximal one to generate a mature mRNA missing exon 3 of 798 bp in length. A database query returned sequencing information of three partial human uricase transcripts (GenBank GI#'s 70957646, 76555529, 76555824).

In the second scenario (Figure 2.2, II), the splicesome completely skips over intron 2 and exon 3, which results in an mRNA transcript of 926 bp, longer than the full-length gene. In scenario III was found in a database query that returned a predicted mRNA (GI# 157412274), which is indicative of both mRNA promoter function and a theoretical translational product [61, 76]. Transcriptional recognition of the aberrant splice site acceptor has been illustrated, providing justification that the transcriptional machinery in a human cell may recognize an alternative "aa" acceptor site. Additionally, ribosomal read through of the stop codon "TGA-G" and "TGA-T" has also been documented as being read-through by the translation machinery some 10 % of the time [75]. Such read-through could result in a full-length properly spliced human uricase of 915 bp (Figure 2.2, IV).

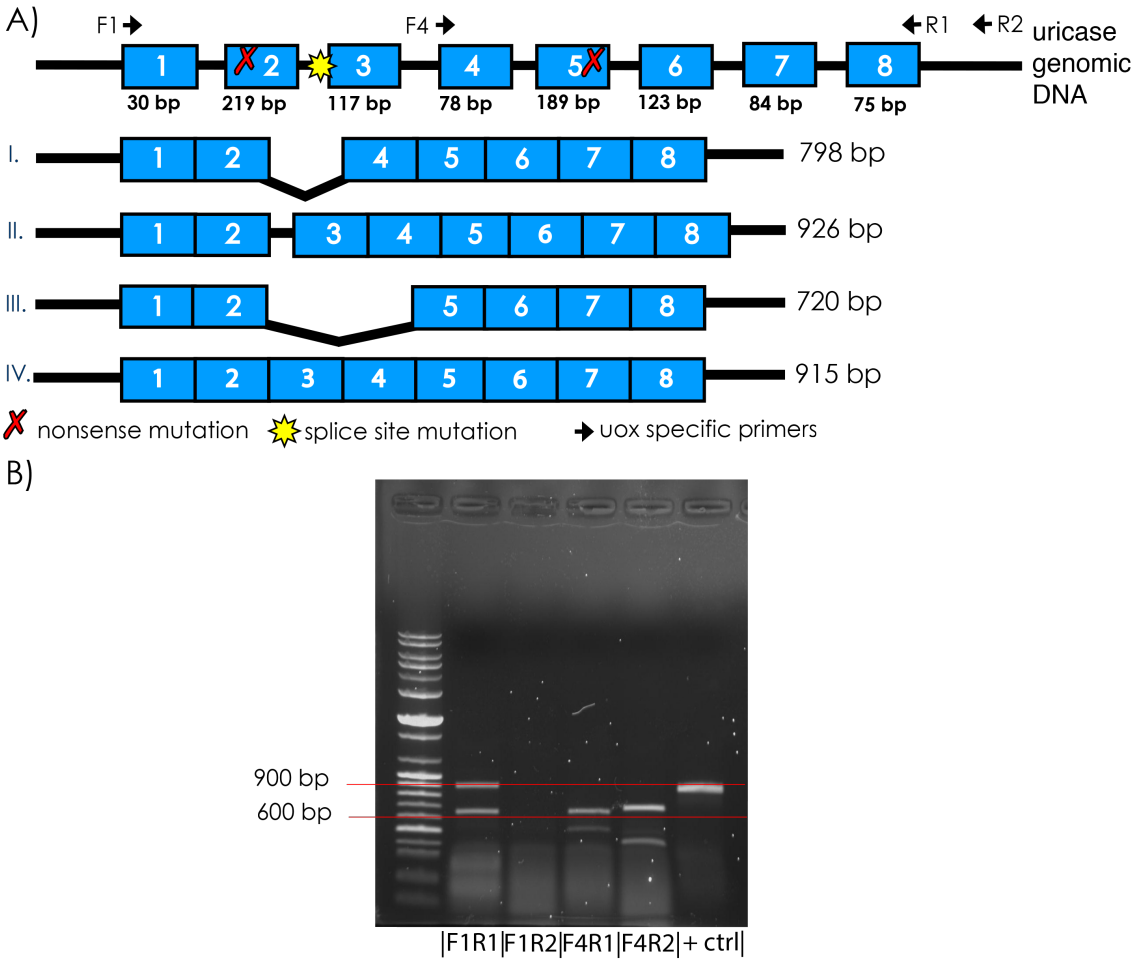


Figure 2.2 Uricase transcripts are present in human tissue (Panel A) At top is the schematic of the unspliced uricase genomic DNA. In this cartoon, each exon is numbered and depicted as a blue box with the corresponding size shown in base pairs (bp) below, and the intervening intronic sequences are shown as a solid black line. The two nonsense mutations are represented by red "Xs" and are located within exon 2 and exon 3, and the splice site mutation is yellow starburst located in intron 2 before exon 3. Four sequence specific primers were designed to query the state of mRNA transcript of hUox in human tissue. Forward primers F1 and F2 bind at the beginning of exons 1 and 4, respectively. Reverse primers R1 and R2 bind to the end of exon 8 and downstream of exon 8, respectively. The four possible mature transcripts that can be generated from hUox are: I. The downstream proximal splice site acceptor located between exon 3 and exon 4 is utilized and a mRNA transcript missing exon 3 of 798 bp is generated. II. The upstream proximal splice site acceptor between exon 2 and exon 3 is used which gives a mRNA transcript of 926 bp (containing the intronic sequence upstream of the splice site mutation along with all of exon 3). III. The downstream splice site acceptor located between exon 4 and exon 5 is used resulting in a mRNA transcript of 720 bp that is missing both exon 3 and exon 4. IV. The uricase genomic DNA is properly spliced giving a full length mRNA transcript of 915 bp. (Panel B) The experimental detection of hUox transcript is shown by the specific

(Figure 2.2 continued)
PCR reactions using fetal cDNA as a template and site specific pairs of PCR primers F1, F4, R1, and R2 run on a 1.5 % agarose DNA gel. A 10,000 bp DNA ladder (lane 1) and actin with a size of 838 bp (lane 6) act as reference points for identification of PCR products obtained with hUox transcript probes. A product corresponding to size of full-length uricase of 915 bp is obtained with primers F1 and R1, along with a smaller product (lane 2). When the reverse primer R2 that binds to the intronic sequence downstream of the last exon 8 is used, no product is obtained showing that the hUox pseudogene is processed (lane 3). The product obtained with the F4 and R1 primer set corresponds to the size of an mRNA transcript minus exon 3 and exon 4 (lane 4) and when R2 is used as the reverse primer an even smaller product is obtained (lane 5).

We queried human cDNA libraries for uricase transcripts using PCR in attempt to better understand the distribution of mRNA transcripts in fetal and adult human cells. This experiment demonstrated that *hUox* of various sizes (including full-length and minus exon 3) is present early in development, from cDNA prepared from human fetal liver tissue. Future PCR experiments, using forward primers specific to only exon 3 may be helpful to further understand the population of *hUox* transcripts present in human tissue – allowing for the amplification and sequencing of uricase transcripts containing all functional transcripts except those that do not contain exon 3. The aberrant splice-acceptor site in intron 2 is ignored by the splicesome in favor of the acceptor site in intron 3. This splicing event generates a transcript (795 bp) that is in-frame, contains the two premature stop codons, but is missing exon 3.

2.3.2 Human cells can express full-length hUox

Considering that human uricase is non-functional, would a human-like uricase offer an improved safety profile over other foreign uricases? It follows that if the human body makes a non-enzymatically active but full-length uricase protein by stop-codon read-through, then the major histocompatibility complex II (MHC-II) and T-cells would recognize a more human-like uricase protein

sequence as 'self-like'. Thus, it is possible that while a human-like uricase will have a lower enzymatic activity compared to other "foreign uricases," it may also elicit a lower immune response making it a safer therapeutic. To explore this hypothesis, we needed to ascertain whether humans ever transcribe and/or translate their endogenous uricase gene. If so, is it ever a full-length protein or is it a partial protein?

To answer this question, we synthesized human uricase genes containing specific combinations of premature and terminal stop codons to determine the read-through ability for the different stop codons. The genes were cloned into the p3XFLAG CMV vector (Sigma) that constitutively expresses the genes with a His-tag at the amino-terminus and a FLAG3x-tag at the carboxyl-terminus when a stop codon is not encountered or read through. The constructs were then transiently transfected into human 293T cells. Cells were grown and harvested to remove recombinant uricases. The uricases were present in the insoluble fraction only, and were purified via denaturing nickel affinity chromatography. Western Blots were performed using an antibody towards the FLAG epitope since this domain is only present upon read-through of the stop codons. The ribosomes in 293T cells do in fact read-through the premature stop-codons with substantial frequency whereas the terminal stop codon is very efficient at terminating translation (Figure 2.3).

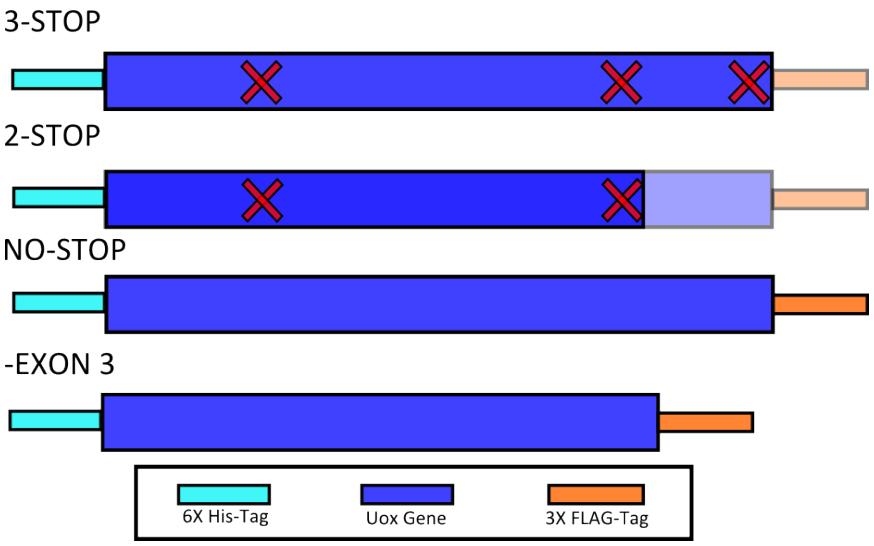


Figure 2.3 Constructs synthesized to test effects of stop codons on hUox expression in humans. The synthesized Uox is shown in blue and there is a 6X His-Tag at the N-terminus for affinity purification, and a 3X FLAG Tag at the C-terminus for full-length protein detection. In the 3-STOP variant the first two nonsense mutations in the human pseudogene are encoded followed by a third strong stop codon at the end of the gene before the C-terminus tag. The 2-STOP variant encodes the human uricase pseudogene with its two nonsense mutations. In the NO-STOP construct, the two nonsense mutations have been substituted with arginine residues that are found in functional homologous uricase sequences. Lastly, the –EXON 3 variant contains the full length human gene in which the bases encoding exon 3 have been removed to model the cellular machinery using a proximal splice site acceptor and bypassing exon 3 in its entirety.

To ascertain whether uricase was expressed in the HEK-293T cells a polyclonal antibody raised against the full murine uricase was tested (Figure 2.4). It would be very advantageous to have a highly selective antibody for going after hUox; however, this polyclonal antibody did not display the requisite selectivity. There was a lot of noise from the chemilluminescent signal, as evidenced by the number of bands in the protein standard lane (Lane 1). In addition, despite employing very stringent wash conditions, the polyclonal antibody used against the highly purified uricase protein sample in Lane 6, gave a strong high molecular weight, non-specific/non-uricase band. This lack of

selectivity and poor sensitivity makes this pAb unsuitable for detecting small amounts of hUox. Therefore, a more selective antibody must be employed along with loading as much hUox as possible.

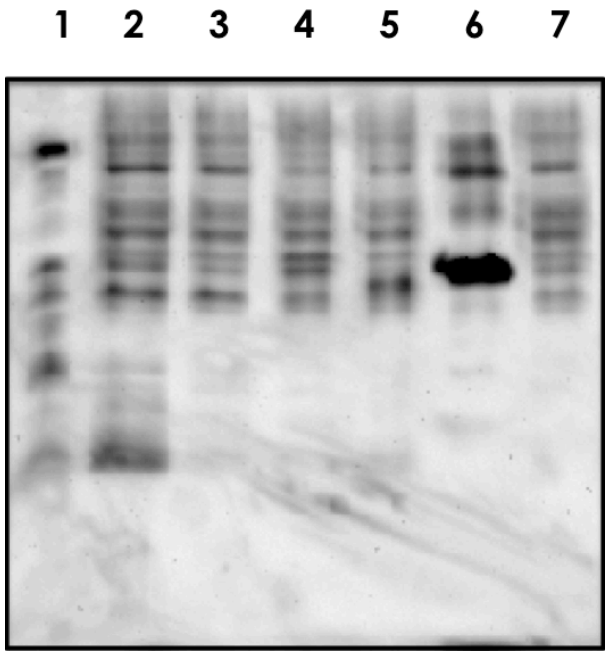


Figure 2.4 The rabbit polyclonal antibody against mouse uricase displayed poor selectivity for hUox detection. The samples that were run on this western blot were obtained from denaturing purification on a HisTrap HP His-binding column. The primary polyclonal antibody (pAb) was raised in rabbits against the full-length mouse uricase sequence. A pAb-HRP conjugate was used for signal amplification. There was a high background despite a high wash stringency. This background can be observed by the nonspecific bands in the 10-250 kDa protein standard, which is free from uricase (Lane 1); Lane 2) 3 –STOP uricase construct; Lane 3) 2-STOP uricase construct; Lane 4) No-STOP uricase construct; Lane 5) hUox-Exon 3 construct; Lane 6) mammalian uricase control (with 93 % sequence identity to hUox); Lane 7) Empty vector.

Since the Uox pAb gave results that were poorly resolved, we next turned to the Anti-FLAG epitope, which is only present when read-through of all stop-codons occurs (Figure 2.5). This signal worked well to determine when full-length product was translated. When a strong terminal stop is employed (Lane 1), no FLAG signal is present which is indicative of no read-through. In contrast, when

the 2 endogenous premature stop codons in hUox are present, the full-length product is generated (Lane 3). As controls, the full-length hUox and the empty vector behaved as expected, lanes 5 and 6, respectively. Lastly, the hUox-exon 3 also generates a product (Lane 6) that runs slightly smaller than the full-length uricase (Lane 5).

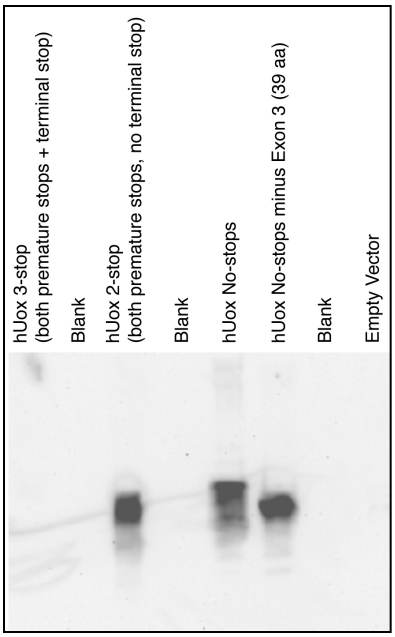


Figure 2.5 Western blot detection of full-length human uricase expression in 293T cells. Samples were run on this gel that were obtained from denaturing purification on a HisTrap HP His-binding column. The primary antibody used is of mouse origin against the FLAG tag that is only present if the human cells produce full-length product. The secondary antibody used for signal amplification is of goat origin and is conjugated to HRP. Note: The samples were all concentrated; therefore, the intensity of the signal between samples cannot be confidently compared. Lane 1) No full-length signal is produced in the 3-STOP construct. Lane 3) Despite the two nonsense mutations found in the human pseudogene, full-length uricase is detected in the hUox 2-STOP transfected sample. Lane 5) A gene product corresponding to the absence of exon 3 is generated by the –EXON 3 construct. Lane 8) The vector sample confirmed that the specific FLAG epitope is required for signal generation by the conjugated HRP, and therefore full-length products have been isolated.

2.4 CONCLUSIONS

To understand the minimum requirements for termination of transcription and translation of human uricase pseudogene we have queried human tissue sources for *hUox* transcripts, and experimentally recreated expression of each of these transcripts in human (HEK-293T) cells. The effects of each pseudogene feature have been explored in this chapter: 1) the aberrant splice site acceptor in intron 2; 2) the nonsense mutations (stop codons) at aa33 and aa187; and 3) the extent to which these lesions might be overcome to produce full-length human uricase protein. As previously mentioned, although transcription products of human uricase have been described, translational products have not been characterized. It is unclear whether the described mRNA transcriptional product is comprised of complete exonic sequences, or if the aberrant splice site acceptor in intron 2 disrupts the transcriptional product and therefore excludes parts of exon 3. This mutational acceptor signal in intron 2 may lead to recognition of a cryptic splice acceptor site downstream in exon 3, therefore excising part of exon 3 in the final mRNA product. Ultimately, this difference would affect the size of the transcriptional product currently characterized, as well as downstream translational abilities.

In some regards, our results are not all that surprising since it is known that sequence signatures flanking known terminal stop codons influence the robustness of the termination signal [77, 78]. Nevertheless, the results suggest that the premature stop codons in uricase have not evolved to serve as strong termination signals. In total, our results suggest that it is likely that humans express non-functional uricase because transcription of the gene takes place, the

uricase transcript has an intact ribosome binding site, and the ribosome can read through the premature stop codons when over expressed in human cell culture. It is not clear at this point how these transcripts evade nonsense-mediated decay. In total, these experiments suggest that human cells can express the *hUox* in spite of the aberrant splice site acceptor and two non-sense mutations. Assuming that the immune system identifies this non-functional uricase as self there its presence is a compelling reason for the development of a "human-like" uricase that may be safer than uricases from other sources by avoiding a dangerous antigenic response.

2.5 REFERENCES

29. Sundy, J.S. and M.S. Hershfield, *Uricase and other novel agents for the management of patients with treatment-failure gout.* Curr. Rheumatol. Rep., 2007. **9**(3): p. 258-264.

46. Oda, M., Y. Satta, O. Takenaka, and N. Takahata, *Loss of urate oxidase activity in hominoids and its evolutionary implications.* Mol. Biol. Evol., 2002. **19**(5): p. 640-653.

57. Li, W.-H., *Molecular evolution*. 1997, Sunderland, Mass.: Sinauer Associates. xv, 487 p.

58. Vanin, E.F., *Processed pseudogenes: characteristics and evolution.* Annu. Rev. Genet., 1985. **19**: p. 253-272.

59. Tutar, Y., *Pseudogenes.* Comp. Funct. Genomics, 2012.

60. Rouchka, E.C. and I.E. Cha, *Current Trends in Pseudogene Detection and Characterization.* Curr. Bioinf., 2009. **4**(2): p. 112-119.

61. Wu, X.W., C.C. Lee, D.M. Muzny, and C.T. Caskey, *Urate oxidase: primary structure and evolutionary implications.* Proc. Natl. Acad. Sci. U. S. A., 1989. **86**(23): p. 9412-9416.

62. Yeldandi, A.V., V. Yeldandi, S. Kumar, C.V. Murthy, X.D. Wang, K. Alvares, M.S. Rao, and J.K. Reddy, *Molecular evolution of the urate oxidase-encoding gene in hominoid primates: nonsense mutations.* Gene, 1991. **109**(2): p. 281-284.

63. Lai, H.M., Y.Y. Chiang, C.C. Hsu, and F. Wu, *A recognition machine for CpG-islands based on Boltzmann model.* J Med Biol Eng, 2008. **28**(1): p. 23-30.

64. Benner, S.A., S.O. Sassi, and E.A. Gaucher, *Molecular Paleoscience: Systems Biology from the Past*, in *Advances in Enzymology*, E.J. Toone, Editor. 2010, John Wiley & Sons, Inc. p. 1-132.

65. Bomalaski, J.S. and M.A. Clark, *Serum uric acid-lowering therapies: where are we heading in management of hyperuricemia and the potential role of uricase.* Curr. Rheumatol. Rep., 2004. **6**(3): p. 240-247.

66. Vogels, G.D. and C. Van der Drift, *Degradation of purines and pyrimidines by microorganisms.* Bacteriol. Rev., 1976. **40**(2): p. 403-468.

67. Needham, J., *Chemical embryology*. 1963, New York,: Hafner Pub. Co.

68. Gersch, C., S.P. Palii, W. Imaram, K.M. Kim, S.A. Karumanchi, A. Angerhofer, R.J. Johnson, and G.N. Henderson, *Reactions of peroxynitrite with uric acid: formation of reactive intermediates, alkylated products and triuret, and in vivo production of triuret under conditions of oxidative stress.* Nucleosides, 2009. **28**(2): p. 118-149.

69. Gersch, C., S.P. Palii, K.M. Kim, A. Angerhofer, R.J. Johnson, and G.N. Henderson, *Inactivation of nitric oxide by uric acid.* Nucleosides, 2008. **27**(8): p. 967-978.

70. Waring, W.S., A. Convery, V. Mishra, A. Shenkin, D.J. Webb, and S.R. Maxwell, *Uric acid reduces exercise-induced oxidative stress in healthy adults.* Clin. Sci., 2003. **105**(4): p. 425-430.

71. Ames, B.N., R. Cathcart, E. Schwiers, and P. Hochstein, *Uric acid provides an antioxidant defense in humans against oxidant- and radical-caused aging and cancer: a hypothesis.* Proc. Natl. Acad. Sci. U. S. A., 1981. **78**(11): p. 6858-6862.

72. Watanabe, S., D.H. Kang, L.L. Feng, T. Nakagawa, J. Kanellis, H. Lan, M. Mazzali, and R.J. Johnson, *Uric acid, hominoid evolution, and the pathogenesis of salt-sensitivity.* Hypertension, 2002. **40**(3): p. 355-360.

73. Lanaspa, M.A., L.G. Sanchez-Lozada, Y.J. Choi, C. Cicerchi, M. Kanbay, C.A. Roncal-Jimenez, T. Ishimoto, N. Li, G. Marek, M. Duranay, et al., *Uric acid induces hepatic steatosis by generation of mitochondrial oxidative stress: potential role in fructose-dependent and -independent fatty liver.* J. Biol. Chem., 2012. **287**(48): p. 40732-40744.

74. Lanaspa, M., L. Sanchez-Lozada, C. Cicerchi, N. Li, C. Roncal-Jimenez, T. Ishimoto, M. Le, G. Garcia, J. Thomas, C. Rivard, et al., *Uric acid stimulates fructokinase and accelerates fructose metabolism in the development of fatty liver.* PLoS One, 2012. **7**(10).

75. Jungreis, I., M.F. Lin, R. Spokony, C.S. Chan, N. Negre, A. Victorsen, K.P. White, and M. Kellis, *Evidence of abundant stop codon readthrough in Drosophila and other metazoa.* Genome Res., 2011. **21**(12): p. 2096-2113.

76. Wu, X.W., D.M. Muzny, C.C. Lee, and C.T. Caskey, *Two independent mutational events in the loss of urate oxidase during hominoid evolution.* J. Mol. Evol., 1992. **34**(1): p. 78-84.

77. Sogaard, T.M., C.G. Jakobsen, and J. Justesen, *A sensitive assay of translational fidelity (readthrough and termination) in eukaryotic cells.* Biochemistry (Mosc), 1999. **64**(12): p. 1408-1417.

78. Manuvakhova, M., K. Keeling, and D.M. Bedwell, *Aminoglycoside antibiotics mediate context-dependent suppression of termination codons in a mammalian translation system.* RNA, 2000. **6**(7): p. 1044-1055.

CHAPTER 3: INDENTIFICATION OF AN ACTIVE HUMAN-LIKE URICASE

From the following publications in preparation:

Kratzer, J. T., Lanaspa, M. G., Johnson, R. J. & Gaucher, E. A. (2013) Management of intracellular triglyceride levels induced by uric acid using ancient uricase enzymes.

Kratzer, J.T., Murphy, M. N., Ortlund, E. A. & Gaucher, E. A. (2013) Evolutionary history of modern and ancient mammalian uricases.

3.1 INTRODUCTION

The transcriptional and translational features reported in chapter 2 supports the presence of pseudogene transcripts (*hUox*) transcripts, and that the pseudogene gene product (hUox) is expressed in human cells. Even though the pseudogene is catalytically inactive, its presence can be therapeutically exploited. We believe that a functional, human-like uricase may be able to evade the inactivating immune response that plagues other clinically explored uricases. It is well established that humans lack an intrinsically active uricase; therefore, those mutations necessary to restore uricase activity to the human pseudogene must be identified. Towards this end, chapter 3 reports our efforts to directly reactivate the hUox, and identify other mammalian uricases that can be explored for further clinical development.

There are 17 amino acid differences between the last active hominoid uricase (baboon) and the inactive human protein. Since none of these mutations are in the active site it is a formidable task to select which mutations will optimize the sequence identity towards the human protein while achieving

maximal enzymatic activity. The highly conserved nature of the uricase protein family affords the opportunity to apply the evolutionary biology rooted approach of ancestral sequence reconstruction (ASR) to explore those sequence combinations that nature has already vetted to obtain active proteins before their inactivation in the hominoid lineage.

3.1.1 Ancestral Sequence Reconstruction (ASR) overview[2]

The ASR approach was first proposed by Pauling and Zuckerkandl in 1963 as a means of testing hypotheses about early life by inferring ancient sequence using modern sequences as input into models of evolution, and then resurrecting (synthesizing) these ancestral proteins in the laboratory [80]. The first experimental ASR study was conducted by Benner and coworkers in 1990 with their resurrection of a catalytically active 5 million year old bovid ribonuclease [81]. This first study validated the potential of ASR to generate functional ancestral proteins for modern-day study. Since its advent, ASR has allowed researchers to "travel back in time," by resurrecting ancient macromolecules in the laboratory to study diverse phenomena ranging from the evolution of ultraviolet vision [82], the origin of steroid hormone signaling [83], and the paleoenvironment of the ancient Earth [84], among others [85].

When put in practice, ASR is a hybrid computational-experimental approach that is comprised of two distinct components (Figure 3.1). The computational component encompasses the analysis of extant homologous protein sequences and putting these sequences within a phylogenetic

[2] Adapted from Kratzer, J.T., Cole, M.F., and E.A. Gaucher, *Protein engineering guided by natural diversity*, in *Protein Engineering Handbook*, S. Lutz and U.T. Bornscheuer, Editors. 2013, Wiley-VCH: Weinheim, Germany.

framework using explicit models of protein evolution. Once a phylogeny is established, statistical methods are then used to infer the most probabilistic ancestral sequences at nodes within the phylogenetic tree. During the experimental component of ASR, the inferred ancestral sequences are synthesized, characterized, and studied to address the posed evolutionary hypotheses [86].

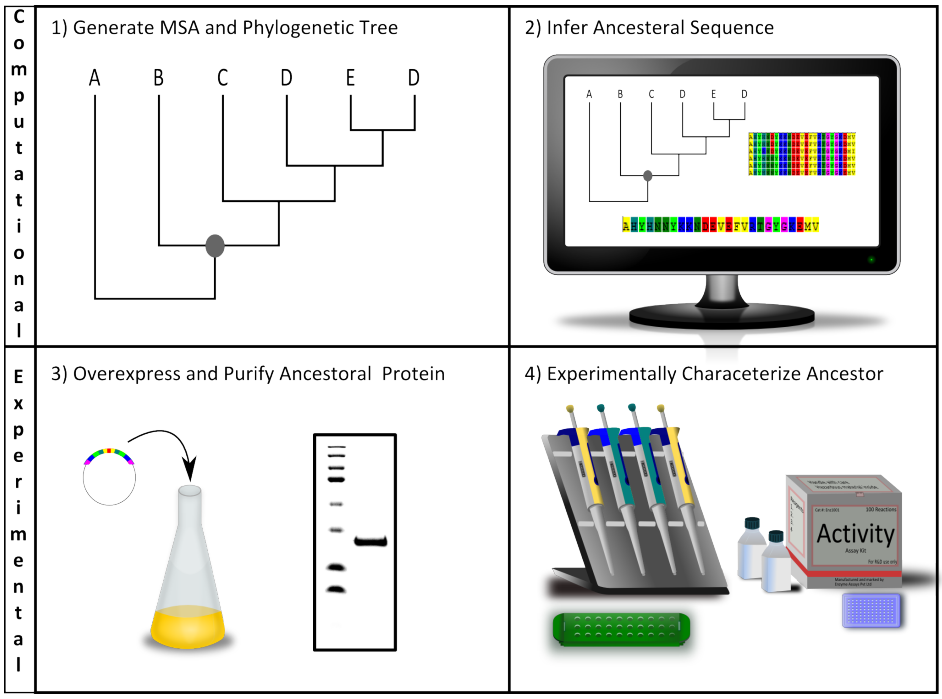


Figure 3.1 Ancestral sequence reconstruction allows researchers to travel back in time and explore ancient proteins. This approach first involves a computational element which involves 1) Generation of a multiple sequence alignment (MSA) to put the proteins of interest into the appropriate phylogenetic context. 2) Evolutionary models are used to infer the ancestral sequences with high statistical support. The second, experimental, component of ASR involves 3) The overexpression and purification of ancestral proteins. Using standard chemical synthesis techniques the inferred ancestral genes are synthesized and put in an expression vector. The proteins are expressed in the system of choice (e.g. *E. coli*) and then purified using standard molecular biology techniques. 4) The proteins are then assayed using a specialized functional test to answer posed questions about the proteins evolutionary past, or to search sequence space for variants with desired biomolecular properties for further development.

The first step in ASR is to collect homologous sequences of the parent protein from sequence databases such as NCBI, PFAM, EBI, GenBank, RefSeq, TPA, SwissProt, PIR, PRF, and PDB. A multiple sequence alignment (MSA) is then created using software such as ClustalW [87] or T-Coffee [88]. This alignment is refined, as needed, to obtain a trustworthy alignment. A high quality MSA is generated by including homologs from a large number of different species, along with an out-group sequence. It is important to choose an out-group sequence that is only moderately related to the query sequences; since its longer genetic distance translates into a longer branch in the inferred phylogenies - aiding in rooting the tree and trimming down the tree-search space which can be computationally demanding [89].

The MSA is then used as the input for phylogenetic analysis to determine the relationships and evolutionary distances among homologous sequences. An evolutionary algorithm is used to construct a phylogenetic tree from the MSA. There are several algorithms for constructing phylogenetic trees such as maximum parsimony[90], maximum likelihood [91], distance-based approaches such as neighbor joining, or Bayesian approaches. While the underlying statistical methods used in each of these tree-building approaches vary, the underlying goal is the same: to search the 'tree-space,' to generate a tree that best recapitulates the evolutionary relationships and histories of the sequences. One popular software tool that applies Bayes Theorem to the tree-building search is MrBayes [92]. A generated gene/protein phylogenetic tree can then be checked against alternative hypotheses of evolutionary relationships (e.g. the species tree derived from 16S ribosomal sequences) to check for consistency. In

order to get the most robust MSA, and accompanying phylogenetic tree, it is often required to repeat cycles of adding/removing sequences from the alignment to correct inconsistencies (disagreements with known evolutionary theory) and ambiguities (resolving polytomies) and then rebuilding the phylogenetic tree until a consistent tree structure is resolved.

The final computational step in ancestral sequence reconstruction is to infer the ancestral sequence. These tools often include various models of molecular sequence evolution that the user must choose to analyze the data with in order to best fit the data to a model (Figure 3.2).

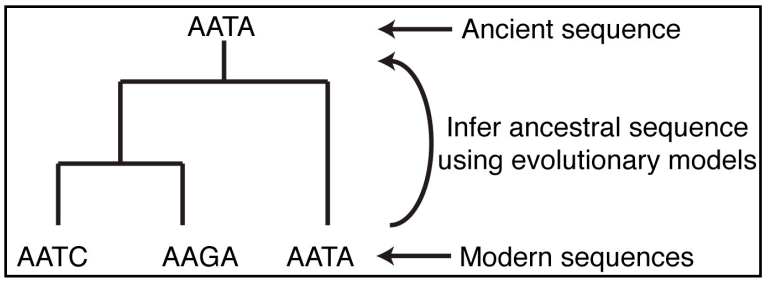


Figure 3.2 Phylogenetic tree highlighting distribution and relationship of modern and ancient sequences. This simple phylogenetic tree shows 3 modern sequences in which there are two bifurcation events, places where the tree has split. The first gave rise to modern sequence AATA and the second gave rise to AATC and AAGA. Using evolutionary models and specific computer algorithms the ancient sequence is inferred to be AATA because it is the most parsimonious answer – since it explains the character states of the modern sequence with the minimum number of changes in the topology.

Once the ancestral sequences have been inferred, they can then be synthesized by traditional DNA synthesis techniques, cloned into an overexpression system, expressed in a modern organism, purified, and then characterized in the laboratory. It is important to note that unique challenges in expressing and purifying protein from ancestral nodes compared to modern-day proteins may exist.

3.1.2 Successful examples of ASR application

The application of ASR in the field of applied protein engineering studies is relatively new, but it offers the advantage of searching sequence space for changes that have already been vetted by natural selection. Scientists may find the greatest utility of ASR when designing more theromostable and/or acid-tolerant proteins because it is hypothesized that ancient life flourished under a much hotter and more acidic environment [93, 94]. Furthermore, ancestral sequences may serve as better 'parent' sequences for directed evolution libraries due to their increased stability. This increased stability allows these proteins to be more receptive to accepting more destabilizing mutations, and this, in turn, may allow the protein to adopt novel biomolecular properties [95, 96].

In the most recent example, a diverse team of researchers set out to answer fundamental questions about how enzymes evolve over time, and the influence that environment had upon their stability and catalytic evolution. Towards this goal, the team chose to study the protein thioredoxin (Trx), which is an oxido-reductase enzyme present within all domains of life and reduces disulfide bonds in proteins [97]. The researchers used ASR to resurrect seven ancient Trxs - including those enzymes from the last bacterial common ancestor, last archael common ancestor, and the archael-eukaryotic common ancestor (which are hypothesized to have last inhabited Earth some 4.2 - 3.5 billion years ago) [98]. The thermostabilities of these ancestors were determined by differential scanning calorimetry and the catalytic ability of these ancestors to reduce disulfide bonds was studied by atom-force microscopy. This is of

particular significance because Trxs have a wide range of industrial applications where enhanced acid- and thermo- stability would be valued. Some of the industrial uses of Trxs include grain processing, allergen deactivation, antibiotic production and venom deactivation [99].

Ancestral Sequence Reconstruction (ASR) follows a present-day-backwards strategy, whereby genomic sequences from extant (modern) organisms are incorporated into evolutionary models to computationally infer the extinct (ancient) sequences of genes no longer present on Earth [84, 98]. To date, approximately 20 studies have emerged where specific molecular systems from extinct organisms have been resurrected for study in the laboratory [64].

3.2 MATERIALS AND METHODS

3.2.1 Direct reactivation of the human uricase pseudogene

The first logical step in reactivating the human pseudogene was to replace the two premature stop-codons with the conserved arginines found in the other functional mammalian uricases. The full-length human uricase where codons 33 and 187 code for arginines in lieu of stops was codon optimized for expression in *E. coli.* This construct was synthesized and cloned into the pET21A+ expression vector. To confirm that human uricase could not encode a highly functional gene, we engineered a human uricase for recombinant expression in which the two stop codons were replaced with arginine (the amino acid found at these two positions in homologs), and the introns were removed from the gene. We confirmed bacterial expression of our engineered human protein by western blot against a 6xHis tag. Unfortunately, the protein resides in the insoluble

fraction and we have not been able to solubilize it under numerous buffer conditions. Regardless, activity assays performed with the buffer suspended insoluble fraction from uricase overexpression showed that no functional human uricase was present within the insoluble fraction.

3.2.2 Computational inference of ancient uricases

Ancestral sequence reconstruction was performed following our laboratory protocols [84, 98]. Modern uricase sequences from 17 taxa were retrieved from public databases and were aligned using ClustalX [87]. The evolutionary relationship between the uricase genes (i.e. DNA sequences) was modeled using MrBayes [92]. The following biological assumptions were used to define the model of character change used: the generalized time reversible DNA substitution model [100], and a proportion of the nucleotide sites are invariable while the remaining sites are drawn from a gamma distribution [101]. To search tree-space, two independent Metropolis-coupled Markov chain Monte Carlo (MCMCMC) simulations, four chains each were performed for 1,000,000 generations with parameter sampling every 100 generations. The first 100 samples were discarded during the burn-in phase of the MCMCMC analysis.

Next, the ancestral sequences were inferred on the codon level using the previously prepared DNA MSA and the phylogenetic topology built by MrBayes. The software application used was Phylogenetic Analysis by Maximum Likelihood (PAML) [102]. The Jukes and Cantor model of nucleotide substitution was used [103]. In addition, the substitution rates were allowed to vary between lineages with omega being the only free parameter. The following parameters were

based upon our laboratory's previous experience and expertise: an alpha parameter of 0.8, an ncatG of 8, and kappa was fixed at 2.789.

3.2.3 Uricase expression and purification

Uricase encoding genes were codon-optimized for expression in *E. coli* and were synthesized by standard methods (Epoch Laboratories). These genes were cloned into the pET-21a vector (Novagen) using the N-terminus restriction site *Nde*I and *Xho*I site at the C-terminus. *E. coli* Tuner (DE3) (EMD Millipore) cells were freshly transformed with uricase containing vector. A single colony was used to inoculate a 5 mL overnight culture. This overnight culture was used to seed a 1 L of Luria broth (LB) with 100 μg/mL carbenicillin (CARB) and 100 μg/mL chloramphenicol (CMP). Cells were grown to an OD_{600} between 0.6 and 0.8 at which point they were induced with 1 mM IPTG. Expression was carried out overnight (16-20 hours) at 37 ° C with shaking of 250 rpm. Cells were then collected by centrifugation at 5,000 x g for 30 min at 6 °C and stored at -80 °C in a 500 mL centrifuge bottle.

The frozen cell pellet was removed from the freezer and allowed to thaw for 15 minutes at room temperature. Two alternative protocols were employed for the lysis of these pellets depending upon the scale of expression. The first protocol for very large preparations (used by our collaborators for crystallization screening) began by resuspending the pellet in 20 mL of Phosphate Buffered Saline (PBS), pH 7.4. The cells were then ruptured by sonication on ice. Sonication was performed with a Branson Sonifier 150 (Emerson) and was carried out in 4 cycles of 40-second bursts at maximum intensity with 1-minute rests in between. After sonication, insoluble material was isolated by centrifugation at

10,400 x g for 10 min at 6 °C. PBS-insoluble contaminants were removed by a series of five PBS washes. For each wash the inclusion body (IB), was resuspended in 20 mL of PBS and vortexed for 1 min at maximum speed, centrifuged at 10,400 x g for 10 min at 6 °C, and the supernatant was discarded. The removal of PBS soluble contaminants was confirmed by SDS-PAGE analysis.

For smaller scale purifications, a non-ionic detergent lysis procedure was employed. Cell pellets were again removed from the freezer and allowed to thaw. These cells were lysed by the addition of BugBuster™ with Benzonase™ (BB) lysis buffer with 1 mM PMSF protease inhibitor added (10 mL of BB per each 1 L worth of cell pellets. This cellular suspension was rocked at room temperature for 30 min. After incubation, the lysate was transferred to a 30 mL microcentrifuge tube and was spun down at 16,000 x g for 15 min at 6 °C. The now clarified extract was saved for SDS-PAGE analysis. The insoluble IB was then washed by first adding a 55% (v/v) BB made in advance, then vortexing for 1 min and centrifuging again at 16,000 x g, the supernatant was saved for analysis. The next wash 3 wash steps were performed in a similar manner but using 10 % (v/v) BB solutions in place of the 55% BB solution.

Regardless of which lysis procedure was employed, the now-clean IBs are resuspended in a 0.1 M to 1 M sodium carbonate buffer at a pH between 10 and 11 and 1 mM PMSF. The resuspended inclusion bodies were rocked at 4 °C overnight to liberate trapped functional uricase from the IB. Carbonate-insoluble debris were then removed by centrifugation at 20,190 x g for 30 min at 6 °C. The carbonate extraction supernatant contained active uricase as confirmed by the Amplex Red Uricase/Uric Acid Activity Kit (Invitrogen).

Uricases were purified on the AKTADESIGN UPC 10 with inline UV and conductivity detectors (GE Healthcare) at 4 °C. Two different purification chromatographic schemes were employed. The first program used was modeled after the two-part purification scheme reported for the Pig Baboon Chimeric uricase [54]. In brief, this method involved an initial extraction step into 1 M Na_2CO_3 that was followed by a complex capture/refining step via anion exchange chromatography (AEX). This AEX step was performed on a HiLoad 16/10 Q-Sepharose HP column (GE Healthcare) pre-equilibrated with 0.075 M Na_2CO_3, pH 10.2. The uricase-containing carbonate extract was filtered through a 0.2 μM filter and diluted to 267 mL with 1 mM PMSF to lower the sample buffer concentration to 0.075 M Na_2CO_3. The sample was applied directly to the column with a sample pump at a flow rate of 1.6 mL/min. After sample application, unbound sample was washed with binding buffer until absorbance at 280 nm approached baseline levels. Next, the pH of the column was lowered by a wash step with 1 column volume of $NaHCO_3$, pH 8.5. An initial wash step with 1 column volume of $NaHCO_3$, pH 8.5 with 0.150 M NaCl was performed. Major contaminants were washed from the column as two large peaks by a 2-column volume salt gradient from 0.15 M NaCl – 1.5 M NaCl in 10 mM $NaHCO_3$, pH 8.5. A 3-column volume wash with 10 mM $NaHCO_3$, pH 8.5, 1.5 M NaCl ensured that all major contaminants had been eluted from the column media. This was followed by a 3-column volume wash with 10 mM $NaHCO_3$, pH 8.5. In preparation for the elution of purified uricase, the pH was raised with 1 column volume of Na_2CO_3, pH 11. Uricase was then eluted by a 10-column volume salt gradient from 0-0.6 M NaCl in 0.1 M Na_2CO_3, pH 11. Uricase was collected in 2 mL

fractions using connected Frac-950 (GE Life Sciences). The eluted uricase peak was pooled and concentrated to less than 5 mL by centrifugation with 20 mL, 9 kDa molecular weight cut-off spin concentrators (Pierce).

The final purification step was performed by size exclusion chromatography on a HiLoad 16/60 Superdex 200 prep grade column pre-equilibrated with 0.1 M Na_2CO_3, pH 10.2 (GE Life Sciences). Sample was applied using a 10 mL superloop at a flow rate of 1 mL/min. This polishing step separated active vs. inactive oligomeric states of uricase. Tetrameric uricase was eluted with 0.1 M Na_2CO_3, pH 10.2 at approximately 140 kDa, and was monitored via an in-line ultraviolet (UV) detector at 280 nanometers (A280). This peak was collected in 2 mL fractions and was pooled and concentrated by centrifugation with spin concentrators (Pierce).

3.2.4 Measuring uricase enzymatic activity

The enzymatic activity of purified tetrameric uricase was determined spectrophotometrically by monitoring the decrease of absorbance at 293 nm (A293), the absorption maxima of uric acid. Reactions were performed in 1 mL reaction volumes in a quartz cuvette at room temperature. A freshly prepared 1 mM uric acid stock in 0.1 M sodium phosphate buffer, pH 7.4, was diluted with 0.1 M sodium phosphate buffer, pH 7.4 to prepare a range of uric acid concentrations. Specifically, the assays were performed at the following uric acid concentrations: 1 μM, 2.5 μM, 5 μM, 10 μM, 15 μM, 20 μM, 25 μM, 50 μM, 75 μM, and 100 μM. The amount of enzyme added to the reaction was empirically determined to give a linear decrease in urate (A293) over the 6 min time course of the assay. Plotting the decrease in A293 versus time (min) and determining the

slope in linear portion of the curve was used to determine the initial velocities of these reactions (MS Excel). The averages from triplicate runs of initial velocities at each urate concentration were used to plot a hyperbolic regression curve to determine the Michaelis constant (K_M) and the maximum velocity (V_{max}) of the purified uricase variant (Hyper32 Kinetics Application). The concentration of purified tetrameric uricase was determined by the Quick Bradford Assay (Bio-Rad). Lastly, the k_{cat} was determined by dividing the V_{max} by the concentration of tetrameric uricase used in the kinetics experiments.

3.2.5 Uricase storage stability assays

Equal masses of cell pellets were lysed and the inclusion bodies were washed with BB detergent. Each sample was separated out into a separate extraction tube to which one of six different carbonate buffers were utilized to extract uricase and subsequently used for its purification by SEC. These carbonate buffers were prepared at concentrations of 0.1 M, 0.5 M, or 1 M and adjusted to either pH 10.2 or pH 11. Once purified, the total protein concentration was determined by Bradford assay using IgG protein mass standards. The specific enzyme activity (SEA) was determined at a urate concentration of 100 μM in triplicate. Uricase preparations were stored at 4 °C in their respective purification buffers and were kept on ice during the SEA determination.

3.2.6 Making uricase variants based upon the inferred ancestral sequences

We explored a range of expression conditions including: growth temperatures ranging from 4 °C to 37 °C, IPTG induction concentrations, auto induction, as well as the CD41 and CD43 overexpression strains for toxic proteins

(Lucigen). All of the conditions explored did yield recombinantly expressed soluble uricase. We therefore employed a protocol to exploit the purity of the insoluble protein in the inclusion body and after several washes with dilutions of BB; the functional uricase was extracted with a pH 11 carbonate buffer.

3.2.7 Uricase structural modeling and solving mammalian uricase crystal structure[3].

A homology model of one of our mammalian uricases was generated using SWISS Model [104]. Briefly, the amino acid sequence of An19/22 (our oldest soluble uricase see Figure 3.6) was thread onto chain A of the X-ray crystal structure of uricase from *Arthrobacter globiformus* (PDB: 2YZB). This procedure was repeated for the remaining four chains of the tetrameric structure. The four generated output PDBs were all loaded into the same PYMOL session from which structural inferences could be made [105].

For crystallization screening experiments, a large 6 L preparation of An19/22 and its variant with 2 surface lysine mutations An19/22_LysII (described later) were expressed and purified following standard experimental procedures (See 3.2.3). This preparation's purity was confirmed by SDS-PAGE analysis and its functional activity was confirmed by *in vitro* assay. The crystal screens were performed using a Phoenix drop setter (Rigaku) and Formulator screen maker (Formulatrix). The X-ray structure was solved for both the inhibitor 8-azaxanthine bound to uricase, and the apo structure. Initial X-ray diffraction data was obtained on a local source at Emory University. High-resolution X-ray diffraction

[3] Crystallization experiments were performed by Emory University collaborators: Dr. Eric Ortlund and Dr. Michael Murphy.

data was obtained from the Advanced Photon Source (Argonne National Laboratory).

3.3 RESULTS AND DISCUSSION

3.3.1 Development of uricase controls

One natural and one engineered uricase were selected to serve as standards for method development and benchmarking of protein activity. As a control for working with mammalian uricases, the pig uricase was overexpressed, purified, and characterized. The pig uricase was chosen because it has 88 % sequence identity to the hUox, but is dissimilar enough to preclude its use as a therapeutic. In addition, pig uricase has long been studied in the context of uricase research, and it is quite active relative to other mammalian uricases [43, 106, 107]. The engineered pig-baboon chimeric (PBC) uricase, is the active part of the FDA-approved Krystexxa®, and serves as a therapeutic benchmark to compare human-like uricases. In addition, PBC uricase has a higher sequence identity to hUox than pig uricase (89 % vs. 87 % identity).

By incubating the insoluble pellet from recombinant overexpression in carbonate buffer pH 11 overnight, uricase could be solubilized for downstream purification and subsequent characterization. The purification scheme resulted in obtaining high purity tetrameric uricase by size exclusion chromatography (SEC) (Figure 3.3).

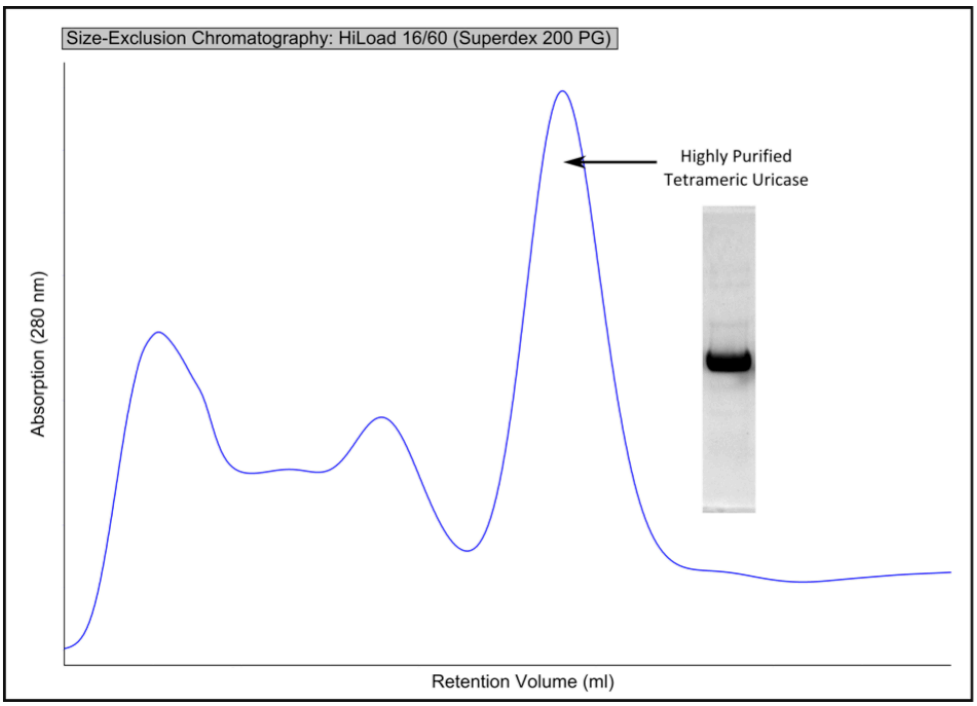


Figure 3.3 Tetrameric uricase is obtained by SEC. The insolubility of the uricase protein was exploited for purification since it was highly enriched for target protein. This size-exclusion chromatography trace on a HiLoad 16/60 (Superdex 200 PG) shows large aggregates that are removed during the purification process on the left the above trace. The large sharp peak contains functional tetrameric uricase (based upon column calibration) whose purity is evidenced in the adjacent SDS-PAGE.

The kinetics assays provide a quantitative means of comparing the extant, ancestral, and engineered uricases. These enzymatic assays were performed in triplicate until a suitable curve was generated. While the free Hyper32 application was used for the majority of the analyses – the same results were obtained when the more feature-rich Origins (OriginLab) was used. To deal with outliers replicates were performed when experimenter error (e.g. pipetting) was likely to have occurred.

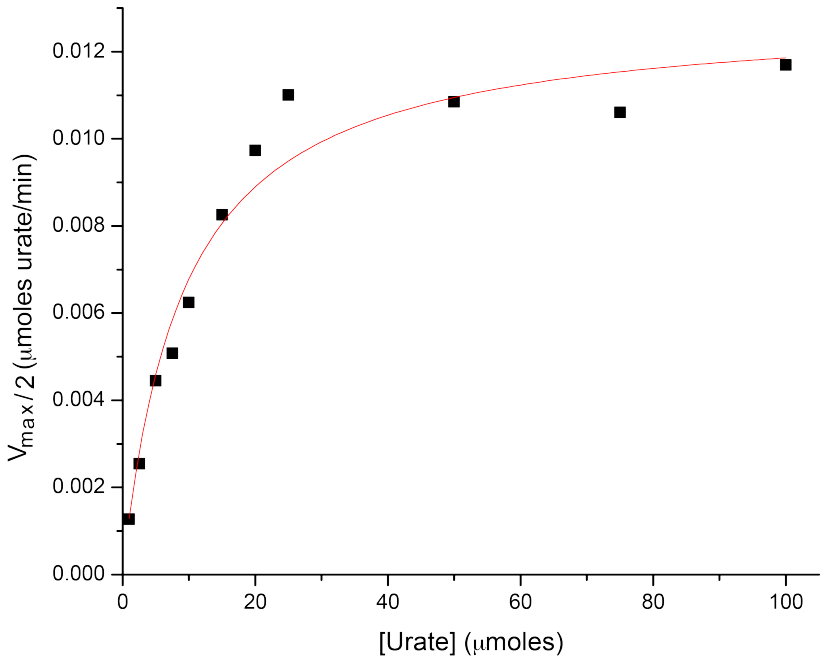


Figure 3.4 The activity of uricase proteins was assessed using UV-Vis spectrophotometry. The above plot represents the average of triplicate runs of enzymatic assays performed on a Cary 50 UV-Vis spectrophotometer. The uricase reaction is monitored by following the reaction via the disappearance of uric acid (as monitored at its emission maxima of 293 nm). Shown above is the average of triplicate runs at a urate concentration of 1 μM, 2.5 μM, 5 μM, 10 μM, 15 μM, 20 μM, 25 μM, 50 μM, 75 μM, and 100 μM plotted as the $V_0/[S]$. From these plots each enzyme's kinetic parameters: the V_{max} maximal velocity determined by hyperbolic regression and K_M the urate concentration at which the enzyme performs at half its maximum velocity.

3.3.2 Human uricase has accumulated deleterious mutations

As discussed in Chapter 2, *hUox* contains two nonsense mutations. To ensure that the human pseudogene could not encode for a functional enzyme these two nonsense mutations were replaced with codons for arginine (the residue in functional homologs). While it could be overexpressed, the human uricase in which the two nonsense mutations were mutated to arginines was insoluble and resistant to extraction in numerous tested buffer systems. In addition, no uricase activity could be detected in the resuspended inclusion bodies of the insoluble fraction. These results suggest that there are deleterious

mutations that are responsible for the inactivation of the human uricase in addition to the two nonsense mutations. The overarching therapeutic goal is to develop a "human-like" uricase and there are 17 amino acid differences between baboon, the last active primate uricase, and the human uricase (Figure S1). The biomedical goal of "humanizing" uricase by minimizing the amino acid differences from hUox while retaining a therapeutic level of activity presents a unique challenge. Since, hUox enzymatic activity cannot be directly rescued by replacing the two nonsense mutations the ASR approach was applied to the Uox family.

3.3.3 Resurrecting ancestral uricases

A multiple sequence alignment containing 17 uricase sequences from modern-day, or extant, sequences was used to generate a robust phylogenetic tree of the uricase protein family (Figure S1). From this MSA and implemented biological assumptions (see Materials and Methods), a consensus tree with branch lengths and posterior probabilities at each bifurcation was obtained using MrBayes (Figure 3.5.) This gene tree is in agreement with the evolution of mammals reported in the literature [108].

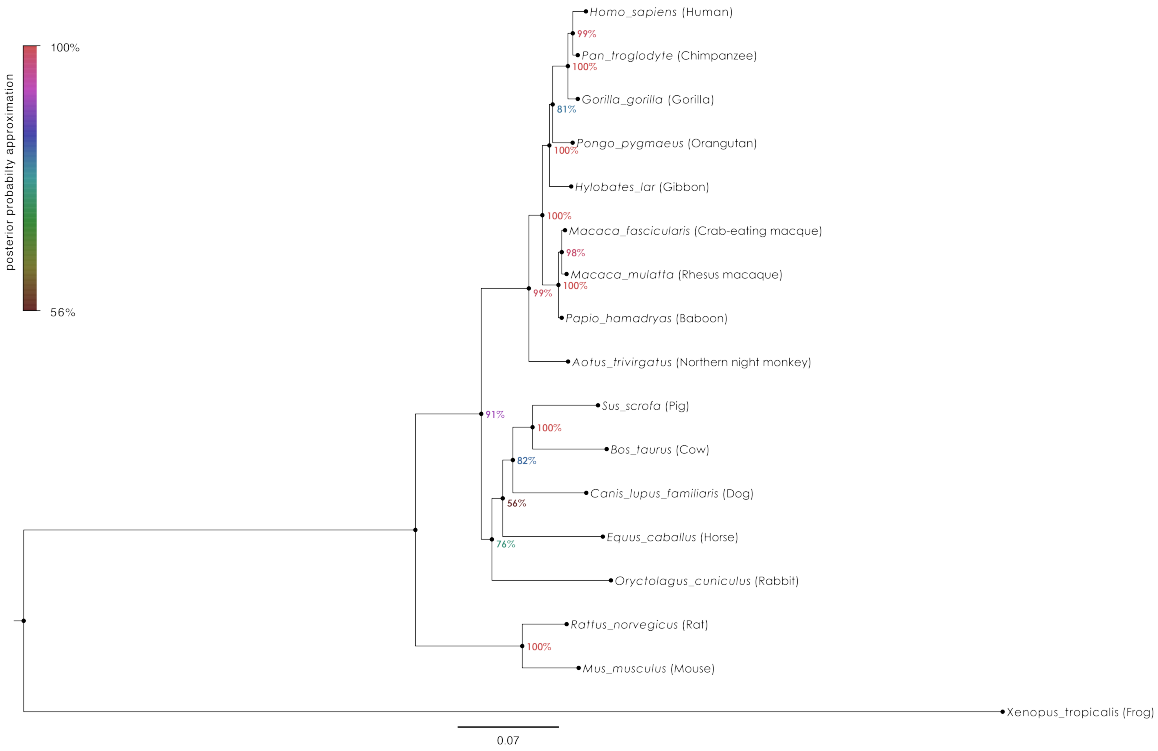


Figure 3.5 Phylogram of uricase genes constructed by the maximum likelihood method This phylogram of uricase proteins generated from a DNA MSA. Branch lengths are scaled to genetic distance between nodes. The internal nodes are labeled with the posterior probability, a measure of statistical support, for each bifurcation event. This generated gene tree agrees with the species tree for these organisms. (See Appendix Figure A.2. for a larger version of this figure)

PAML was used to infer the ancestral sequences with the topology shown above. Note: the models and the input data did not differ in their output of inferred ancestral sequences. These sequences are from a diverse set of taxa that represent mammals with a functional uricase, diminished uricase activity, and ultimately abolished uricase activity [48, 49]. Nine internal nodes were inferred from the uricase mammalian phylogeny; however, two of these internal nodes were conserved on the amino acid level. The first set of ancestors that coding for identical uricase proteins are An19 and An22, and only one ancestor, An19/22, was resurrected. The second set of ancestors coding for identical

uricase proteins are An32 and An33, again only ancestor, An32/33, resurrected (Figure 3.6).

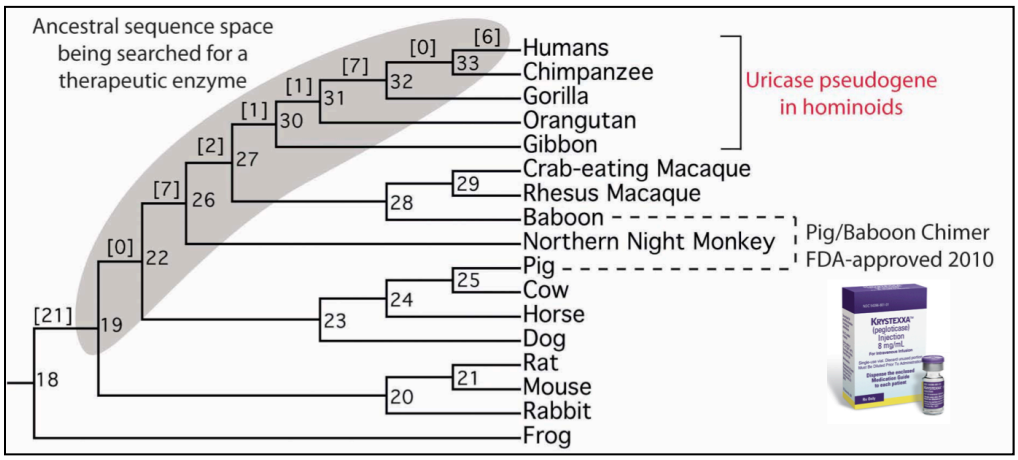


Figure 3.6 Cladogram tree showing the ancestral sequence space explored. The same topology (e.g. evolutionary relationship between uricase sequences) as in Figure 3.5 is shown – only the branches are no longer scaled for clarity and the common organismal names are used at the leaf nodes. The ancestral nodes are numbered (above a gray background) and represent the sequence space that has been searched for a novel uricase therapeutic. The protein sequence of nodes 19 and 22 is identical and is referred to as An19/22. In addition, the protein sequence of nodes 32 and 33, are identical as well and are referred to as An32/33. The numbers in square braces are the number of amino acid changes that occur along each branch (the lines connecting the ancestral nodes in the cladogram). Lastly, as a point of reference the Pig/Baboon Chimer (PBC) uricase sequence is shown on the phylogenetic tree, as well as the five hominoids which were used in ASR in which uricase is a pseudogene (human, chimpanzee, gorilla, orangutan, and gibbon).

The accuracy of the inferred ancestral mammalian uricases (excluding An18) ranged from 97.6 % to 99.9 % indicating high confidence in the predictions of ancient residues based on the models implemented (Table 3.1).

Table 3.1 The age of and accuracy of inferred uricase sequences. The time tree of life was utilized to approximate an age for the ancestral proteins studied in the laboratory [109].

Ancestral Node	Divergence Time (Ma)	Inference Accuracy (posterior probability)	Percent Identity to hUox
18	371.2	52.6 %	86.8 %
19/22	92.4	97.6 %	92.7 %
26	42.6	99.1 %	95.0 %
27	29.2	99.9 %	95.7 %
30	20.4	99.9 %	96.0 %
31	15.7	99.9 %	96.4 %
32/33	6.4	99.9 %	98.0 %

Uricolytic activity assays were first performed on inclusion bodies generated from the overexpression of each ancestral uricase. These assays involved the resuspension of the insoluble fraction in buffer to determine whether functional uricase was sequestered within these insoluble masses (Figure 3.7). These assays were performed in a 96-well plate format and could be measured by a multiplate reader. Therefore, we could get data simultaneously from multiple ancestral uricases. One will notice that An18, the oldest inferred uricase has been omitted from the cladogram in Figure 3.7 because it could not be experimentally characterized. While, An18 could be overexpressed it could not be isolated as a purified preparation under any of the conditions explored. Furthermore, the inclusion bodies from An18 overexpression were void of detectable uricase activity. Since, we are interested in a human-like (i.e. mammalian) uricase, it was not very disheartening that An18 was unable to be

characterized, because it likely would be highly immunogenic and unsuitable for therapeutic development. However, it was essential to include the amphibian sequence from our ASR protocol in order to confidently root the uricase tree. Working with ancestral proteins affords the researcher the opportunity to survey unique sequence space and provides insights into which human residues that can be tolerated in a functional uricase. Like their descendant proteins, ancestral uricases differ greatly in their activities and solubilities.

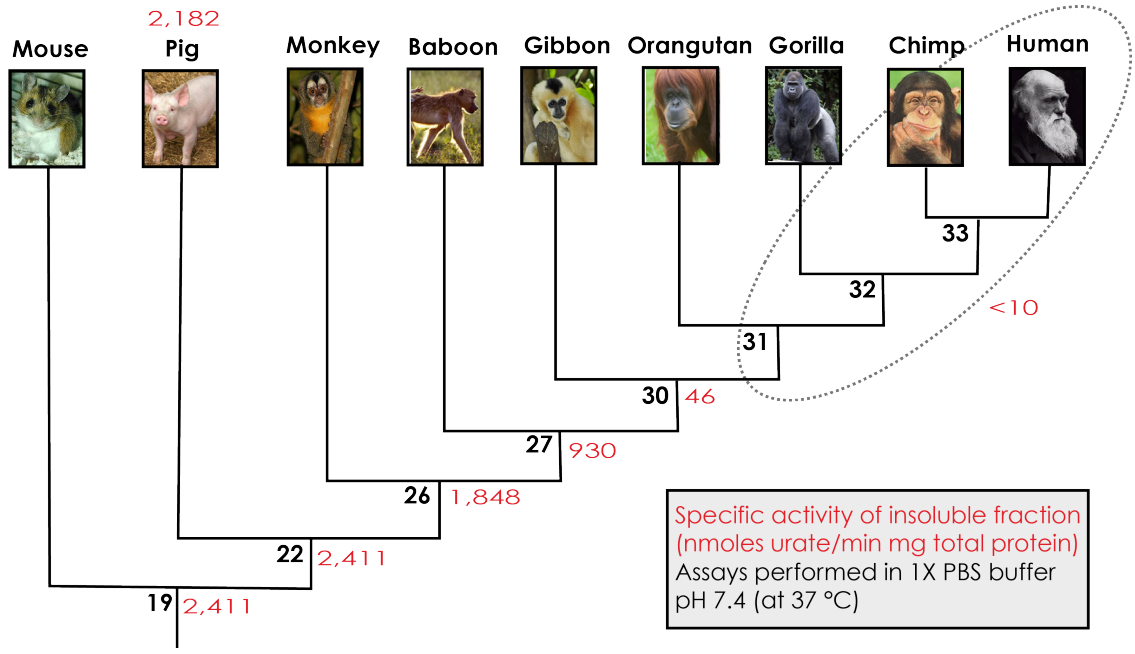


Figure 3.7. Enzymatically active mammalian uricase is trapped within the inclusion body produced during overexpression in *E. coli*.. Equal masses of resuspended insoluble fractions were assayed using the Amplex Red fluorometric assay for uricase activity. The soluble fractions had no observable activity in the assay (data not shown). The activity is expressed in terms of specific enzyme activity in units of nmoles urate/min mg of total protein. The modern day pig uricase has a comparable amount of uricase trapped within its inclusion body compared to An19/22. A substantial hit in activity is observed and An30 is just above the detection limit of the employed high throughput assay.

The ancestral uricases insoluble SEA serves as a proxy for their intrinsic activity. When assayed under physiological conditions (37 °C and pH 7.4), a

stepwise decrease in SEA along the branches from An19/22 to An26 and to An27 was observed. While An30 barely registered activity in the insoluble fraction and the remaining ancestral nodes of An31 and An32/33 were below the detection limit of the Amplex Red fluorometric uricase activity assay. It seems that not only are the more recent ancestral uricases largely inactive but they also present a unique experimental challenge due to their insolubility. These ancestral uricases were determined to be insoluble under an array of expression conditions. To further explore the effects of these apparent decreases in SEA in the insoluble fraction, the extraction efficiency of each ancestral protein was explored. While the uricase inclusion bodies are enriched with uricase, they do contain large uricase aggregates along with other contaminants. These contaminants are first removed by a series of detergent washes and then uricase is extracted with a high pH carbonate buffer, and large aggregates are subsequently removed by subsequently by size exclusion chromatography (SEC) (Figure 3.8).

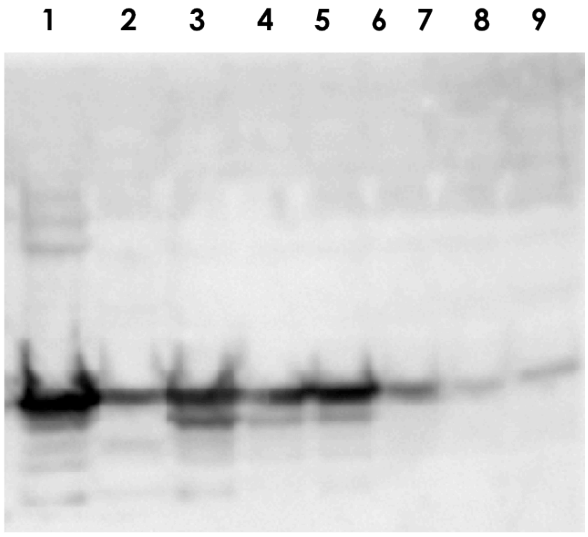


Figure 3.8 More recent ancestors are the least soluble under experimental conditions. Western blot using a 6xHis primary antibody conjugated to HRP. The samples were each expressed overnight and then subjected to a 4 hour

(Figure 3.8 continued)
extraction in 0.1 M carbonate, pH 11. Equal volumes were loaded into each lane. While the 6xHis hindered our ability to isolate enzymatically active ancestral proteins, this analysis provides a qualitative assessment of the decrease in solubility during the evolution of the uricase protein family. Lane 1) Pig uricase Lane 2) An18 Lane 3) An19/22, Lane 4) An26 Lane 5) An27 Lane 6) An30 Lane 7) An31 Lane 8) An32/33.

3.3.4 The structure of mammalian uricases

At the inception of this research program, no mammalian uricase structure had been published; therefore, we relied upon a homology model generated using SWISS Model [104]. To date no mammalian uricase structure has been published. This may be due to the poor solubility and the difficulty in obtaining crystals with high-resolution diffraction. However, as we seek to identify the effects of individual mutations, a solved structure of a mammalian uricase would prove invaluable. Our collaborators at Emory University (Dr. Eric Ortlund and Dr. Michael Murphy) solved the structure of the ancestral uricase variant An19/22_LysII (An19/22 with the following two lysines introduced by mutations R147K and E220K) was solved at a resolution of 2.4 Angstroms (Figure 3.9). The structure of An19/22 did not diffract to as high a degree as An19/22_LysII (personal communication). This crystal structure was instrumental in analyzing the effects of individual mutations and developing additional variant.

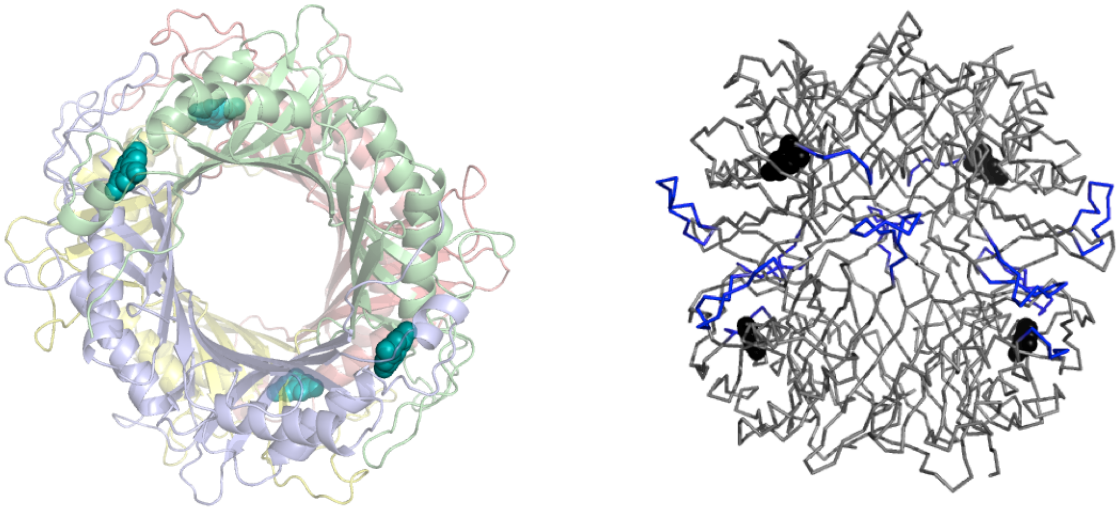


Figure 3.9 Solved mammalian crystal structure shows structural similarity to published microbial structures. (Left) Shown is the homology model of An19/22 uricase generated by Swiss Model. The amino acid sequence of An19/22 was thread onto the published structure of uricase from *Arthrobacter globiformus* (PDB: 2YZB). Each of the four monomers that make up the functional uricase are shown in the cartoon representation in a different color, and the bound substrate uric acid is shown as cyan spheres. (Right) Shown is the wire representation of the overlay of our solved mammalian uricase, An19/22_LysII, on-top of the microbial (*Bacillus sp, Aspergillus flavus, and Arthrobacter globiformus*) structures available from the Protein Data Bank (PDB: 1J2G, 3LD4 ,2YZB). Shown in blue are the variable regions located in the hinges. This variability is important to consider as the effects of individual mutations are considered in a structural context.

We have obtained kinetic data for purified An19/22, An26, and An27. However, An30, An31, and An32/33 were insoluble under tested conditions and showed negligible activity in their insoluble fraction (Table 3.2). An19/22 is the most promising of the ancient uricases to serve as a potential therapeutic. This enzyme displayed comparable kinetic and stability properties in parallel enzymatic preparations and assays [110]. Advantageously, An19/22 has a greater sequence identity to the human sequence than other therapeutic uricases (discussed below) and may elicit a weaker immune response when introduced to human patients and be a safer treatment for the management of gout.

The most active mammalian uricase is An19/22. This ancestor is inferred to have diverged from humans some 94 Mya old [109]. Its catalytic efficiency was determined by Michaelis-Menten kinetics experiments to be 1.4E6 $M^{-1}S^{-1}$, and has a stability in excess of 1 year when stored at 6 °C in carbonate pH 11 buffer (Figure 3.10). There are 21 amino acid differences between An19/22 and hUox.

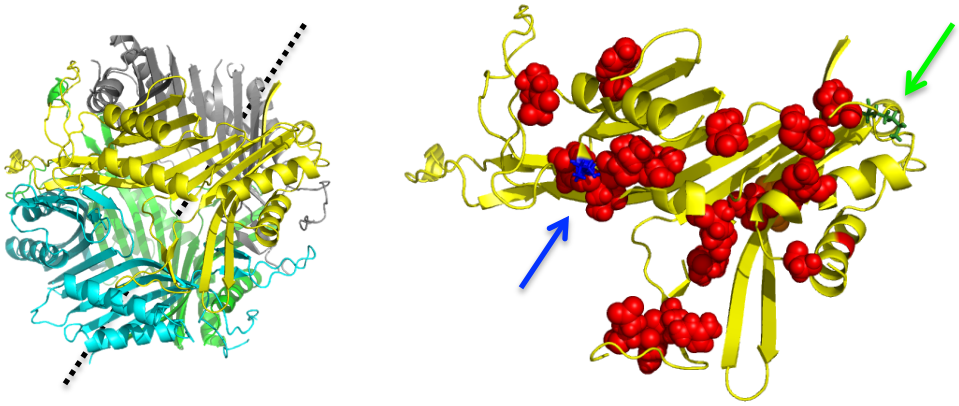


Figure 3.10 An19/22 is the common ancestor of placental mammals. This uricase protein is estimated to correspond to a protein that is 94 million years old. It possesses a high catalytic activity and a stability that makes it an ideal protein for further engineering. Shown at left is the active uricase tetramer, An19/22_LysII, where each identical monomer is shown as a different color: cyan, sea green, gray and yellow. The dashed line runs through the center of the barrel of the active tetramer. At right, is one monomer in the identical orientation as the yellow monomer in the tetrameric representation. The structure differs by two surface lysine residues that are both shown as sticks: site K147 is in blue and site K220 is in green. The 21 amino acids that differ from An19/22 and human uricase are shown as red spheres.

To establish the best conditions for working with An19/22 uricase, various buffer conditions were tested for its purification and long-term storage at 4 °C. These stability assays were routinely performed over the course of three months (even samples that were stored over a year displayed a minimal change in their specific activities). For An19/22, the greatest extraction efficiency, in terms of the specific activity of uricase activity recovered, was observed in 0.1 M carbonate buffer was at pH 11 was used for both extraction and SEC purification

Figure 3.11). This same extraction buffer was employed with other uricase constructs, and the extraction efficiency varied between variants (**Figure 3.8**). A correlation exists between those IBs isolated from preparations of uricases in which uricolytic activity could be detected and those ancestors that could be solubilized using a carbonate buffer.

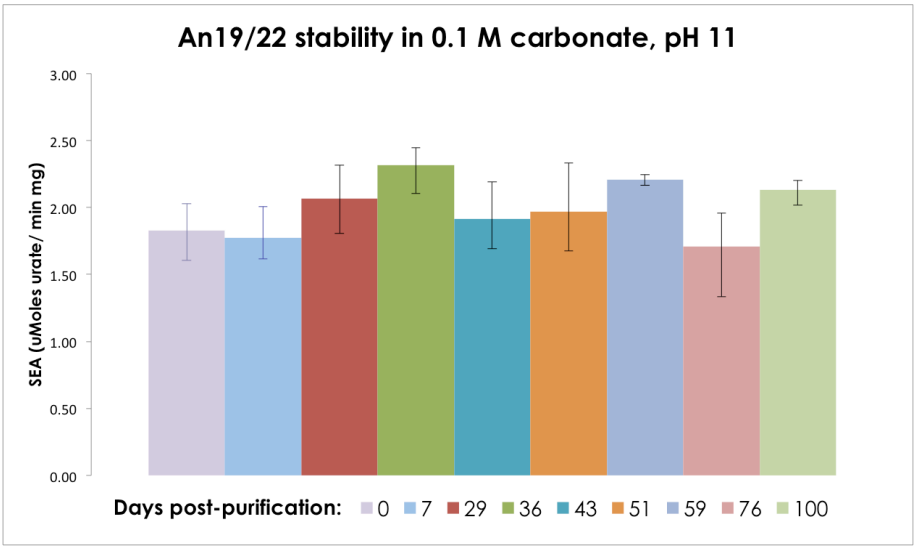


Figure 3.11 An19/22 extraction efficiency and storage buffer stability. Assayed with 100 µM urate in 1X PBS, pH 7.4. Samples were kept in their purification buffer at 4 – 6 ° C. Assays were performed in triplicate and error bars correspond to sampling range.

The next ancestor that was experimentally characterized was the common ancestor of primates, An26 (Figure 3.12). This uricase contains seven human residues in addition to those found in An19/22. The functional uricase was much more difficult to extract from An26 than it had been for An19/22. The catalytic efficiency of An26 is 7.7E5 $M^{-1}s^{-1}$, which is a reduction of approximately 50 % relative to An19/22. An26 is stable or at least 3 months at 4 ° C in 0.1 M carbonate buffer, pH 11.

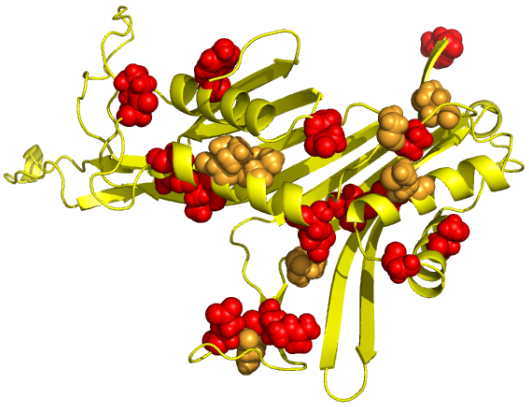


Figure 3.12 An26 is the common ancestor of primates. Shown is the monomer from our solved An19/22_LysII variant. The seven human residues that are introduced in An26 are highlighted in bright orange. The remaining 14 residues that differ from hUox are shown in red.

The last ancestor that could be isolated from its inclusion bodies and purified for enzymatic characterization is An27 (Figure 3.13). This ancestor is 29 Ma and only substituted one human residue with the consequence of dropping the catalytic efficiency by another order of magnitude to 7.9E+04 $M^{-1}s^{-1}$. Furthermore, An27 was too unstable to serve as a therapeutic (upon retesting 4 days after initial kinetics experiments there was no detectable activity with either the cuvette based assay, or the enzyme-coupled Amplex Red Assay (Invitrogen) which amplifies the uricase signal yielding a greater sensitivity.

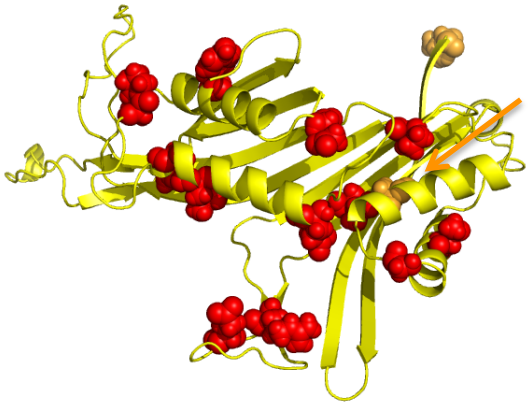
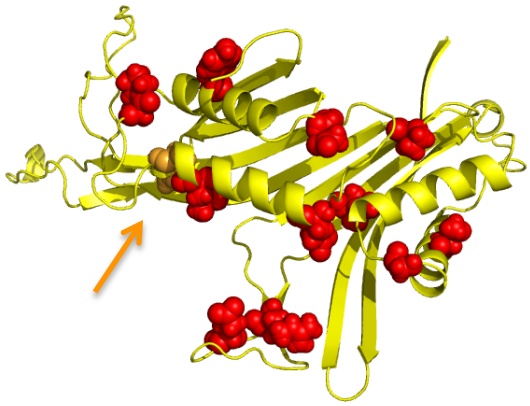


Figure 3.13 An27 is The common ancestor of the Old World monkeys and hominoids. Shown is the monomeric form our solved An19/22 variant (yellow).

(Figure 3.13 continued)
Two additional human residues are introduced in An27 (shown as orange spheres). These two mutations significantly comprised both the activity and the stability of An27. When the purified sample was re-assayed less than a week after purification, it was enzymatically dead – no activity could be detected by laboratory activity assay.

It had already been observed that An27 was an unstable protein, but purified functional uricase had been obtained for enzymatic characterization. An30 contains one additional human residue compared to An27 (Figure 3.14). This single point mutation yielded a protein that could not be solubilized from the inclusion body, even though its overexpression was confirmed by denaturing the IB. However, a small amount of uricase activity was determined in the insoluble fraction near the detection limit of the Amplex Red assay kit.



F**igure 3.14 An30 the common ancestor of apes is highly unstable.** Shown is the monomer (yellow) from our solved An19/22_LysII variant. The single human residue is introduced in An30 (shown as an orange sphere) is shown. This single mutation completely abolished the solubility of the An30 uricase in the employed carbonate buffers, and the activity of the insoluble fraction was at the detection limit.

The use of ASR to study ancient proteins can be a double-edged sword. On the one hand, you are able to explore sequence space that has the potential to bring about new functionality [79]. On the other hand, as we traversed the branches of the uricase phylogeny towards the human sequence,

we were unable to obtain soluble protein to characterize enzymatically. An30 barely gave a fluorescence signal above background. The most recent uricases An31, An32/33 and the hUox (in which the stop-codons have been mutated to arginines) uricase were inactive, or at the very least near the detection limit of the employed enzymatic assays.

An19/22 has the greatest activity under the employed experimental conditions (Figure 3.15). This ancient protein can also be extracted in its active form from inclusion bodies in high yield by overnight incubation at 4 °C in 0.1 M carbonate buffer. As evidenced by the results from each ancestor, as residues from hUox are incorporated, both the stability and activity are compromised. By increasing the "human-like" character of the uricases studied (i.e. resurrecting more recent ancestors), the catalytic efficiency drops by an order of magnitude from An26 to An27, and is completely demolished in An30.

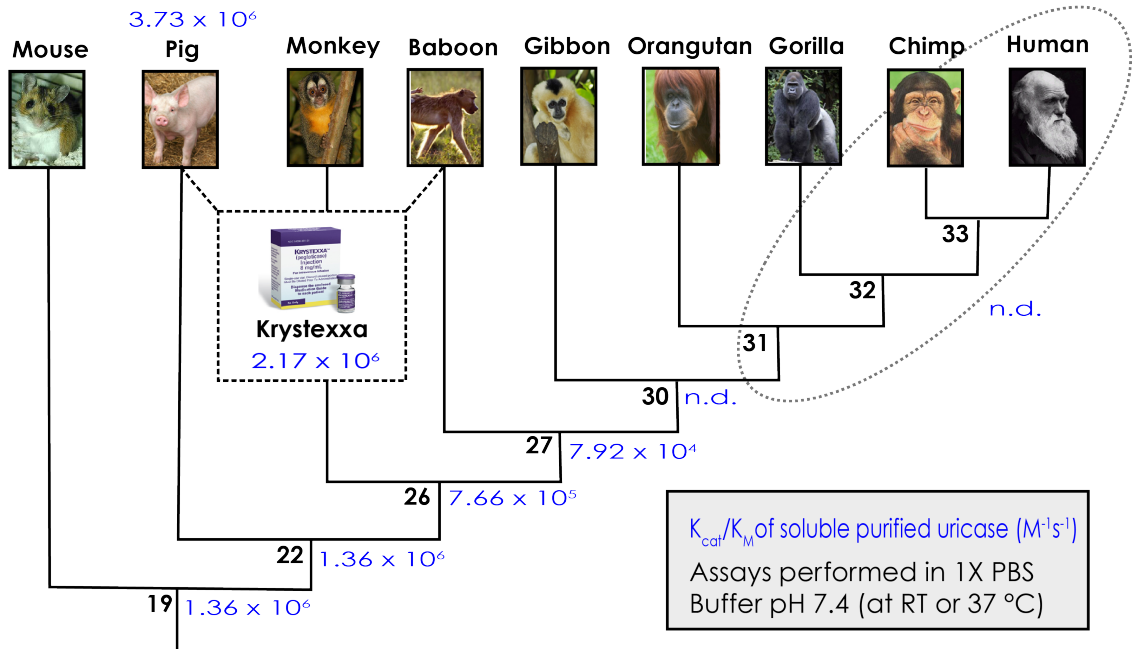


Figure 3.15 Soluble activity of purified tetrameric uricase. The catalytic efficiency (k_{cat}/K_M) for the modern-day pig uricase control, the PBC chimera (Krystexxa®),

(Figure 3.15 continued)
and 3 ancestral uricases were obtained. Assays were performed in triplicate in 1 mL reactions in a 1X PBS buffer at pH 7.4 and at 37 °C. There is a stepwise decrease in activity before the introduction of the two nonsense mutations, the first of which occurs on the branch leading from An30 to An31.

The main appeal of using ASR is that we can identify a uricase with a greater balance of human residues, while retaining a greater catalytic activity to have a therapeutic effect in patients. Considering the results obtained from resurrecting ancient mammalian uricases several observations can be made (Table 3.2). The first is that the most recent ancestors (i.e. An30, An31, An32/33) were insoluble under the experimental conditions employed. The next observation, is that last ancestor with detectable activity was An27, which was highly unstable, losing all of its activity within 4 days of purification. The absence of activity in the most recent uricases (An30, An31, An32/33, and hUox) has halted our path to obtaining a humanized functional uricase. In total there are twelve human residues that are found exclusively in these nonfunctional uricases. The fact that these "inactive" uricases all contain the destabilizing mutation present in An27, suggests that some of these dozen human residues might be tolerated in a more stable uricase background.

Table 3.2. Summary of modern and ancient uricases. Each uricase was assessed for its solubility in the 0.1 M carbonate, pH 11 buffer used for extracting uricase from inclusion bodies as either low, medium, or high based upon the extracted yield of protein as a fraction of total inclusion body. In addition, the results from Michaelis-Menten kinetics experiments are tabulated showing that Pig, PBC, and An19/22 all have a catalytic efficiency of 10^6 $M^{-1}s^{-1}$ and comparable stabilities at 4 °C in the extraction buffer of around 4 months. As additional human residues arise in the resurrected ancestors, the activities, as well as the stabilities decrease – the 12 human residues from An30 onward were not present in any functional resurrected uricase.

Uricase	Solubility	k_{cat}/K_M ($M^{-1}s^{-1}$)	Stability (4 °C)	Number of different residues relative to human sequence
Pig	Medium	4.1E+06	3-4 months	38
PBC	Medium	2.2E+06	> 4 months	32
An19/22	High	1.4E+06	> 4 months	22
An26	High	7.7E+05	3 months	15
An27	High	7.9E+04	3-4 days	13
An30	Low	n.d.	n.d.	12
An31	Low	n.d.	n.d.	11
An32/33	Low	n.d.	n.d.	6
Human	Low	n.d.	n.d.	n/a

3.3.5 Properties of chimeric uricases

To this point, we have focused exclusively upon our main objective of identifying a "human-like" uricase for the treatment of gout. We had not considered any mutations that were not present in the human sequence. We were encouraged by the stability of An19/22 and next sought to explore whether or not this variant could tolerate additional modifications. We turned our attention to the literature to identify specific mutations that were reported to have a stabilizing effect upon other uricases. The rationale being that those mutations might counterbalance the destabilizing effect that some of the

human residues impart on the protein. The first mutant of An19/22 uricase incorporated three mutations that had been identified in *Candida utilis* to enhance oxidative stability and thermostability: K175R, Q177E, and F178Y [111]. The second round of mutations incorporated a set of mutations in An19/22 uricase that were reported to improve the *Bacillus sp. (TB-90)* specific activity: Y265L, I276G and D286S [112]. With the final variant we sought to fully test An19/22's resilience to mutations by mutating a number of residues to those found in the soybean uricase (*Glycine max*): Y111N, V112I, E113V, K158E, K161S, F178Y, V184T, K185R, F189V, Q192E, Y194T, C195A and K196L.

While the previous variants sought to enhance the stability and/or activity of An19/22, this next set of variants was focused on introducing solvent-accessible lysine residues that might prove useful for conjugating moieties to the protein surface at later stages of development (e.g. PEGs). Furthermore, charged surface residues (i.e. lysines) can act as gatekeepers to prevent aggregation of the hydrophobic patches on a protein's surface [113]. By aligning mammalian uricase sequences, we identified sites in homologs where a lysine was introduced and based upon our homology model selected those sites that were likely on the protein's surface. The first lysine variant is, An19/22_LysI, introduces three lysines on the protein surface with the following three mutations: N103K, E230K, R303K. The second lysine variant, An19/22_LysII introduces two surface lysines with by mutating sites R147K and E220K. These lysine variants were expressed, purified, and characterized using our standard laboratory protocol.

Table 3.3. Using An19/22 as a stable backbone for exploring non-humanizing mutations. The first set of variants LysI and LysII sought to introduce additional lysine residues on the protein surface. The introduction of these sites will be valuable for later modification and it was important to determine whether they had a deleterious effect upon the protein's affinity for the substrate and catalytic efficiency. The USPTO variant introduced 3 residues that were reported to improve the stability of a bacterial uricase and resulted in a 50 % decrease in catalytic efficiency compared to the An19/22 background. Both variant JPO and *Glycine max* (*G. max*) introduced to many "foreign" mutations that led to uricases with no detectable activity.

An19/22 Variants	Mutations	K_M (M^{-1} urate)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)
Unmodified reference		1.34E-05	5.46	4.07E+05
LysI*	N103K, E230K, R303K	1.92E-05	4.62	2.41E+05
LysII*	R147K, E220K	1.90E-05	6.29	3.31E+05
USPTO	K175R, Q177E, F178Y	1.22E-05	2.45	2.01E+05
JPO	Y265L, I276G, D286S	**Purified tetrameric enzyme, but it was inactive.**		
G.max	Y111N, V E113V, K158E, K161S, F178Y, V184T, K185R, F189V, Q192E, Y194T, C195A, K196L	**Could not solubilize under reaction conditions. Insoluble, and no activity detected in the insoluble fraction.**		

3.4 CONCLUSIONS

Working with ancestral uricases effectively allowed us to travel back in time to the point before the uricase protein became functionally absent in hominids. It was the longstanding belief that the two nonsense mutations were solely responsible for the inactivation of the human uricase. However, we demonstrated that there was a significant decrease in uricolytic activity prior to An30, when the two premature stop codons are thought to have occurred. This data supports the gradual step-wise decrease in the activities of uricases among mammals, especially those that we have studied so far.

By working with these ancestral uricases, we have shed some light upon the history of uricase inactivation in the human lineage. In addition to exploring this evolutionary story we have also identified two "human-like" uricases that have a greater sequence identity to the hUox than Krystexxa®, the PBC uricase. The first is An19/22 which has a comparable *in vitro* activity and stability, at 4 ° C, to PBC uricase. This most ancient resurrected uricase, An19/22, contains 22 amino acid differences from hUox - versus PBC uricase, which differs from the hUox by 32 amino acids. Furthermore, while less catalytically active, An26 contains only 15 amino acid differences from hUox, and while it has an order of magnitude lower catalytic activity compared to An19/22 it is still quite stable and may be another viable lead for a safer uricase therapeutic.

3.5 REFERENCES

43. Mahler, H.R., G. Hubscher, and R. Baum, *Studies on uricase. I. Preparation, purification, and properties of a cuproprotein.* J. Biol. Chem., 1955. **216**(2): p. 625-641.

48. Logan, D.C., D.E. Wilson, C.M. Flowers, P.J. Sparks, and F.H. Tyler, *Uric acid catabolism in the woolly monkey.* Metabolism., 1976. **25**(5): p. 517-522.

49. Usuda, N., M.K. Reddy, T. Hashimoto, M.S. Rao, and J.K. Reddy, *Tissue specificity and species differences in the distribution of urate oxidase in peroxisomes.* Lab. Invest., 1988. **58**(1): p. 100-111.

54. Hershfield, M. and S.J. Kelly. Urate oxidase. 7056713 USPTO application. 2006.

64. Benner, S.A., S.O. Sassi, and E.A. Gaucher, *Molecular Paleoscience: Systems Biology from the Past*, in *Advances in Enzymology*, E.J. Toone, Editor. 2010, John Wiley & Sons, Inc. p. 1-132.

79. Kratzer, J.T., Cole, M.F., and E.A. Gaucher, *Protein engineering guided by natural diversity*, in *Protein Engineering Handbook*, S. Lutz and U.T. Bornscheuer, Editors. 2013, Wiley-VCH: Weinheim, Germany.

80. Pauling, L. and E. Zuckerkandl, *Chemical Paleogenetics Molecular Restoration Studies of Extinct Forms of Life.* Acta Chem. Scand., 1963. **17**: p. S9-S16.

81. Stackhouse, J., S.R. Presnell, G.M. Mcgeehan, K.P. Nambiar, and S.A. Benner, *The Ribonuclease from an extinct bovid ruminant.* FEBS Lett., 1990. **262**(1): p. 104-106.

82. Yokoyama, S. and Y.S. Shi, *Molecular analysis of the evolutionary significance of ultraviolet vision in vertebrates.* Proc. Natl. Acad. Sci. U. S. A., 2003. **100**(14): p. 8308-8313.

83. Thornton, J.W., E. Need, and D. Crews, *Resurrecting the ancestral steroid receptor: Ancient origin of estrogen signaling.* Science, 2003. **301**(5640): p. 1714-1717.

84. Gaucher, E.A., S. Govindarajan, and O.K. Ganesh, *Palaeotemperature trend for Precambrian life inferred from resurrected proteins.* Nature, 2008. **451**(7179): p. 704-707.

85. Chen, F., E.A. Gaucher, N.A. Leal, D. Hutter, S.A. Havemann, S. Govindarajan, E.A. Ortlund, and S.A. Benner, *Reconstructed evolutionary adaptive paths give polymerases accepting reversible terminators for sequencing and SNP detection.* Proc. Natl. Acad. Sci. U. S. A., 2010. **107**(5): p. 1948-1953.

86. Thornton, J.W., *Resurrecting ancient genes: experimental analysis of extinct molecules.* Nat. Rev. Genet., 2004. **5**(5): p. 366-375.

87. Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-2948.

88. Notredame, C., D.G. Higgins, and J. Heringa, *T-Coffee: A novel method for fast and accurate multiple sequence alignment.* J. Mol. Biol., 2000. **302**(1): p. 205-217.

89. Mount, D.M., *Bioinformatics: Sequence and Genome Analysis* 2nd ed. 2004, Cold Spring Harbor, NY.: Cold Spring Harbor Laboratory Press.

90. Fitch, W.M., *Toward defining the course of evolution: minimum change for a specific tree topology.* Syst. Zool., 1971. **20**(4): p. 406-416.

91. Yang, Z., S. Kumar, and M. Nei, *A new method of inference of ancestral nucleotide and amino acid sequences.* Genetics, 1995. **141**(4): p. 1641-1650.

92. Huelsenbeck, J.P., F. Ronquist, R. Nielsen, and J.P. Bollback, *Bayesian inference of phylogeny and its impact on evolutionary biology.* Science, 2001. **294**(5550): p. 2310-2314.

93. Di Giulio, M., *The universal ancestor was a thermophile or a hyperthermophile: Tests and further evidence.* J Theor Biol, 2003. **221**(3): p. 425-436.

94. Woese, C.R., *Bacterial evolution.* Microbiol. Rev., 1987. **51**(2): p. 221-271.

95. Tawfik, D.S., S. Bershtein, and K. Goldin, *Intense neutral drifts yield robust and evolvable consensus proteins.* J. Mol. Biol., 2008. **379**(5): p. 1029-1044.

96. Tawfik, D.S. and N. Tokuriki, *Stability effects of mutations and protein evolvability.* Curr Opin Struc Biol, 2009. **19**(5): p. 596-604.

97. Perez-Jimenez, R., A. Ingles-Prieto, Z.M. Zhao, I. Sanchez-Romero, J. Alegre-Cebollada, P. Kosuri, S. Garcia-Manyes, T.J. Kappock, M. Tanokura, A. Holmgren, et al., *Single-molecule paleoenzymology probes the chemistry of resurrected enzymes.* Nat. Struct. Mol. Biol., 2011. **18**(5): p. 592-596.

98. Gaucher, E.A., J.M. Thomson, M.F. Burgan, and S.A. Benner, *Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins.* Nature, 2003. **425**(6955): p. 285-288.

99. Joudrier, P., M.F. Gautier, F. de Lamotte, and K. Kobrehel, *The thioredoxin h system: potential applications.* Biotechnol. Adv., 2005. **23**(1): p. 81-85.

100. Tavare, S., *Some probabilistic and statistical problems in the analysis of DNA sequences.* Lectures Math. Life Sci., 1986. **17**: p. 57-86.

101. Lio, P. and N. Goldman, *Models of molecular evolution and phylogeny.* Genome Res., 1998. **8**(12): p. 1233-1244.

102. Yang, Z., *PAML 4: phylogenetic analysis by maximum likelihood.* Mol. Biol. Evol., 2007. **24**(8): p. 1586-1591.

103. Jukes, T.H. and C.R. Cantor, *Evolution of Protein Molecules*. 1969, New York: Academic Press.

104. Arnold, K., L. Bordoli, J. Kopp, and T. Schwede, *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.* Bioinformatics, 2006. **22**(2): p. 195-201.

105. Schrodinger, LLC, *The PyMOL Molecular Graphics System, Version 1.3r1*, 2010.

106. London, M. and P.B. Hudson, *Uricolytic activity of purified uricase in two human beings.* Science, 1957. **125**(3254): p. 937-938.

107. Conley, T.G. and D.G. Priest, *Purification of uricase from mammalian tissue.* Prep. Biochem., 1979. **9**(2): p. 197-203.

108. Delsuc, F., G. Tsagkogeorga, N. Lartillot, and H. Philippe, *Additional molecular support for the new chordate phylogeny.* Genesis, 2008. **46**(11): p. 592-604.

109. Hedges, S.B., J. Dudley, and S. Kumar, *TimeTree: a public knowledge-base of divergence times among organisms.* Bioinformatics, 2006. **22**(23): p. 2971-2972.

110. Kelly, S.J., M. Delnomdedieu, M.I. Oliverio, L.D. Williams, M.G.P. Saifer, M.R. Sherman, T.M. Coffman, G.A. Johnson, and M.S. Hershfield, *Diabetes insipidus in uricase-deficient mice: A model for evaluating therapy with poly(ethylene glycol)-modified uricase.* J. Am. Soc. Nephrol., 2001. **12**(5): p. 1001-1009.

111. Koyama, Y. and T. Ichikawa. Mutant uricase, a mutant uricase gene, a novel recombinant DNA, and a process for producing mutant uricase. US 5376545 US 5700674, USPTO application. 1997.

112. Yoshiaki, N., T. Astsushi, K. Takahide, and H. Takao. Method for improving stability of uricase and modified uricase having improved stability. 198289 JPO application. 2006.

113. Reumers, J., F. Rousseau, and J. Schymkowitz, *Multiple evolutionary mechanisms reduce protein aggregation.* Open Biol., 2009(2): p. 176-184.

CHAPTER 4: OPTIMIZING A HUMAN-LIKE URICASE

From the following publication in preparation:

Kratzer, J.T., Murphy, M. N., Ortlund, E. A. & Gaucher, E. A. Evolutionary history of modern and ancient mammalian uricases.

4.1 INTRODUCTION

The goal of this chapter is to take the most active resurrected uricase, An19/22, and begin to individually incorporate human residues into its backbone. By examining each individual point mutation, will allow us to better pinpoint which mutations are the most least deleterious and should be incorporated in humanized variants of An19/22, and conversely those mutations that are the most deleterious and should be avoided entirely. Furthermore, by robustly stepping through the mutations that occurred in the This will also help to parse out the specific mutations that occurred after the nonsense mutations arose and explain why in Chapter 3 simply replacing the stop-codons with arginines was insufficient to restore activity to hUox.

4.2 MATERIALS AND METHODS

4.2.1 Synthesis of branch mutants

Including one revertant there are 22 amino acid differences between An19/22 and hUox; therefore, 21 variants were synthesized and cloned into the pET21A+ expression vector using the *Nde*I and *Xho*I restriction sites.

4.2.2 Protein expression and purification

The protocol utilized with ancestral uricase was again applied was employed for the IPTG-induced expression, and SEC purification of each variant (3.2.3). For each of the single mutants a 250 mL LB culture was expressed and the insoluble IB, after cell lysis, was isolated by centrifugation. The IB was washed with BB detergent dilutions and then extracted into 0.1 M Na_2CO_3 pH 11 for 4 hours at 4 °C. After extraction, the uricase in the soluble fraction was clarified by centrifugation and purified by SEC.

4.2.3 Enzymatic activity assay

The enzyme activity assays were performed as described previously (3.2.4).

4.2.4 Quantifying effect of individual branch mutations

Several properties of each branch mutation were utilized to quantify the effect that the effect of each human residue substitution had relative to the robust An19/22 background. The first metric considered was the ratio of tetrameric uricase to other larger aggregates. An approximation for this ratio was determined by using the peak integration feature of the evaluation module of the Unicorn™ chromatography software (GE Life Sciences). This feature integrates the UV absorbance at 280 nm as the protein is eluted from the size-exclusion column. Specifically, the total area in milliabsorbance units (mAU) that corresponds to tetrameric uricase, which elutes around 69 mL, divided by the protein absorbance that elutes before the tetrameric uricase peak. The specific activity for each uricase preparation was determined by running triplicate cuvette-based assays at 100 **μM** urate and 37 °C. The last parameter used to

characterize these single mutant variants was the catalytic efficiency, which was determined by doing a full set of Michaelis-Menten kinetics experiments.

4.2.5 Humanizing ancestral uricase

Based upon the characterization of the individually introduced mutations in An19/22, additional humanized variants were synthesized (Table 4.2). The sites chosen where those that promoted the purification of tetrameric uricase versus larger aggregates, a specific enzyme activity close to that of An19/22, and where the catalytic activity was marginally diminished relative to the parent An19/22. These An19/22-Human chimeras were then purified using standard methods and characterized by enzymatic assays.

4.3 RESULTS AND DISCUSSION

4.3.1 An19/22-Human chimeras: introducing solvent-accessible human residues into the An19/22 background

The most active ancestral uricase, An19/22, had the highest stability and activity under the assayed conditions. The rationale behind selecting the stable and catalytically active An19/22 as the backbone to test subsets of surface accessible branch mutations was that the effects of these mutations would be easily interpreted. Therefore, we introduced several of the mutations from the more recent ancestral uricases to determine if they would be tolerated within the An19/22 background. To further explore the effects of these mutations, several variants were designed in which a few select mutations were introduced into the background of An19/22. Sites selected for mutation included those that were solvent accessible, and not at the oligomerization interfaces (dimer or tetramer). We primarily focused on surface residues because of their potential

role in an immune response, and how they might impact aggregation of the uricase protein. The first variant addresses a subset of mutations along the branch of An32 to the human protein (Figure 4.3). In variant A19/22-26 three of the 7 mutations that occur along the branch from An19/22 to An26 were introduced: M92V, M146L, and G202C. The second variant An19/22_31-32 mutates two of the seven sites that occur along the branch of An31 to An32. A third variant An19/22_32-H with mutations E121G and P233T was also explored.

In variant An19/22_19-26, three out of the six mutations that occur along the branch from the common ancestor of non-primate mammals (An19/22) to the common ancestor of primates (An26) were introduced: M92V, M146L, G202C. This resulted in a protein that was insoluble under experimental conditions and whose insoluble fraction had no detectable catalytic activity. However, by introducing these 3 mutations, we actually observed a reduction in activity by one order of magnitude. Perhaps the effect of all 7 mutations allows for some compensation of the deleterious effect observed with these chosen three.

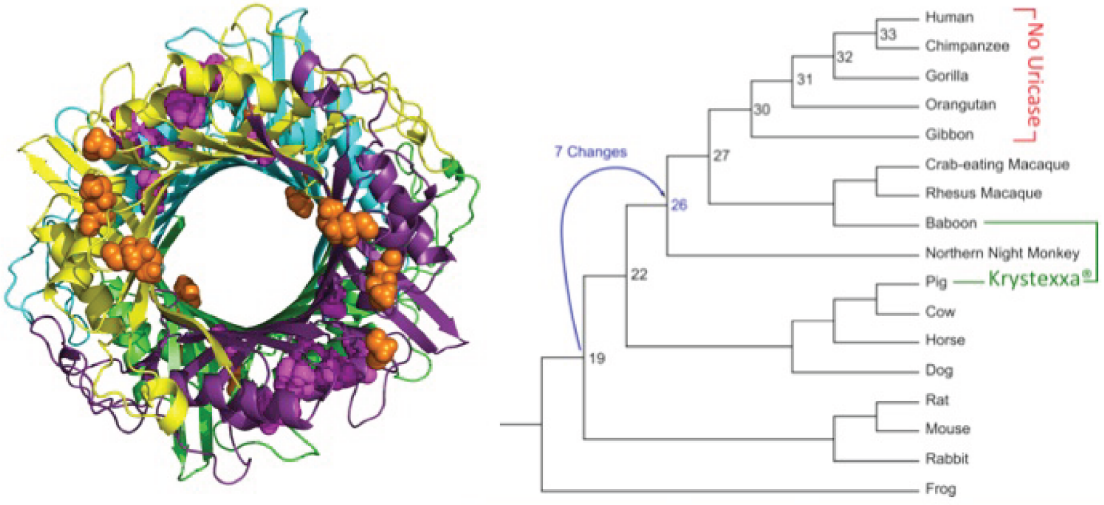


Figure 4.1 Dissecting branch changes: An19/22_19-26 Shown is the crystal structure of our An19/22_LysII variant. The homotetramer is shown with each identical chain shown in the cartoon representation in a different color. The seven mutations that occur along the branch from An32-H are shown as spheres on one uricase monomer (purple). The subset of mutations introduced in variant An19/22_19-26 is shown in orange.

In variant An19/22_31-32, two out of the seven mutations that occur along the branch from the common ancestor of hominoids (An31) to the common ancestor of the great apes (An32/33), G83E and E208K, were introduced (Figure 4.2). This resulted in a protein that had a diminished catalytic efficiency by two orders of magnitude from 4.07E+05 $M^{-1}s^{-1}$ for An19/22 to 1.05E_04 $M^{-1}s^{-1}$ when the two sites were introduced (Table 4.1). However, it was a functional protein with the combination of these sites introduced into An19/22; whereas, An32 itself was both insoluble and catalytically inactive.

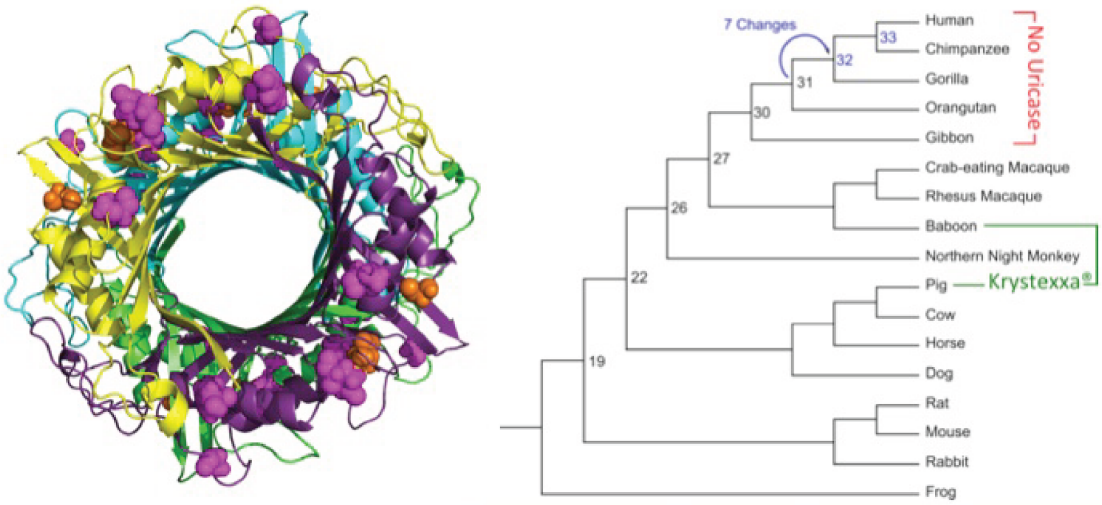


Figure 4.2 Dissecting branch changes: An19/22_31-32 Shown is the crystal structure of our An19/22_LysII variant. The homotetramer is shown with each identical chain shown in the cartoon representation in a different color. The seven mutations that occur along the branch from An32-H are shown as spheres on one uricase monomer (purple). The subset of mutations introduced in variant An19/22_31-32 is shown in orange.

In variant An19/22_32-H, two out of the six mutations that occur along the branch from the common ancestor of great apes (An32) to humans the E121G and P233T are introduced (Figure 4.3). This resulted in a protein that was insoluble under experimental conditions and suggested that these two sites in combination were deleterious (Table 4.1).

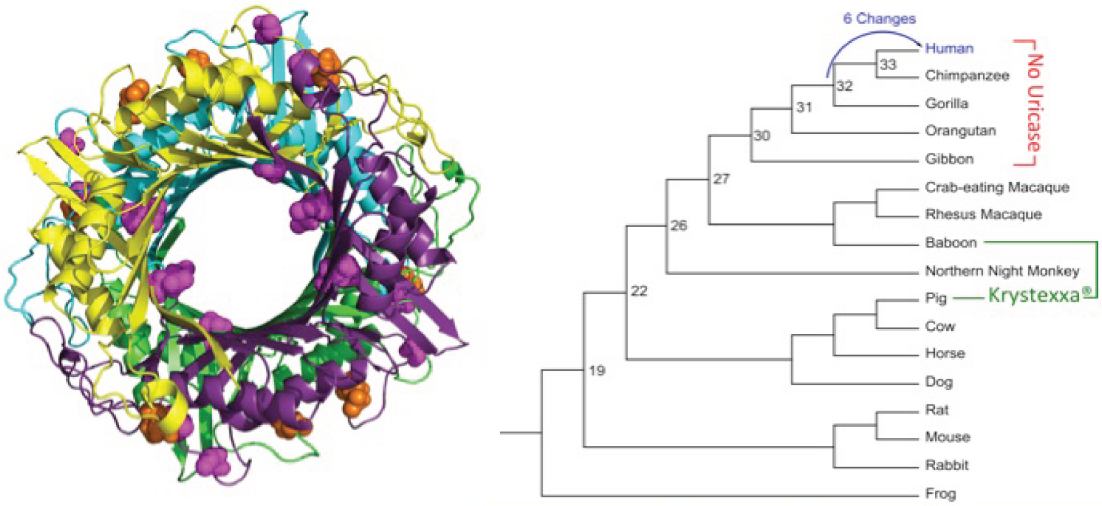


Figure 4.3 Dissecting branch changes: An19/22_32-H. At left is the solved crystal structure of An19_LysII (with each monomer shown as a different color. The seven mutations that occur along the branch from An32-H are shown as spheres along the purple uricase monomer. The subset of mutations introduced in variant An19/22_32-H is shown in orange.

Table 4.1 Summary of An19/22-Human chimeras' kinetics Given the high stability and catalytic activity of An19/22 several variants were constructed to introduce mutations from later branches in the phylogeny.

Variant	Mutations in An19/22	K_M (M^{-1} urate)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)
An19/22 unmodified		1.34E-05	5.46	4.07E+05
19-26	M92V, M146L, I217L	1.81E-05	2.15	1.19E+05
31-32	G83E, E208K	3.50E-05	0.37	1.05E_04
32-H	E121G, P233T	**Unstable and aggregates readily**		

It became apparent that by introducing only two mutations in the An19/22_3-H variant. By simply including two point mutations that occur along the branch from An32 to the human protein, the recovered variant was highly unstable. Thus revealing that An19/22 stability and activity to be quite sensitive to a single set of mutations (which are not in the active site).

4.3.2 Exhaustively testing the effects of individual branch mutations within the An19/22 background

In the first An19/22 variants only the solvent accessible residues were considered. This focus was largely governed by the erroneous assumption that those surface residues play a major role in the immune response to Uox However, the human body's mechanism for generating antibodies to non-self proteins often involves their phagocytosis and processing [114, 115]. Therefore, any residue within the 304 amino acid sequence of a recombinant uricase could serve as an epitope to which and immune response is mounted. Since An19/22 is the most robust ancestor characterized (in terms of both its activity and stability profile), it was chosen as the template into which each of the 21 mutations that occur along the phylogeny from An19/22 to the human sequence. The three properties studied were the ratio of tetramer to larger aggregates therapeutic uricases are more effective as highly purified tetramers [116], and large uricase aggregates can lead to rapid uricase clearance [117]. The next property reported is

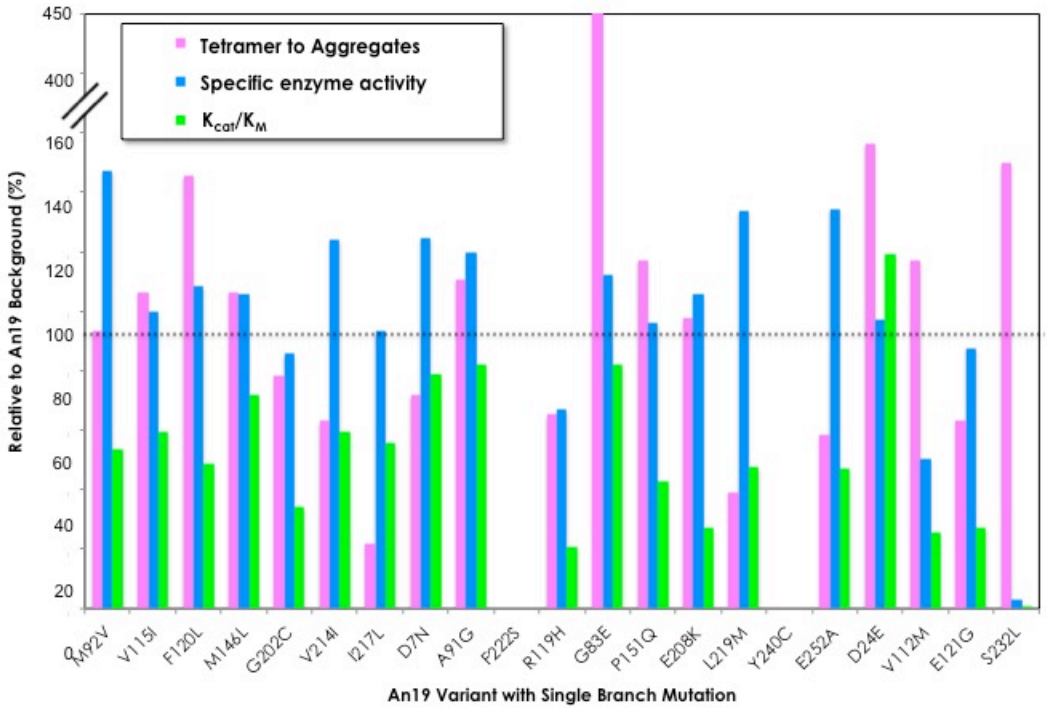


Figure 4.4 Individual branch mutations in the stable An19/22 showed that single human residues could completely kill activity. To better understand the effects of individual mutations, which occur during the uricase protein's evolution, each of the 21 point mutations (the number of differences between An19/22 and the human uricase pseudogene) were made. Each variant was characterized as follows: 1) The proportion of tetrameric uricase that was obtained from SEC purification (pink bar); 2) The specific enzyme activity of the purified preparation (blue bar); 3) The catalytic efficiency (k_{cat}/K_M) of the purified single mutant. All of the reported values are relative to the An19/22 protein (shown by the dashed line at 100 %). Two mutations, F222S and Y240C completely abolished activity and the ability to purify the An19/22 mutants. The single An19/22 mutant with the S232L could be purified, however, its relative activity was severely diminished.

Table 4.2 Effects of individual branch mutations in An19/22 background. Below are summarized the results of each single branch mutation relative to its unmodified An19/22 background.

Branch	Mutation	K_M (M^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)	Relative enzyme efficiency (Anc19 = 100 %)
An19		9.10E-06	6.42	7.07E+05	100 %
26	M92V	2.72E-05	10.3	3.80E+05	54 %
	V115I	1.46E-05	6.11	4.19E+05	59 %
	F120L	2.13E-05	7.33	3.45E+05	49 %
	M146L	1.35E-05	6.88	5.09E+05	72 %
	G202C	1.58E-05	3.79	2.41E+05	34 %
	V214I	2.24E-05	9.43	4.21E+05	60 %
	I217L	1.44E-05	5.66	3.95E+05	56 %
27	D7N	1.44E-05	8	5.56E+05	79 %
	A91G	1.37E-05	7.95	5.82E+05	82 %
30	F222S	Not Active			
31	R119H	3.63E-05	5.31	1.46E+05	21 %
32	G83E	1.32E-05	7.64	5.80E+05	82 %
	P151Q	1.84E-05	5.54	3.03E+05	43 %
	E208K	1.84E-05	3.49	1.91E+05	27 %
	L219M	3.02E-05	10.12	3.03E+05	48 %
	Y240C	Unable to purify			
	E252A	2.83E-05	9.4	3.33E+05	47 %
H	D24E	7.46E-06	6.29	8.44E+05	119 %
	V112M	1.95E-05	3.5	1.80E+05	26 %
	E121G	3.37E-05	6.47	1.93E+05	27 %
	S232L	6.43E-05	0.32	5.05E+03	1 %
	P233T	1.44E-05	3.49	2.44E+05	34 %

One of the likely reasons that the replacement of the two stop codons with arginines did not rescue uricase activity is that once hUox became a pseudogene, it was free from selective pressure and accumulated additional deleterious mutations. Three single mutants (F222S, S232L, and Y240C) and exhibited significant decreases in activity relative to An19/22 (Figure 4.4 and Table 4.2).

4.3.3 Human residue that kills the activity of An19/22

The mutation S232L is located within a loop and has a solvent accessible surface area (SASA) of 6 % calculated using the solvent accessible calculation program GET AREA [118]. The surrounding residues that may make contact are

S168, Y231, and P233. Through the introduction of a nonpolar residue, this polar environment may be responsible for the abolished activity. Furthermore, it resulted in the lowest relative enzyme efficiency of any of the active branch mutations.

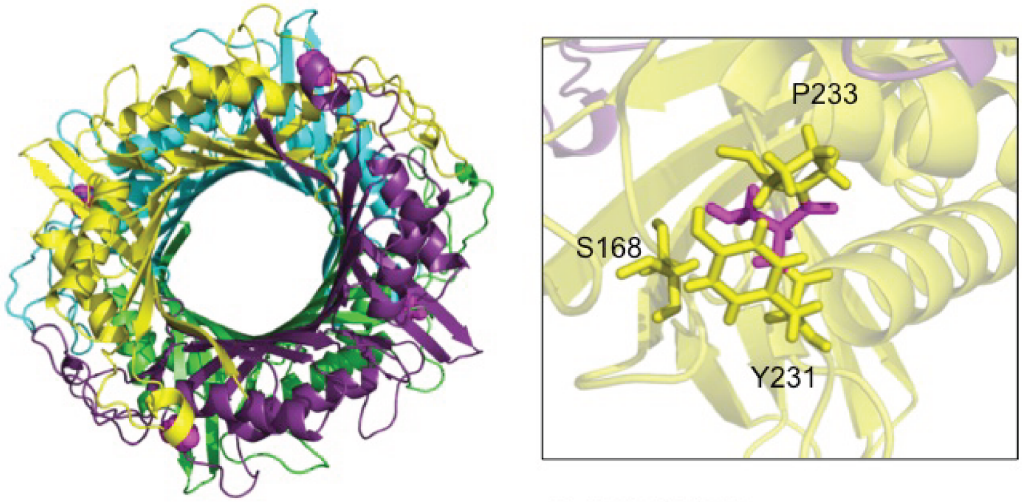


Figure 4.5 Deactivating mutation S232L is only present in the human pseudogene. At the left is the An19/22_LysII crystal structure showing the global context of this mutation at the dimer interface of two uricase monomer subunits. Site 232 is shown as spheres within one monomer (purple). This residue is located at the dimer interface of two uricases and its local environment includes two polar resides: S168 and Y231. The introduction of a threonine at this position is a non-conservative mutation.

The deleterious mutation Y240C did not yield any tetrameric enzyme during the SEC purification. This mutation is located at the dimer interface and has a SASA of 23 %.

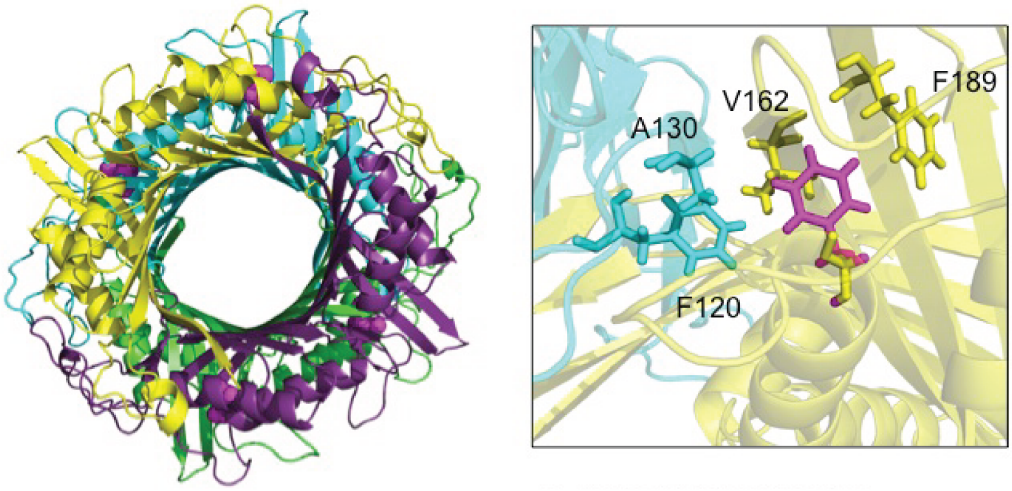


Figure 4.6 Deleterious mutation Y240C arose in the last common ancestor of gorilla, chimpanzee, and human Shown on the left is the global context of this mutation at the dimer interface of two uricase monomer subunits (shown as purple spheres).

The deleterious mutation F222S is located at the tetrameric interface and is buried with a SASA of 7 %, and is surrounded by hydrophobic residues. A tetrameric uricase was purified but was not active. It is possible that this mutation has destabilized the protein complex, and as a result, killed activity.

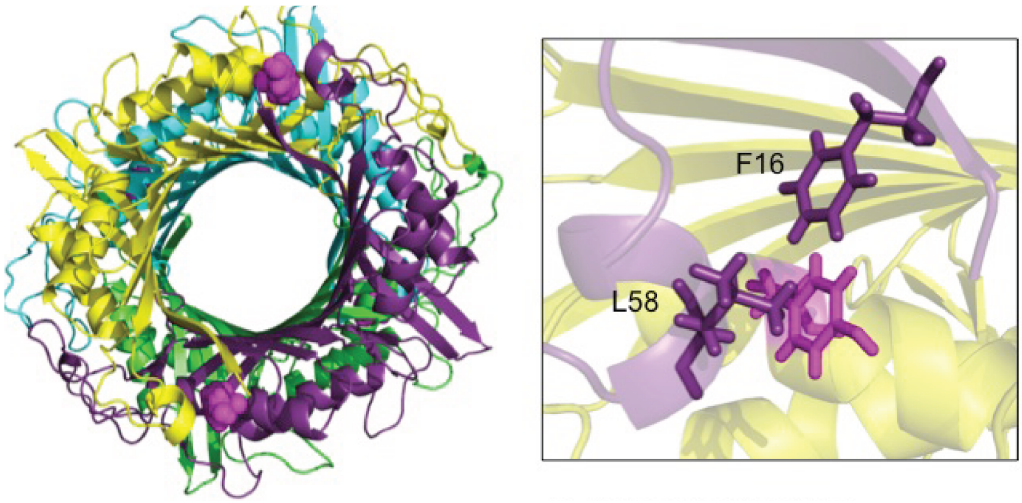


Figure 4.7 Deleterious mutation F222S arose in the common ancestor of hominoids. The An19/22 lysine crystal structure is shown with residue 222 shown as spheres in the purple monomer. This mutation occurs at the tetramer interface, and this residue is surrounded by hydrophobic residues (F120, F189, and V162). The introduction of a polar serine may be disruptive to the assembly of the four

monomers to form the active tetramer. This may explain why this variant could not be extracted or purified.

In contrast to these detrimental mutations, several sites only seemed to exert a mild to even slightly advantageous effect compared to An19/22. The first mutation that seemed to confer a boost in the desired properties of our "human-like" uricase was G83E (a much longer residue and a charged residue) seems likely to impact packing at the tetramer interface (Figure 4.8).

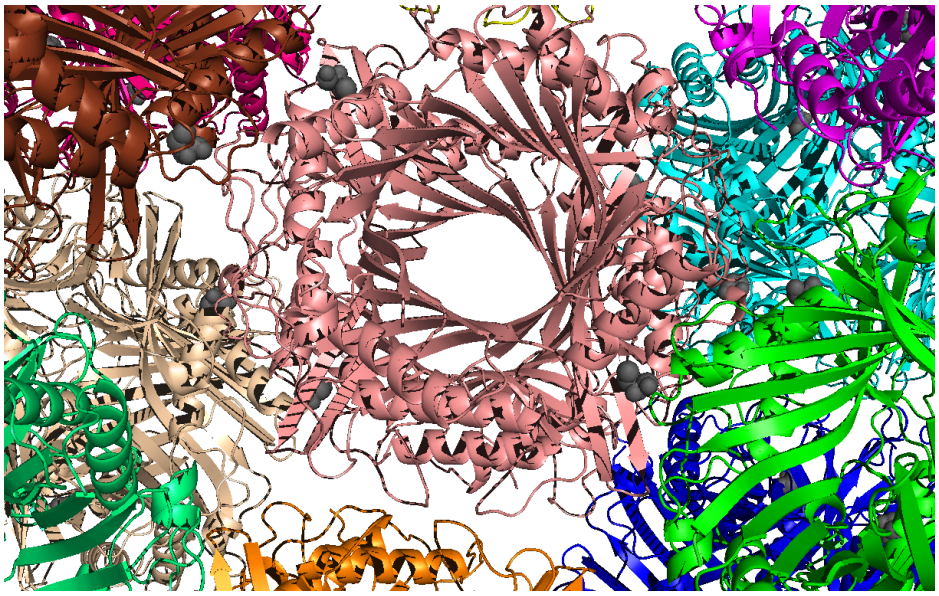


Figure 4.8. The G83E mutation is located at the packing interface of uricase tetramers. The crystal packing of solved variant An19/22-LysII is shown. Each tetrameric uricase is shown as a different color cartoon representation.

4.3.4 Properties of "human-like" uricases

By taking into consideration the effects of individual mutations in the background of An19/22, we developed several chimeric proteins. From the exhaustive single mutant variants, we identified D24E, occurring in the branch from An32 to human as the only mutation that improved the performance of An19/22 (Table 4.3). We therefore elected to include the D24E mutation in all of the chimeric proteins. In our first chimera, An19/22Med, those sites that had a

neutral to slightly negative effect relative to An19/22 were selected: M92V, F120L, and V214I (from An26); D7N (from An27); D24E (from the human sequence). This resulted in a protein that retained 97 % of the SEA of An19/22 while incorporating five additional human residues into the backbone.

Secondly, we introduced a set of sites that seemed to have a neutral to a positive effect upon the properties of An19/22 in a variant referred to as An19/22Plus. This chimeric protein had the following four mutations introduced into the An19/22 backbone: from An26 M146L; from An27 A91G; from An32 G83E; and from the human sequence D24E. This mutation increased the SEA to 120 % of An19/22. An26 contains seven additional human residues than are present in An19/22; at the expense of roughly half the enzyme activity.

Lastly, we wondered whether incorporating our sole "advantageous" mutation D24E into An26 would improve the enzymatic properties of An26. Indeed this mutation resulted in an An26 mutant that retained 80 % of An19/22 specific activity.

Table 4.3. The stable An19/22 and An26 will accommodate select subsets of human residues while retaining its specific enzyme activity. The An19/22Plus variant actually had a higher SEA than its parent An19/22.

Humanized Variants	SEA relative to An19/22	k_{cat} (s^{-1})	K_M (M^{-1} urate)	k_{cat}/K_M ($M^{-1}s^{-1}$)
An19/22Med	97 %	9.4	3.69E-05	2.55E+05
An19/22Plus	120 %	7.8	1.93E-05	4.04E+05
An26Plus	80 %	7.3	2.72E-05	2.69E+05

4.4 CONCLUSIONS

By rigorously querying each mutation that occurred along the tree topology from An19/22 to hUox, the individually detrimental residues were identified. This information is valuable on multiple levels. First, it tells us which residues to avoid as we seek to further humanize An19/22. Secondly, it also tells us something about the evolution of the protein family – in particular, that there were individual mutations that led to a stepwise decrease in uricase activity from An19/22 to An26, to An27, and it's deadening at An30. Of particular importance is the fact that these mutations that diminished uricase activity occurred before the two nonsense mutations silenced the *hUox*. While evaluating only single mutants does not account for allosteric effects or other interactions, it did shed some light upon several residues that were particularly deleterious.

This research walked thorough the point mutations that have occurred during the evolution of the human uricase. By doing this, we identified chimeric ancestral human uricases that still maintain catalytic activity, and may serve as viable therapeutics.

4.5 REFERENCES

114. Weber, C.A., P.J. Mehta, M. Ardito, L. Moise, B. Martin, and A.S. De Groot, *T cell epitope: friend or foe? Immunogenicity of biologics in context.* Adv. Drug Delivery Rev., 2009. **61**(11): p. 965-976.

115. De Groot, A.S. and D.W. Scott, *Immunogenicity of protein therapeutics.* Trends Immunol., 2007. **28**(11): p. 482-490.

116. Sherman, M.R., M.G.P. Saifer, L.D. Williams, M.S. Hershfield, and S.J. Kelly. Aggregate-free urate oxidase for preparation of non-immunogenic polymer conjugates. 7927852 USPTO application. 2011.

117. Zhang, C., K. Fan, X. Ma, and D. Wei, *Impact of large aggregated uricases and PEG diol on accelerated blood clearance of PEGylated canine uricase.* PLoS One, 2012. **7**(6).

118. Fraczkiewicz, R. and W. Braun, *Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules.* J. Comput. Chem., 1997. **19**(3): p. 319-333.

CHAPTER 5: Modifying and testing uricases *in vivo*

Will support the filing of a US patent application:

Kratzer, J. T., Gaucher, E. A. *Use of PEGylated ancestral uricases*. US PATENT APPLICATION. In preparation.

5.1 INTRODUCTION

Polyethylene glycol (PEG) is a hydrophilic functional group $(CH_2CH_2O)_n$ with FDA approval as an additive for food and as a protective group for drugs destined for human consumption because it is largely regarded as being biologically inert [119]. PEGs are long-chain polymers of controlled length. This control is exerted during the synthesis process to enrich for PEG polymers of a desired average molecular weight and are then purified to varying degrees of homogeneity (depending on the size of the chain) [120]. PEGylation, or the attachment of a PEG moiety to a protein, is achieved by incorporating an activating group at one or both ends of the PEG molecule [121]. By utilizing different activating groups and modifying the reaction conditions, a protein chemist can exert control over the number of PEG molecules that attach to the monomeric subunit of a protein, also known as the degree of PEGylation (N) [122].

By coating a protein surface with PEG, several therapeutically valuable characteristics may be conferred to the biologic such as reduced toxicity and immunogenicity, prolonged circulation time, resistance to proteolysis, increased chemical stability and improved solubility [123]. The benefits of enzyme PEGylation can come at the cost of reduced catalytic power. This reduction

may, for example, be attributed to limiting a substrate's access to an enzyme's active site. Therefore, by varying reaction conditions and utilizing activated PEG groups of diverse chemistries, a balance can be struck between improving a proteins drug-like character while retaining a sufficient therapeutic level of activity. Notwithstanding, there are scenarios where a substantial loss in enzymatic activity is perfectly acceptable, provided a therapeutically relevant increase in elimination half-life, a reduction in the clearance rate, and an increase in the absorption half-life[124].

There are examples in the literature that describe the empirical nature of the PEGylation process [125]. It is highly protein-specific because the number and identity of which surface residues are modified can have a grossly different effect on protein activity. Several of the key parameters that direct the reaction are the activating group on the PEG molecule, the reaction buffer pH, and the molar excess (ME) of activated PEG to target protein [126]. The PEGylation reaction is a stochastic process whose heterogeneity is dictated by the length of each PEG, the number of PEGs, and the location of PEGs on each natively folded protein [127].

There are reports that the advantages conveyed by PEGylation are a function of the net mass of PEG that has been attached to the protein rather than the number of attached strands [128]. For the same benefit of improved solubility, the attachment of a single 30 kDa PEG can be replaced by attaching 6 different 5 kDa at six different sites on the protein surface. However, in the later case, the activity is significantly diminished compared to the former [129].

Modification of uricases by the addition of PEG groups is an area of research that has garnered a fair amount of consideration in the scientific literature[130]. A Web of Science (Thomson Reuters) citation search for "PEGylation" and "urate oxidase" or "uricase" returned an average of 50 new publications every year for the last decade. As mentioned earlier, uricase from hog liver, has been a staple in uricase research and the effects of modification by PEG with a molecular weight of 5,000 g/mol shows that it suffers a decrease in specific activity ten-fold with a benefit of doubling the half-life from only 3 h to 6 h [131]. A more beneficial outcome occurred with the modification of *C. utilis* uricase by PEG of a molecular weight of 20,000 g/mol. This modified microbial uricase's SEA more than doubled from 3.3 IU/mg to 8.5 IU/mg (international unit (IU) is defined as the amount uricase required to oxidize 1 μmole of uric acid to allantoin in one minute under the specified assay conditions). Not only did these researchers achieve an increase in the SEA they also increased the enzymatic half-life from 3 hours to 3 days - 24-fold increase [52].

Since we are working with a mammalian uricase that is fairly similar to the PBC uricase (having only 13 amino acid differences between the two) we sought to first screen the conditions PBC researchers report [132]. Namely, what they screened was the use of 5, 10, or 20 kilodalton (kDa) PEG groups attached to uricase. While they used an in-line scattering light detector to develop a universal calibration curve to determine the extent of PEGylation, we relied on a more qualitative measure by following the reaction based on a SEC trace and its motility in a native gel [128].

To improve the drug-like characteristics of PBC uricase (i.e., establishing the requisite stability and tolerance necessary for human administration), these researchers covalently modified the protein via the addition of an average of 10 strands of molecular weight (MW) 10,000 Dalton (Da) PEG groups [129, 133-135]. This work ultimately resulted in the first FDA-approved uricase for the management of TFG due in part to improved half-life and bioavailability achieved by masking the protein through the covalent attachment of PEG) groups on surface-exposed lysines. The PEGylation reduces antigenicity and thus prolonged the circulating half-life of the protein. It is curious that these researchers elected to use a Pig/Baboon chimer since the pig uricase is substantially more active than the baboon uricase. Maybe the key is that this chimera is more 'human-like' compared to pig alone.

5.2 MATERIALS AND METHODS

5.2.1 PEGylation of "human-like" uricases

Three activated PEGs were selected based on the literature of uricase PEGylation [131, 136-138], and purchased from Laysan Bio, Inc . Each of these activated PEGS were used in a series of experiments to modify An19/22 uricase, and the PBC uricase (Laysan Bio Inc.). All of these were methoxy-poly (ethylene-glycol) chains with an average molecular weight of 5,000 g/mol; however, the analysis sheet provided with the activated PEGs (which was determined by gel filtration chromatography by the manufacturer) was used when calculating the mass of activated-PEG required for each reaction. These activated PEGs

covalently attach to the epsilon amine of solvent exposed lysine residues in a non-specific manner.

The solved crystal structure of the An19/22 variant was used as input for the web implementation of the GET AREA solvent accessible area calculator. There are a total of 29 and 30 lysines in An19/22 and PBC, respectively. Three key parameters govern the setup of these PEGylation parameters: the buffer pH selected, the activated PEG group used, and the molar excess of activated-PEG to target protein.

The PEGylation reactions were performed on freshly purified An19/22 and PBC uricase (using methods described in 3.2.3). The concentrations of both uricases were adjusted to 5 mg/mL by concentration using a 30 molecular weight cut-off (MWCO) centrifugal filter (Thermo Scientific). The following calculation was employed to determine the amount of activated PEG for each PEGylation reaction:

$$Activated\ PEG\ (mg) = \left(\frac{M_n}{S_{NMR}}\right) * \left(\frac{P}{M_w * K_{SA} * M_r}\right)$$

- M_n: average molecular weight (g/mol) of activated PEG determined by gel filtration chromatography
- S_{NMR}: degree of substitution (%) of activated PEG determined by nuclear magnetic resonance (NMR)
- P: the mass (mg) of target protein to react
- M_w: the molecular weight (g/mol) of each target protein subunit
- K_{SA}: the number of surface (i.e. solvent accessible lysine residues) in target protein
- M_r: the molar ratio of activated PEG to target protein

The reactions were performed in 2 mL microcentrifuge tubes at a total volume of 1.5 mL in SEC buffer of 0.1 M Na_2CO_3 pH 11. Three activated PEG groups were utilized PEG-NPC, PEG-SG, and PEG-SVA. For An19/22 reactions with the three activated PEG groups were used at five molar excesses to solvent

accessible surface lysines An19/22: 0.5x, 1x, 2x, 3x, or 4x. For PBC uricase the three activated PEG reactions were performed at only three molar excesses relative to solvent accessible surface lysine residues of: 1x, 2x, or 3x (Supplemental Table S1). The reactions were gently mixed by micro-stir bars and incubated for either 1 hour or overnight at 4 °C. The next day, the reaction was quenched by the addition of a large molar excess of 250 mM glycine (pH 7.4), to scavenge any unreacted PEGs from the solution. The reaction was then passed through a size-exclusion column to remove small side products and separate the various degrees of modified uricase.

Both denaturing and native PAGE were employed to follow the PEGylation reactions. Standard 12 % Tris-Glycine gels with 1X SDS-PAGE running buffer were used. For native page, the CAPS/Ammonia buffer system was used as described in the Mini-Protean manual (BioRad).

5.2.2 Pharmacokinetics experiments with unmodified and PEGylated uricases

Sprague Dawley (SD) rates (Charles River Laboratories) were obtained and acclimated after delivery for at least one week before testing. The rats were housed in pairs in a hermetically sealed cage, fed a standard laboratory chow, and given free access to water. Each set of pharmacokinetic experiments was completed during a single 6 – 8 hour period. For each uricase tested, 5 age and weight matched male rats served as biological replicates. The rats were anesthetized using isofluorane and minor surgery was performed to place a catheter made of silicone rubber tubing into the rat's jugular vein, and a

baseline 100 µL blood samples were collected. A sterile 24-gauge tail vein catheter is placed into the tail of the rats in preparation for uricase delivery.

The uricase preparations were kept on ice and then immediately diluted in room temperature PBS, pH 7.4 to a concentration of 0.2 mg/mL. Using a syringe and ensuring no air bubbles are present, the tail catheter was used to inject 1 mL of a 0.2-micron (PALL) filter sterilized uricase preparation. Post-injection blood samples were collected via the jugular catheter at 5 minutes and every 60 minutes for a minimum of 5 hours. Between each sampling, the surgically placed jugular catheter was flushed with a heparin/saline solution to avoid clotting, and a syringe was used to remove a minimum of 200 µL of blood and set aside to be replaced immediately after sample collection. A fresh syringe was then used to remove 100 µL of fresh blood which was immediately transferred to lithium heparin coated Microtainer™ blood collection tubes (BD) and mixed by rocking at room temperature and centrifuged for 5 min at 5,000 rpm to separate plasma from whole blood. Samples were stored on ice until being assayed.). After final collection, all animals are euthanized by the carbon dioxide method.

Uricase remaining in plasma was measured with a commercial fluorometric uricase assay kit (Cell Bio Labs). This kit utilized an optically pure flat-bottom cuvette and the concentration of uricase was determined from a standard curve of a uricase of known activity provided with the kit. One milliunit (mU) is defined as the amount of uricase needed to oxidize 1 nanomole of urate/min under assay conditions.

5.3 RESULTS AND DISCUSSION

5.3.1 The effects of PEGylation on the uricases

The extent of modification via the covalent attachment of PEG was attempted by removing aliquots after the first hour of the reaction, quenching the reaction, and comparing the migration of the band of various reacted samples versus an unreacted sample. First, denaturing SDS-PAGE gels were employed which is a standard approach used in the literature. The expectation was that a heterogeneous mixture containing PEG-uricase would run as a series of bands: the unmodified enzyme has a MW of 34 kDa and the PEGs used had a MW of approximately 5 kDa. There are approximately 15 solvent accessible lysine residues, therefore a denatured samples should run the gambit from 34 kDa (unmodified) to over 100 kDa. However, as will be shown in both the native and denaturing gels presented in this section PEG groups add a much larger surface area to proteins than a globular protein of the same molecular weight. Despite the qualitative nature of gel electrophoresis it is routinely used to assess the "degree" of protein modification by PEG group attachment [139, 140].

A standard denaturing PAGE gel can grossly differentiate between protein samples that have been modified by PEG groups from unmodified proteins (Figure 5.1). Notably, lane 2 contains the quenched 4-hour reaction mixture of An19/22 with a two-fold molar excess of PEG-NPC. While the six bands between 250 kDa and 37 kDa in the molecular weight ladder in lane 1 were resolved, the modified uricase barely migrated within the gel. This same crude sample was separated from the un-reacted protein and excess PEG by SEC into 1X PBS at pH 7.4, and is loaded in lane 4. Also on this gel are the crude reaction

mixtures of An19/22 with PEG-SG or PEG-SVA (lanes 8 and 9, respectively). Lastly, lane 10 contains the unreacted An19/22. So, at least qualitatively, a standard 12 % SDS-PAGE gel can be used to differentiate un-PEGylated from PEGylated uricase based upon the retarded mobility of the modified protein through the gel. The fact that a highly modified uricase can be calculated to be around 100 kDa in but runs at over twice that size on a SDS-PAGE gel is an example of the phenomenon of how PEG groups retard protein migration via molecular sieving due to their large surface area to mass ratio. As a result, when a protein is modified with a large number of relatively small 5 kDa PEG groups, it becomes increasingly difficult to separate proteins with only a few PEG modifications, or heterogeneous mixtures proteins with high degrees of modification [125].

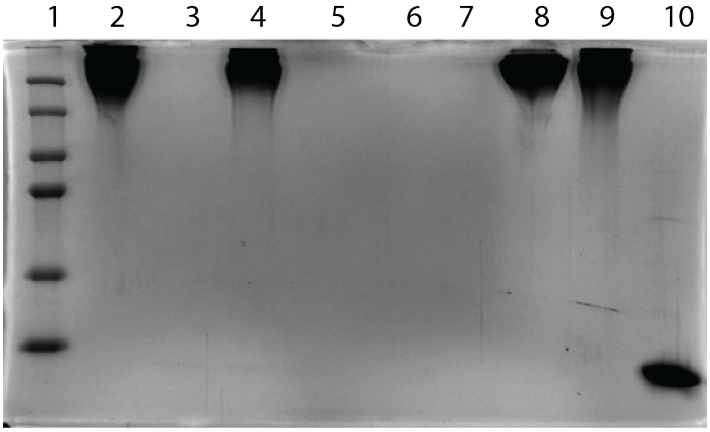


Figure 5.1 12 % SDS-PAGE gel can only differentiate between reacted and unreacted protein samples. The PEGylated products are too large to migrate into this SDS-PAGE gel All reactions are with 2x molar excess activated PEG to protein: Lane 1) 250 kDa protein standards Lane 2) An19/22 + PEG-NPC 4 hr reaction Lane 4) Size-exclusion An19/22 + PEG-NPC Lane 8) An19/22 + PEG-SG 4 hr reaction Lane 9) An19/22 + PEG-SVA 4 hr reaction Lane 10) Unmodified An19/22 control.

By running a standard concentration SDS-PAGE gel it became apparent that the large degree of modification coupled with the relatively low molecular weight PEG group only allowed a qualitative comparison of high to low degrees

of modification. A 6 % SDS-PAGE gel was also attempted but the unmodified uricase control ran off the gel while the PEGylated samples still barely migrated into the gel (not shown).

Based the literature, there are several examples of using native PAGE to obtain a better separation between proteins of various degrees of PEGylation [141]. The theoretical isoelectric point (pI) of An19/22 and PBC uricase was computed using the ExPASy ProtPram tool to be around 8 [142]. A new buffer system had to be selected for this purpose. To this end, a 4 % continuous native PAGE was run using the CAPS-NH_4OH running buffer. This buffer system has a pH of around 11 and this was necessary to ensure the folded uricase would have a net negative charge, since there is no detergent added to ensure a net negative charge on the protein surface in native gel electrophoresis. Samples from several different PEG-uricase reactions that differ in both the molar excess of activated PEG to protein present in the reaction, as well as, the activated PEG employed were determined (Figure 5.2). Each pair of lanes from 2 – 7 are loaded with the An19/22 PEGylated uricase while the PBC PEGylated uricase is in the next lane to the right. Lanes 2 and 3 contain the PEG-NPC reactions. Lanes 4 and 5 contain the PEG-SG reaction, and the final two lanes contain PEG-SVA reactions in lanes 6 and 7. The PEG-SG reactions ran as a much broader smear than the other uricase-PEG reactions and suggests that under the reaction conditions PEG-SG covalently modifies these uricases to a lesser extent than its activated PEG counterparts (PEG-NPC or PEG-SVA). The samples in lanes 9 and 10 contain unmodified tetrameric uricase and also SEC separated octomeric

uricase thereby confirming that this native buffer system does in fact effectively separate native uricase.

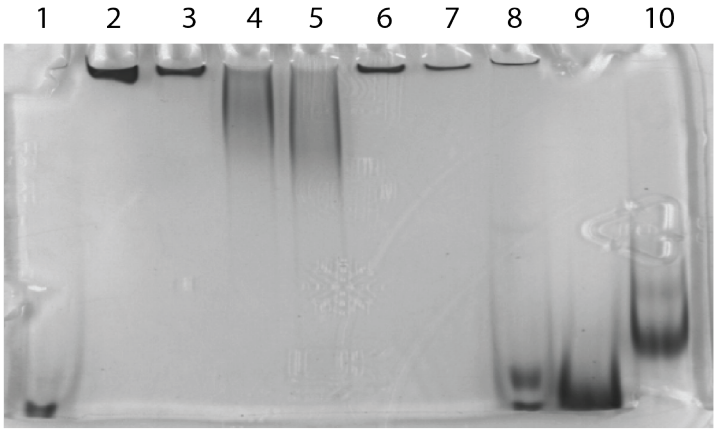


Figure 5.2 A 4 % Native PAGE gel run with CAPS-NH_4OH running buffer shows a better separation between PEG-protein species. Lane 1)Native protein standard of 158 kDa Lane 2) An19/22 + PEG-NPC 4 hour reaction Lane 3) PBC + PEG-NPC 4 hour reaction PBC Lane 4) An19/22 + PEG-SG 4 hour reaction Lane 5) PBC + PEG-SG 4 hour reaction Lane 6) An19/22 + PEG-SVA 4 hour reaction Lane 7) PBC + PEG-SVA 4 hour reaction Lane 8) Native protein standard of 158 kDa Lane 9) Unmodified An19/22 Lane 10) An19/22 unmodified octomer peak from SEC.

Perhaps the experimental conditions of employing a molar excess of activated PEG dictated that all of the reactions went to completion within a short amount of time, and the separation techniques employed did not allow us to differentiate between a heterogeneous mixture of highly modified uricase tetramers. The following set of gels look at this scenario by evaluating at both a range of molar excess of activated PEGs as well as comparing a single hour reaction with one that was allowed to progress for much longer before being quenched (Figure 5.3).

This first gel shows a series of An19/22 (lanes 2-5) and PBC (lanes 7-10) uricase reactions with NPC-PEG molar excess from left to right of 3x, 2x, 1x, and 0.5x. Separating these two sets of reactions is the unmodified An19/22 uricase at 5 mg/mL – the same concentration at which the PEGylation reactions were

performed. It is apparent that a 2-fold molar excess is sufficient to get the maximum covalent modification of either uricase since using the samples reacted with a 3-fold molar excess ran almost identically (comparing lanes 2 and lanes 3). As the molar excess was decreased gradually to 0.5-fold, a small amount of modified uricase was now visible on the gel.

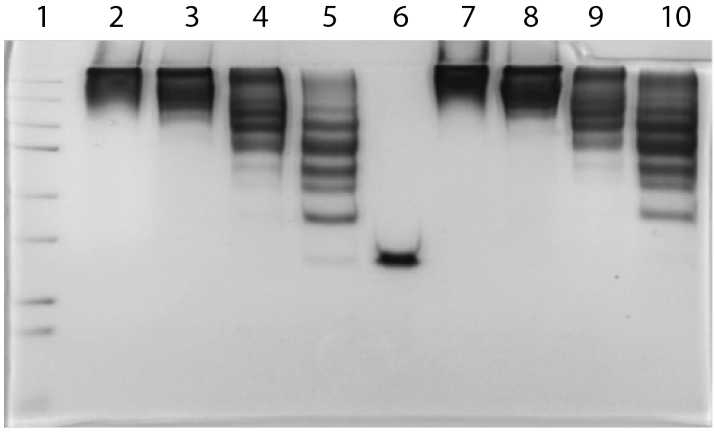


Figure 5.3 Extent of PEGylation can be modulated by altering the molar excesses of activated PEG relative to the protein's surface lysines available for conjugation. This SDS-PAGE gel (4 % Stack and 6 % resolve) looks at varying the molar excess of activated PEG and the reaction is quenched after four hours. Lane 1) 250 kDa protein ladder Lane 2) 3 x molar excess (ME) PEG-NPC + An19/22 Lane 3) 2 x ME PEG-NPC + An19/22 Lane 4) 1 x ME PEG-NPC + An19/22 Lane 5) 0.5 x ME PEG-NPC Lane 6) Unmodified An19/22 Lane 7) 3 x ME PEG-NPC + PBC Lane 8) 2 x ME PEG-NPC + PBC Lane 8) 2 x ME PEG-NPC +PBC Lane 9) 1 x ME PEG-NPC + PBC Lane 10) 0.5 x ME PEG-NPC + PBC.

Figure 5.4 examines the effect of incubation length on the extent of PEGylation as visualized by denaturing gel electrophoresis. In this case each pair of lanes from left to right are for molar excesses of PEG-SVA reacted with An19/22 of 3x ME, 2x ME, 1x ME, and 0.5x ME. Within each set of lanes, the left lane is half that reaction removed after 1 hour and quenched, and the right lane of that pair is the remaining reaction quenched after a total of 19 hours of incubation. For all molar excesses of PEG-NPC, PEG-SG, and PEG-SVA reacted

conducted with either An19/22 uricase or PBC uricase the reaction was complete after only an hour of incubation.

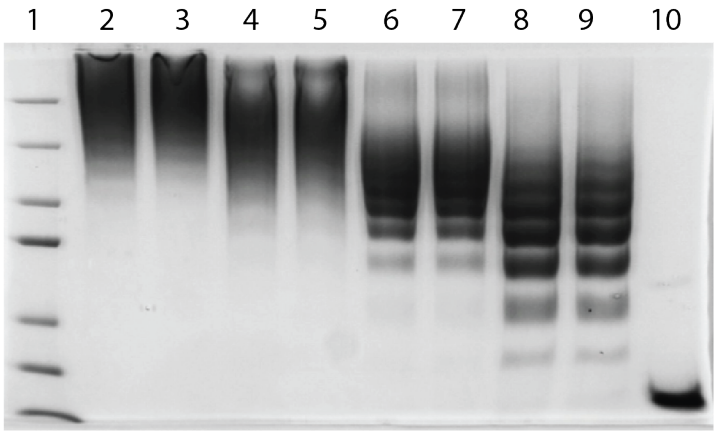


Figure 5.4 The PEGylation reactions reach completion after a single hour. This SDS-PAGE gel (4 % Stack and 6 % resolve) looks at varying the effect of incubation time prior to reaction quenching on PEGylation extent of An19/22 uricase. Lane 1) 250 kDa protein ladder Lane 2) 3 x molar excess (ME) PEG-NPC + An19/22 incubated for 1 h Lane 3) 3 x molar excess (ME) PEG-NPC + An19/22 incubated for 19 h Lane 4) 2 x molar excess (ME) PEG-NPC + An19/22 incubated for 1 h Lane 5) 2 x molar excess (ME) PEG-NPC + An19/22 incubated for 19 h Lane 6) 1 x molar excess (ME) PEG-NPC + An19/22 incubated for 1 h Lane 7) 1 x molar excess (ME) PEG-NPC + An19/22 incubated for 19 h Lane 8) 0.5 x molar excess (ME) PEG-NPC + An19/22 incubated for 1 h Lane 9) 0.5 x molar excess (ME) PEG-NPC + An19/22 incubated for 19 h Lane 10) Unmodified An19/22

The next set of experiments were performed to get a sense for the effects that the three activated PEG groups have on the solubility and activity profile of An19/22 or PBC. Each uricase was exchanged into a physiological buffer of 1x PBS, pH 7.4 for 1 hour, and after dialysis was complete the samples were collected and the protein concentration was determined by a Bradford assay. Finally, the initial velocity of the enzymatic reaction was measured to determine whether or not PEGylation had improved the stability and activity of uricases incubated at pH 7.4 (Figure 5.5). For both An19/22 and PBC, the catalytic activity

did improve upon PEGylation by either NPC-PEG or SVA-PEG relative to the unmodified uricase.

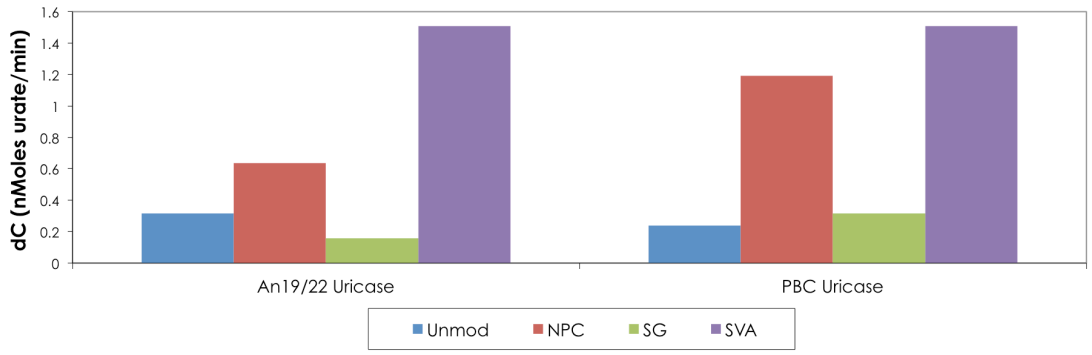


Figure 5.5 PEGylation improves uricase activity at sub-optimal pH. This quick assay reveals the enhancement that PEGylation has upon the pH stability. All samples have been dialyzed against 1X PBS (pH 7.4) and were assayed at 0.3 mg/mL. Shown in the chart are the unmodified uricases (Unmod) in blue, the uricases modified with the nitrophenylcarbonate (NPC) PEGs in red, uricases modified with succimidylgluterate (SG) PEGs in green, and uricases modified with succimidylvalerate (SV) PEGs in purple.

To ultimately decide which PEGylation scheme would be pursued for the first animal studies, we decided to compare the specific activity of each 1 hour PEGylation reaction that had been quenched by an excess of a neutral pH glycine stock solution. Both An19/22 and PBC were reacted with molar excesses of activated PEG of 0.5x, 1x, and 2x. Since a greater yield was obtained from the preparation of An19/22 additional PEGylation reactions using molar excesses of 0.25x and 4x were also explored. The crude reaction mixtures were first assayed to determine their specific activities via enzymatic assays at 100 μM urate performed in triplicate (Figure 5.6). This initial set of assays exhibited the trend that, at all molar excesses, explored reactions with PEG-NPC retained the highest extent of specific activity relative to unmodified uricase. This suggested that a 1x

ME of PEG-NPC was optimal for An19/22 whereas a 3x ME of NPC-PEG maintained the highest residual activity when reacted with PBC.

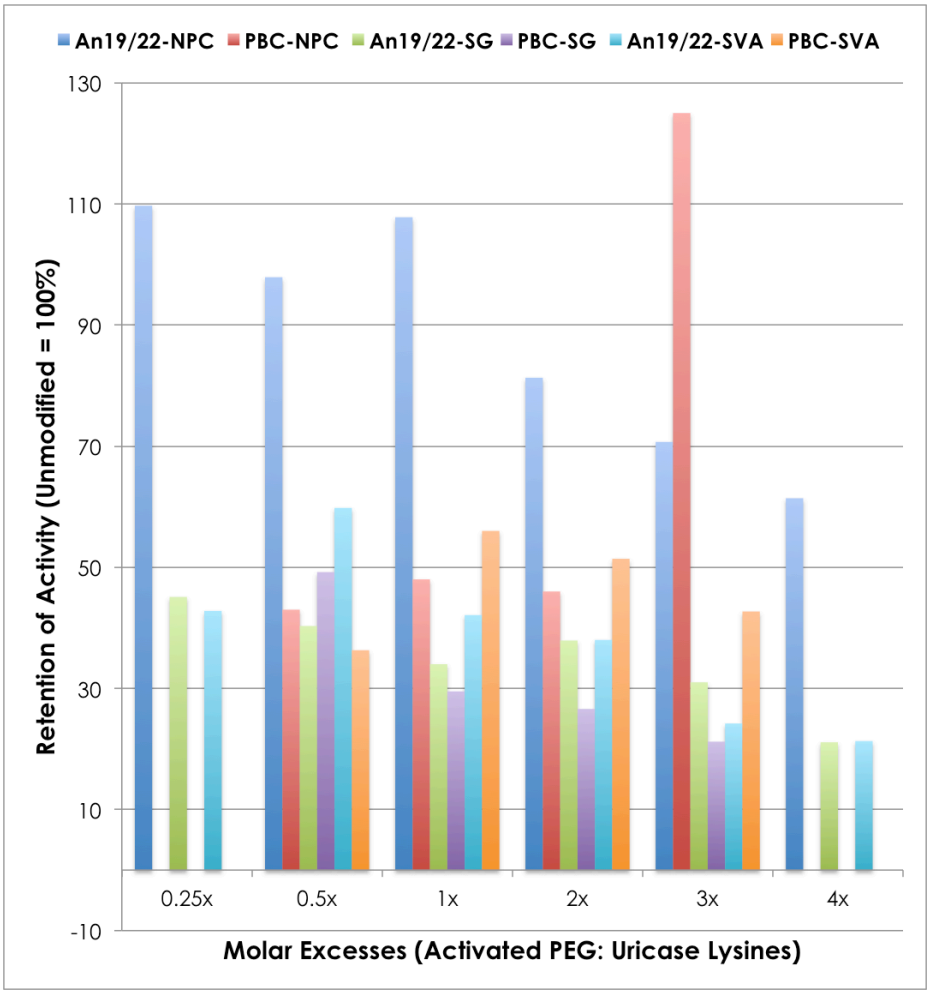


Figure 5.6. Crude PEGylation reaction mix in 0.1 M carbonate buffer pH 11 Assayed at 100 µM urate; 1X PBS pH 7.4 in triplicate. The retention of activity of both An19/22 and PBC uricase was greatest when the NPC-PEG was employed compared to SG-PEG or SVA-PEG. Note: due to unequal yields in starting protein material (reactions for 0.5x and 3x molar excesses of activated PEG were only performed with An19/22).

The next is that the separation of the PEGylated uricases from unreacted protein and PEG groups, minimally modified proteins, or even large protein/PEG aggregates - that can trigger an immune response themselves [116, 143]. This is achieved by taking the crude reaction mixtures whose assays are shown above

and performing SEC chromatography under physiological conditions of 1X PBS, pH 7.4 (Figure 5.7).

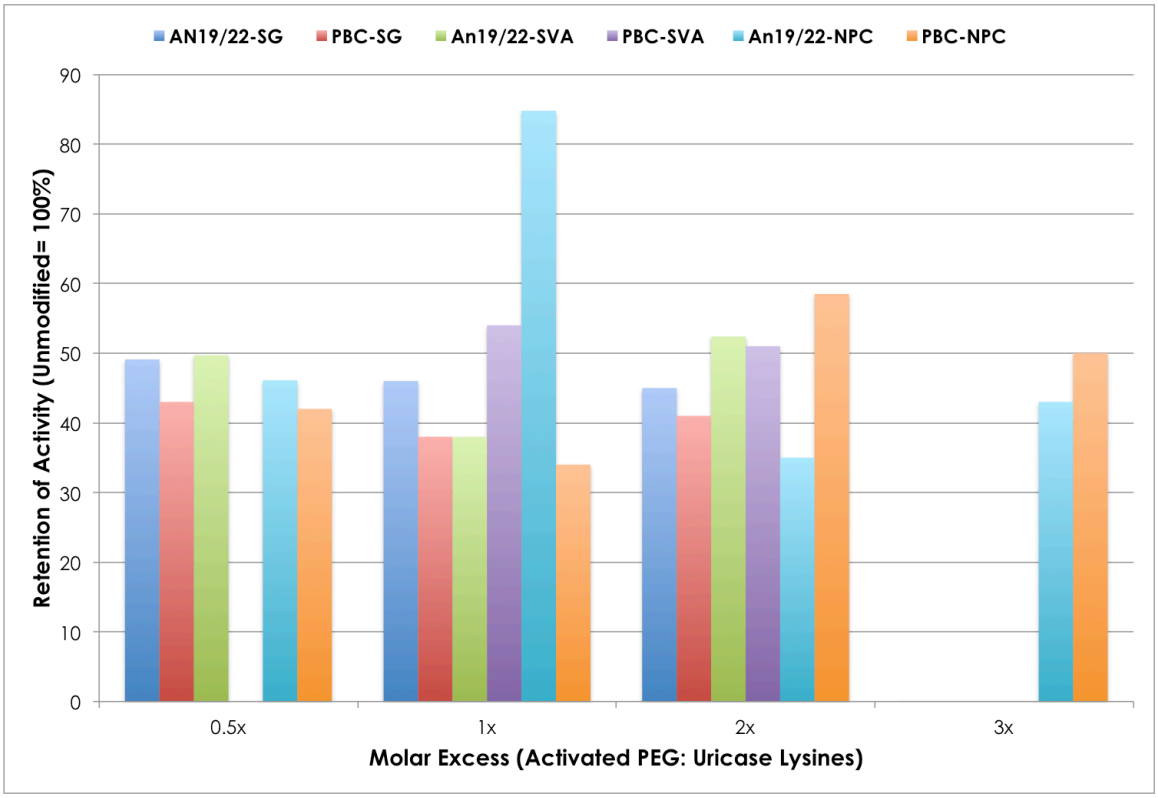


Figure 5.7 Various PEGylation strategies effect on the SEA An19/22 uricase at physiological pH Each quenched PEGylation was purified into pH 7.4 1X PBS. Based upon the SEA activity assays of the crude reaction mixtures several of the reaction mixtures were not purified into 1X PBS pH 7.4 and do not appear in the above graph (e.g. An19/22-SG at 3x ME). Note: due to unequal yields in starting protein material (reactions for 0.5x and 3x molar excesses of activated PEG were only performed with An19/22).

5.3.2 Pharmacokinetic data of unmodified and PEGylated uricases

The goal of this first *in vivo* testing of uricases is to obtain simple pharmacokinetic data regarding the safety and stability of the various uricases. The animal model selected for these experiments was the rat since it is commonly used for such early studies. These animals were each injected with equal amounts of An19/22 or PBC uricase and the stability of these enzymes was

determined by assaying the uricase activity present in blood samples taken over the course of 5 hours (Figure 5.8).

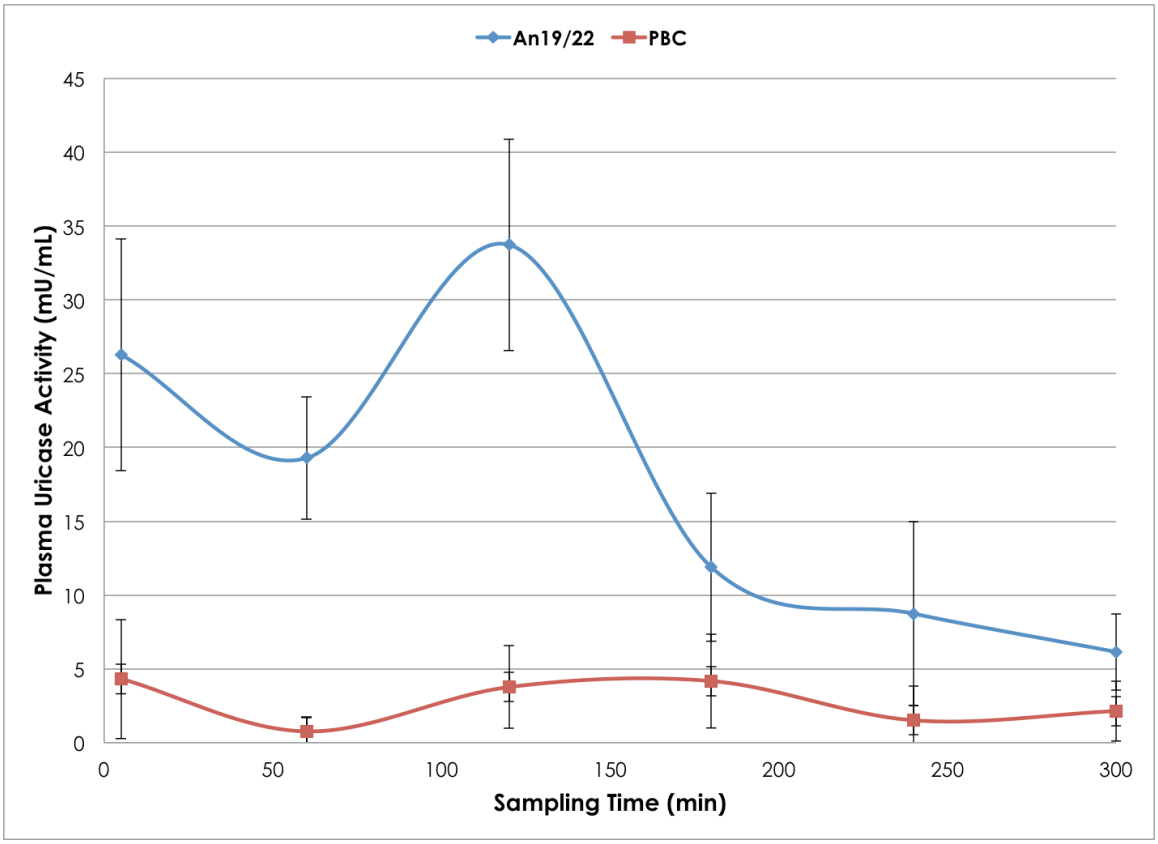


Figure 5.8. In healthy rats unmodified An19/22 displayed enhanced pharmacokinetics over the PBC uricase. 10 male SD rats (5 rats for each uricase) were each injected with 1 mL (0.2 mg/mL) of recombinant uricase preparations (An19/22, or PBC the basis for Krystexxa©). One milliunit (mU) is the amount of uricase needed to oxidize 1 nanomole of urate/min under assay conditions. The An19/22 uricase had a statistically improved half-life compared to PBC uricase. A student T-test was performed with a p-value of 0.02.

The An19/22 uricase shows a higher activity than PBC uricase in plasma, of healthy rats, at all samples taken post injection. A student T-test was performed with a p-value of 0.02. The following assumptions were employed in the calculation of the plasma uricase activity estimation: The average weight of the rats used was 284 g in the An19/22 set of replicates and 272 g in the PBC replicates. Using the reported value for the Norwegian rat of 3.38 mL plasma/100

g gives a plasma volume of 9.6 and 9.2 mL for An19/22 and PBC, respectively. The specific enzyme activities were determined before dosing to be 1.4 for An19/22 and 0.9 for PBC. A total of 0.230 mg of each uricase was administered which comes out to a hypothetical initial concentration of 34 mU/mL for An19/22, and 24 mU/mL for PBC given the above assumptions.

Based upon our i*n vitro* work presented earlier (Figure 5.7) An19/22 reacted with a 1x ME of PEG-NPC retained the greatest percentage of the unmodified enzyme's specific activity. Therefore, we took both An19/22 and PBC modified with 1xME PEG NPC and injected it into rats to ascertain the effects PEGylation had upon these uricases' *in vivo* stabilities (Figure 5.9). Unfortunately, the uricase standards that were run during these assays had become degraded so was not possible to report these assays in terms of uricase activity. However, by comparing the assay's raw output, relative fluorescent units (RFUs) plasma samples from rats injected with An19/22 are reporting a higher signal than those collected from animals injected with PBC uricase. Furthermore, the residual uricase activity in An19/22 injected rats seems to be decreasing at a slower rate than PBC uricase injected rats.

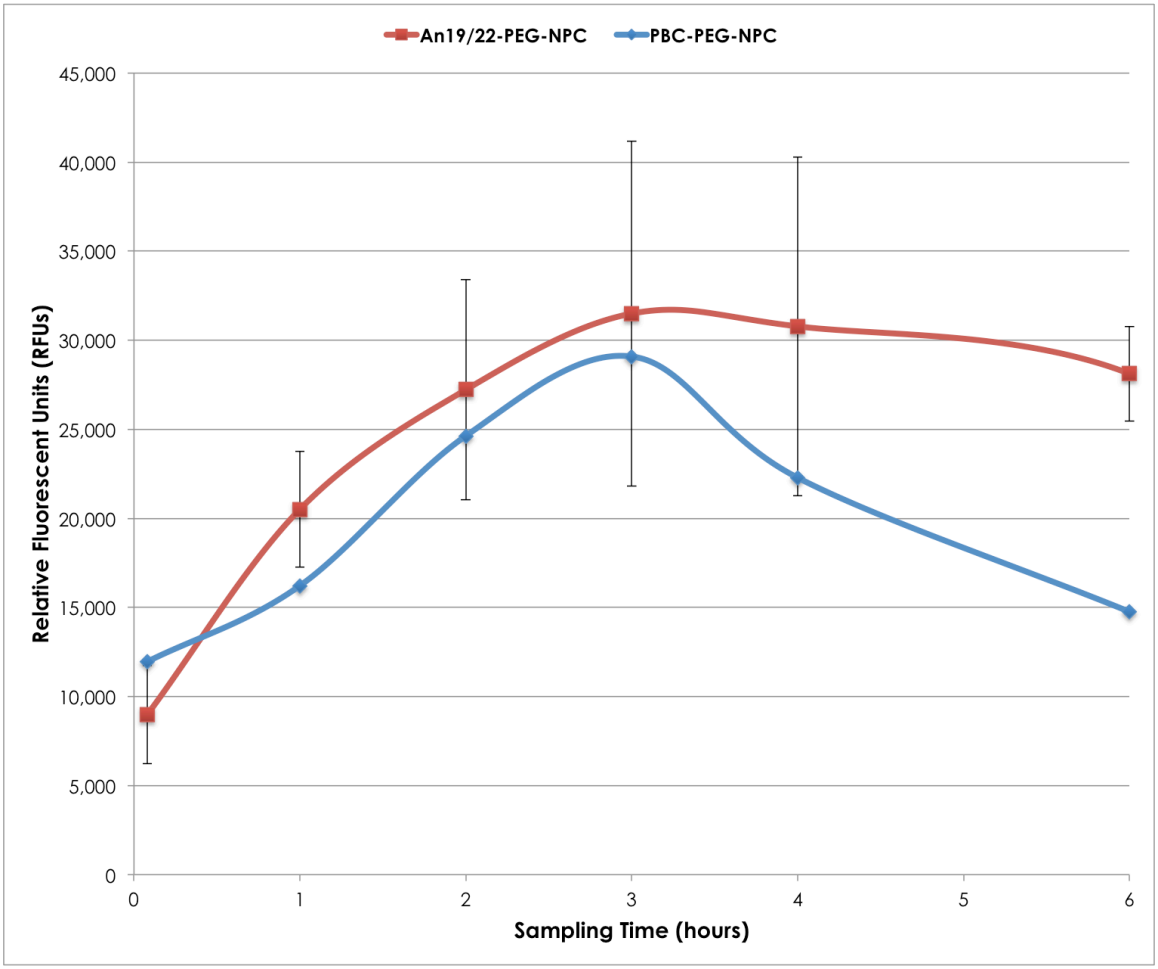


Figure 5.9. PEGylating an ancestral uricase improves its circulation time in healthy rats Biological duplicates were each dosed with PEG-NPC modified An19/22 or PBC. The raw fluorescent signal (RFU) is reported because the uricase standards did not generate a well-fit standard curve. A student T-Test (2-tailed, heteroscedastic = two-samples of unequal variance) was performed and the error bars are the standard deviations among samples ($p = 0.24$*).

These experiments need to be repeated with a larger sample set and as the PEGylation scheme/purification is better optimized, there will be a need to follow the reactions for a greater duration of time. This obviously will require that the animals not be anesthetized for the entire period – since a previous run for a longer time period of 10 hours led to irregularities with the laboratory animals and difficulty in interpreting the results. As cited in the literature, PEGylation has been reported to dramatically increase the half-life of PBC uricase [144] . Therefore, it

will be necessary to allow the animals to recover so that a longer duration of sampling may be conducted.

5.4 CONCLUSIONS

These experiments demonstrate that ancestral uricase An19/22 appears to be more stable in rat regardless of whether it is PEGylated or not. The non-PEGylated An19/22 is statistically more stable after injection into rats than the non-PEGylated PBC protein. Although the PEGylated uricase tests suggest that An19/22 retains its longer half-life in rats, these experiments will require follow-up tests in order to determine the robustness of the results. The homogeneity of the PEGylation reactions that were performed by covalent modification of both An19/22 and PBC uricase were not determined. All that could really be said is that, even after 1 hour the proteins were fully modified based on PAGE gel electrophoresis. While the FDA does not require that all uricases be covalently modified with PEG uniformly they do require reproducibility in the batch-to-batch range of this modification, usually expressed as the average number of covalently attached PEG moieties per protein subunit [145]. Therefore, it will be necessary for further refinement to explore other more analytical methods such as tandem mass spec to characterize these reactions.

Furthermore, the small scale (in terms of volume) at which these reactions were performed will likely not translate well to larger scale preparations. Therefore, other separation schemes such as diafilitration/ultrafiltration should be explored since SEC is both time-consuming and there is a risk that PEG will foul the column.

In these experiments, we looked at three attachment chemistries; NPC, SG, and SVA, all at a constant molecular weight. The next step is to look at using additional molecular weights. While historically PEG has been deemed "biologically inert", there are reports that antibodies against the PEG moiety itself have resulted in rapid clearance of PEG-biotherapeutics [146]. Furthermore, a recent publication by the researchers that developed the PEGylation technology employed in Krystexxa® report that the antibodies resulting in the accelerated clearance of mPEG protein conjugates recognize the methoxy group of the PEG moiety, and they suggest that using a hydroxy-PEG (HO-PEG) conjugate may prevent may prevent this unwanted consequence [147] [148]. Excitingly, this initial work suggests that An19/22 is more active in rat blood than PBC over the 5 hours sampled.

5.5 REFERENCES

52. Bomalaski, J.S., F.W. Holtsberg, C.M. Ensor, and M.A. Clark, *Uricase formulated with polyethylene glycol (uricase-PEG 20): Biochemical rationale and preclinical studies.* J. Rheumatol., 2002. **29**(9): p. 1942-1949.

116. Sherman, M.R., M.G.P. Saifer, L.D. Williams, M.S. Hershfield, and S.J. Kelly. Aggregate-free urate oxidase for preparation of non-immunogenic polymer conjugates. 7927852 USPTO application. 2011.

119. Zalipsky, S. and J.M. Harris. *Introduction to chemistry and biological applications of poly(ethylene glycol)*. in *Poly(ethylene glycol) Chemistry and Biological Applications: ACS Symposium Series 680*. 1997. Washington, DC: American Chemical Society.

120. Roberts, M.J., M.D. Bentley, and J.M. Harris, *Chemistry for peptide and protein PEGylation.* Adv. Drug Delivery Rev., 2002. **54**(4): p. 459-476.

121. Sherman, M.R., D.W. Williams, M.G. Saifer, J.A. French, L.W. Kwak, and J.J. Oppenheim. *Conjugation of high-molecular weight poly(ethylene glycol) to cytokines: granulocyte-macrophage colony-stimulating factors as model substrates*. in *Poly(ethylene glycol) Chemistry and Biological Applications*. 1997. San Francisco, CA: American Chemical Society.

122. Conan, J.F. and M.V.A. James, *PEG-proteins: Reaction engineering and separation issues.* Chem. Eng. Sci., 2006. **61**.

123. Veronese, F.M. and G. Pasut, *PEGylation, successful approach to drug delivery.* Drug Discov. Today, 2005. **10**(21): p. 1451-1458.

124. Veronese, F.M. and A. Mero, *The impact of PEGylation on biological therapies.* Biodrugs, 2008. **22**(5): p. 315-329.

125. Fee, C.J. and J.A. Van Alstine, *PEG-proteins: Reaction engineering and separation issues.* Chem. Eng. Sci., 2006. **61**(3): p. 924-939.

126. Sezer, A.D. and A.F. Yagci, *Overview of peptide and protein PEGylation: properties and general strategies.* Acta Pharm. (Zagreb, Croatia), 2010. **52**: p. 377-389.

127. Kunitani, M., G. Dollinger, D. Johnson, and L. Kresin, *On-line characterization of polyethylene glycol-modified proteins* J. Chromatogr., 1991(588): p. 125-137.

128. Fee, C.J. and J.M. Van Alstine, *Prediction of the viscosity radius and the size exclusion chromatography behavior of PEGylated proteins.* Bioconjugate Chem., 2004. **15**(6): p. 1304-1313.

129. Williams, D.L., M.S. Hershfield, S.J. Kelly, M.G.P. Saifer, and M.R. Sherman. PEG-urate oxidase conjugates and use thereof. 8067553 USPTO application. 2011.

130. Pasut, G. and F. Veronese, *State of the art in PEGylation: the great versatility achieved after forty years of research.* J. Controlled Release, 2012. **161**(2): p. 461-472.

131. Chen, R.H., A. Abuchowski, T. Van Es, N.C. Palczuk, and F.F. Davis, *Properties of two urate oxidases modified by the covalent attachment of poly(ethylene glycol)*. Biochim. Biophys. Acta, 1981. **660**(2): p. 293-298.

132. Hartman, J. and S. Mendelovitz. Variant forms of urate oxidase and use thereof. 8188224 2012.

133. Williams, D.L., M.S. Hershfield, S.J. Kelly, M.G.P. Saifer, and M.R. Sherman. PEG-urate oxidase conjugates and use thereof. 6576235 2003.

134. Sherman, M.R., M.G.P. Saifer, D.L. Williams, M.S. Hershfield, and S.J. Kelly. Aggregate-free urate oxidase for preparation of non-immunogenic polymer conjugates. 7927852 2011.

135. Hartman, J.H., IL), Mendelovitz, Simona (Ramat Aviv, IL, US), Rehrig, Claudia D. (Plainsboro, NJ, US), Huang, William (Florham Park, NJ, US), Hershfield, Michael (Durham, NC, US). Methods for lowering elevated uric acid levels using intravenous injections of PEG-uricase. 8148123 2012.

136. Bomalaski, J.S., D.H. Goddard, D. Grezlak, M.A. Lopatin, F.W. Holtsberg, C.M. Ensor, and M.A. Clark, *Phase I study of uricase formulated with polyethylene glycol 3 (Uricase-PEG 20)*. Arthritis Rheum., 2002. **46**(9): p. S141-S141.

137. Zhang, C., K. Fan, H. Luo, X. Ma, R. Liu, L. Yang, C. Hu, Z. Chen, Z. Min, and D. Wei, *Characterization, efficacy, pharmacokinetics, and biodistribution of 5kDa mPEG modified tetrameric canine uricase variant.* Int J Pharm, 2012. **430**(1-2): p. 307-317.

138. Zhang, C., K. Fan, W. Zhang, R. Zhu, L. Zhang, and D. Wei, *Structure-based characterization of canine-human chimeric uricases and its evolutionary implications.* Biochimie, 2012. **94**(6): p. 1412-1420.

139. Morpurgo, M. and F.M. Veronese, *Conjugates of peptides and proteins to polyethylene glycols*, in *Bioconjugation Protocols: Strategies and Methods*, C.M. Niemeyer, Editor. 2004, Humana Press, Inc: Totowa, NJ. p. 45-70.

140. Mero, A., C. Clementi, F.M. Veronese, and G. Pasut, *Covalent conjugation of poly(ehtylene glycol) to proteins and peptides: Strategies and methods*, in *Bioconjugation protocols: Strategies and Meethods*, S.S. Mark, Editor. 2011, Springer Science and Business Media. p. 95-129.

141. Zheng, C., C. Zheng, G. Ma, and Z. Su, *Native PAGE eliminates the problem of PEG-SDS interaction in SDS-PAGE and provides an alternative to HPLC in characterization of protein PEGylation.* Electrophoresis, 2007. **28**(16): p. 2801-2807.

142. Artimo, P., M. Jonnalagedda, K. Arnold, D. Baratin, G. Csardi, E. de Castro, S. Duvaud, V. Flegel, A. Fortier, E. Gasteiger, et al., *ExPASy: SIB bioinformatics resource portal.* Nucleic Acids Res., 2012. **40**: p. 603.

143. Zhang, C., K. Fan, X.F. Ma, and D.Z. Wei, *Impact of Large Aggregated Uricases and PEG Diol on Accelerated Blood Clearance of PEGylated Canine Uricase.* PLoS One, 2012. **7**(6): p. 11.

144. Sherman, M.R., M.G. Saifer, and F. Perez-Ruiz, *PEG-uricase in the management of treatment-resistant gout and hyperuricemia.* Adv. Drug Delivery Rev., 2008. **60**(1): p. 59-68.

145. Alconcel, S.N.S., A.S. Baas, and H.D. Maynard, *FDA-approved poly(ethylene glycol)-protein conjugate drugs.* Polym. Chem., 2011. **2**: p. 1442-1448.

146. Armstrong, J.K., G. Hempel, S. Koling, L.S. Chan, T. Fisher, H.J. Meiselman, and G. Garratty, *Antibody against poly(ethylene glycol) adversely affects PEG-asparaginase therapy in acute lymphoblastic leukemia patients.* Cancer, 2007. **110**(1): p. 103-111.

147. Garay, R.P., R. El-Gewely, J.K. Armstrong, G. Garratty, and P. Richette, *Antibodies against polyethylene glycol in healthy subjects and in patients treated with PEG-conjugated agents.* Expert Opin. Drug Deliv., 2012. **9**(11): p. 1319-1323.

148. Sherman, M.R., L.D. Williams, M.A. Sobczyk, S.J. Michaels, and M.G. Saifer, *Role of the methoxy group in immune responses to mPEG-protein conjugates.* Bioconjug. Chem., 2012. **23**(3): p. 485-499.

CHAPTER 6: FUTURE DIRECTIONS AND CONCLUDING REMARKS

6.1 FUTURE DIRECTIONS

6.1.1 Going after the expressed human "pseudogene"

Based upon our HEK-293T experiments, we believe that human cells may naturally express a uricase gene product. However, the translation process is encumbered by two premature stop codons lying in the way of full-length uricase. Therefore, we would require an antibody in order to detect the small amount of human uricase that might be expressed in human cells naturally. In our early work with HEK-293T cells, we tried using a polyclonal Ab (pAb) from rabbits raised against rat uricase. This polyclonal antibody gave a very nonspecific signal to a large number of proteins on the membrane. The noise of this signal abrogated the sensitivity required to detect if small amounts of uricase are present and in are earlier studies we turned to the much more selective Anti-FLAG and Anti-His antibodies. A natural extension therefore would be to raise our own anti-hUox antibody so that we can detect, or pull-down, human uricase. To this end, we have already shown that, under denaturing conditions, cation exchange chromatography can be employed to isolate full-length human uricase recombinantly expressed in *E. coli.*

6.1.2 Optimization of PEGylation strategy

We have only scratched the surface using conventional non-specific chemistries with three linear active PEGs of molecular weight 5,000 g/mol. Chemistry that is more specific can be explored in the future. Also, additional

sized PEG moieties also need also to be explored in hopes to better evade a human immune response.

Furthermore, there are reports about antibodies against the PEG moiety itself. These antibodies compromise the active protein component by essentially promoting recognition by the body and enhancing clearance from the body (the exact opposite of their designed intent). Another area that would have been beneficial is to develop an in-house expertise for the ability to determine the average molecular weight, and by extension, the degree of PEG-modification (N) of the different uricases.

For these small-scale preparations, we relied heavily on preparative size-exclusion. It may be worthwhile to perform analytical grade size-exclusion in conjunction with an in-line light scattering detector. Methods are available to generate two calibration curves in the same experimental system (i.e., same sizing column and buffer conditions) whereby these two calibration curves can then be combined to generate a universal calibration curve[128]. Once such a calibration curve is generated, the sample of interest can be run and the amount of PEG can be more quantitatively determined. Furthermore, we have only performed the most preliminary studies in terms of the number of biological replicates, as well as the duration of plasma sampling. Since PEGylation has been shown to dramatically increase the half-life of other uricases, it will be important to alter our PK studies to accommodate a longer monitoring/sampling period.

6.1.3 Testing therapeutic effect

We have already begun using rats as a healthy model to study the *in vivo* stability of our uricase preparations, rodents are commonly used in the field for such testing [131, 149-151]. The only true disease gout-like model is a strain of mice having a double uricase knockout [110]. Mice usually have an active uricase so the uricase knockout mouse's diet must be supplemented with allopurinol, otherwise it will not live to adulthood (The Jackson Laboratory, Bar Harbor, ME). The FDA approval process for new drugs, including a new uricase protein to treat gout and manage hyperuricemia, requires first in human testing in the form of a Phase I clinical trial. For gout, it is possible to conduct clinical studies to get some preliminary efficacy data from first in human studies in addition to safety data.

6.1.4 Determining the safety of uricases

While we have tested our most ancient uricase An19/22 in healthy rats, it is a poor proxy for the human – especially considering that we are interested in avoiding a uniquely human immune response. Therefore, it will be invaluable to determine if approaching the human uricase protein sequence has imparted any reduction in the immunogenicity over An19/22 versus the FDA approved PBC uricase. There are commercially available cell-based Dendritic/T-cell activation assays from a representative donor population, epitope mapping and T-cell activation assays (ProImmune). These assays would allow for the direct comparison between the humanized ancestors and the PBC uricase of the black box prescribed Krystexxa®.

6.1.5 Explore the co-administration of uricase's downstream enzymes

Those organisms with functional uricases also have the downstream enzymes, HIU hydrolase and OHCU decarboxylase to catalyze the stereospecific conversion of urate to *S*-allantoin. While HIU decomposes non-enzymatically into racemic allantoin, there is evidence that the accumulation of HIU may be potentially harmful [152]. Therefore, it may be worthwhile to explore the possibility of concomitantly administering these hydrolases and decarboxylases to rapidly generate *S*-allantoin, and avoiding the accumulation of HIU.

6.2 CONCLUDING REMARKS

Our first major endeavor was to explore the transcriptional and translational features of a human uricase pseudogene. To this end, we queried a fetal human cDNA library and obtained a product that corresponds to human uricase minus exon 3. The translational part of this study was addressed by transiently transfecting Human Embryonic Kidney cells with constructs containing the two premature stop codons. Using western blot, we confirmed that, despite the two premature stop codons, the human uricase pseudogene can be translated as a full-length product.

Another key aim of this research was to develop a "human-like" uricase to exploit the possibility that a functional uricase with a high sequence identity to the human pseudogene would be recognized as self and would not elicit an immune response. We initially attempted to recombinantly express in *E. coli* the human uricase where the two stop codons had been mutated to the arginines found in functional orthologs. However, this protein could not be solubilized and

there was no detectable uricase activity within the insoluble fraction. Since we could not directly reactivate the human uricase, we applied Ancestral Sequence Reconstruction to identify the last functional ancestral uricase whose descendent path leads to the human lineage. We then were able to understand the effects individual mutations had on a functional ancestral mammalian uricase (An19/22) as we traversed the evolutionary path to the human sequence.

The final thrust of this research focused on the identification of functional human-like ancestors and a knowledge of which mutations have a neutral or slightly positive effect on preventing protein aggregation, promoting protein stability, or enhancing protein activity in order to generate additional Ancestral-Human chimeric proteins. In addition, we tested our most active ancestral uricase An19/22, which has a sequence identity of 94 % compared to the human uricase, in healthy rats against PBC which is the active protein component of Krystexxa® - the FDA approved uricase for the treatment failure gout. These pharmacokinetics studies showed that unmodified An19/22 is more stable at physiological pH than PBC and displays a statistically significant increase in its half-life. In addition, we have begun to identify functional PEG groups that should be explored to further improve the solubility, stability and activity of our engineered uricases.

In total, this research has encompassed an evolutionary synthetic biology approach by bringing the evolutionary grounded backwards-to-today approach of Ancestral Sequence Reconstruction to bear on the medical

problems caused by elevated uric acid levels, especially gout to develop a novel uricase that is both safer and more effective for human patients.

6.3 REFERENCES

110. Kelly, S.J., M. Delnomdedieu, M.I. Oliverio, L.D. Williams, M.G.P. Saifer, M.R. Sherman, T.M. Coffman, G.A. Johnson, and M.S. Hershfield, *Diabetes insipidus in uricase-deficient mice: A model for evaluating therapy with poly(ethylene glycol)-modified uricase*. J. Am. Soc. Nephrol., 2001. **12**(5): p. 1001-1009.

128. Fee, C.J. and J.M. Van Alstine, *Prediction of the viscosity radius and the size exclusion chromatography behavior of PEGylated proteins.* Bioconjugate Chem., 2004. **15**(6): p. 1304-1313.

131. Chen, R.H., A. Abuchowski, T. Van Es, N.C. Palczuk, and F.F. Davis, *Properties of two urate oxidases modified by the covalent attachment of poly(ethylene glycol)*. Biochim. Biophys. Acta, 1981. **660**(2): p. 293-298.

149. Tsuji, J., K. Hirose, E. Kasahara, M. Naitoh, and I. Yamamoto, *Studies on antigenicity of the polyethylene glycol (PEG)-modified uricase*. Int. J. Immunopharmacol., 1985. **7**(5): p. 725-730.

150. Caliceti, P., O. Schiavon, and F.M. Veronese, *Biopharmaceutical properties of uricase conjugated to neutral and amphiphilic polymers.* Bioconjug. Chem., 1999. **10**(4): p. 638-646.

151. Caliceti, P., O. Schiavon, and F.M. Veronese, *Immunological properties of uricase conjugated to neutral soluble polymers.* Bioconjug. Chem., 2001. **12**(4): p. 515-522.

152. Stevenson, W.S., C.D. Hyland, J.G. Zhang, P.O. Morgan, T.A. Willson, A. Gill, A.A. Hilton, E.M. Viney, M. Bahlo, S.L. Masters, et al., *Deficiency of 5-hydroxyisourate hydrolase causes hepatomegaly and hepatocellular carcinoma in mice*. Proc. Natl. Acad. Sci. U. S. A., 2010. **107**(38): p. 16625-16630.

SUPPORTING MATERIAL

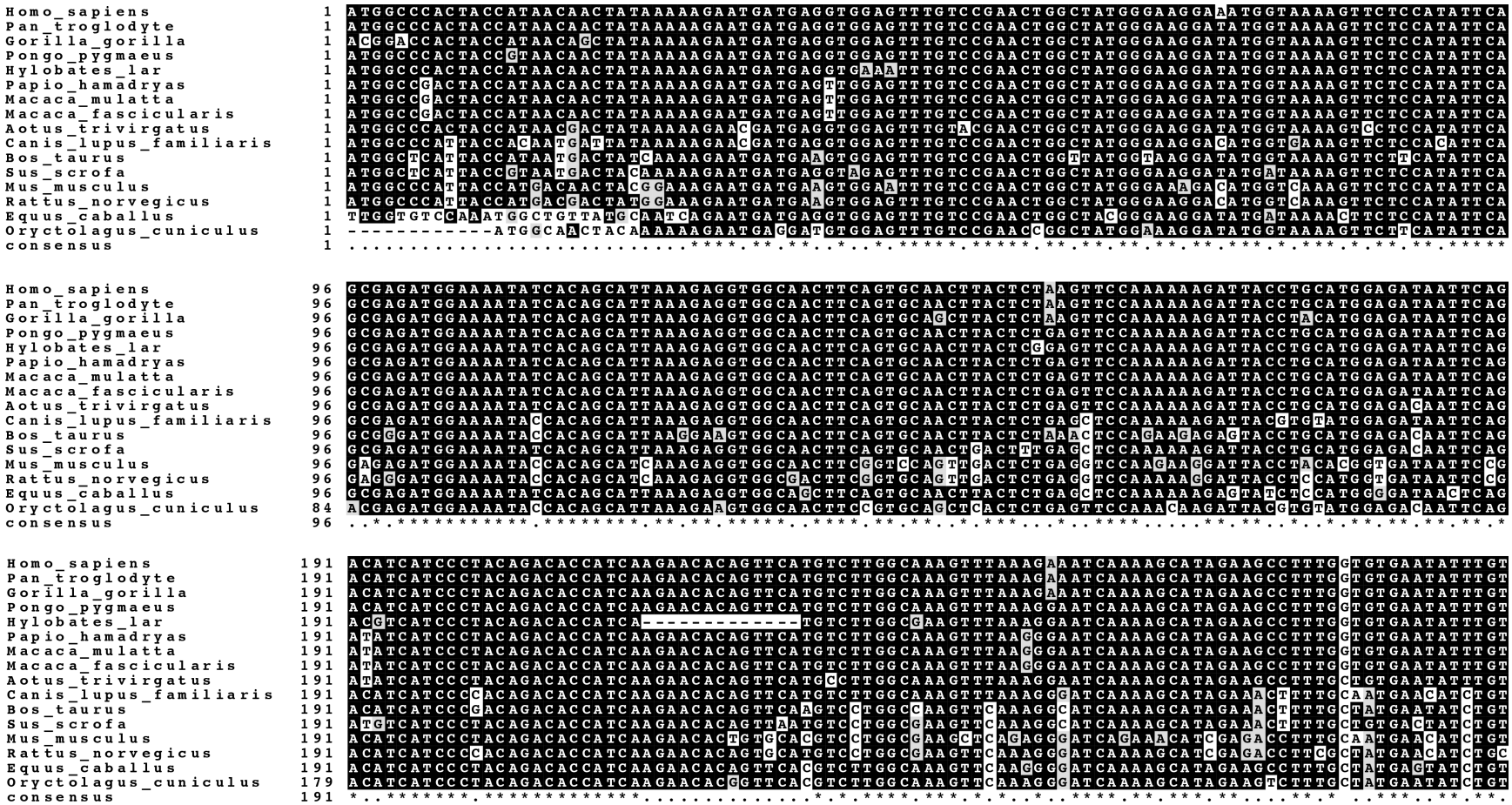
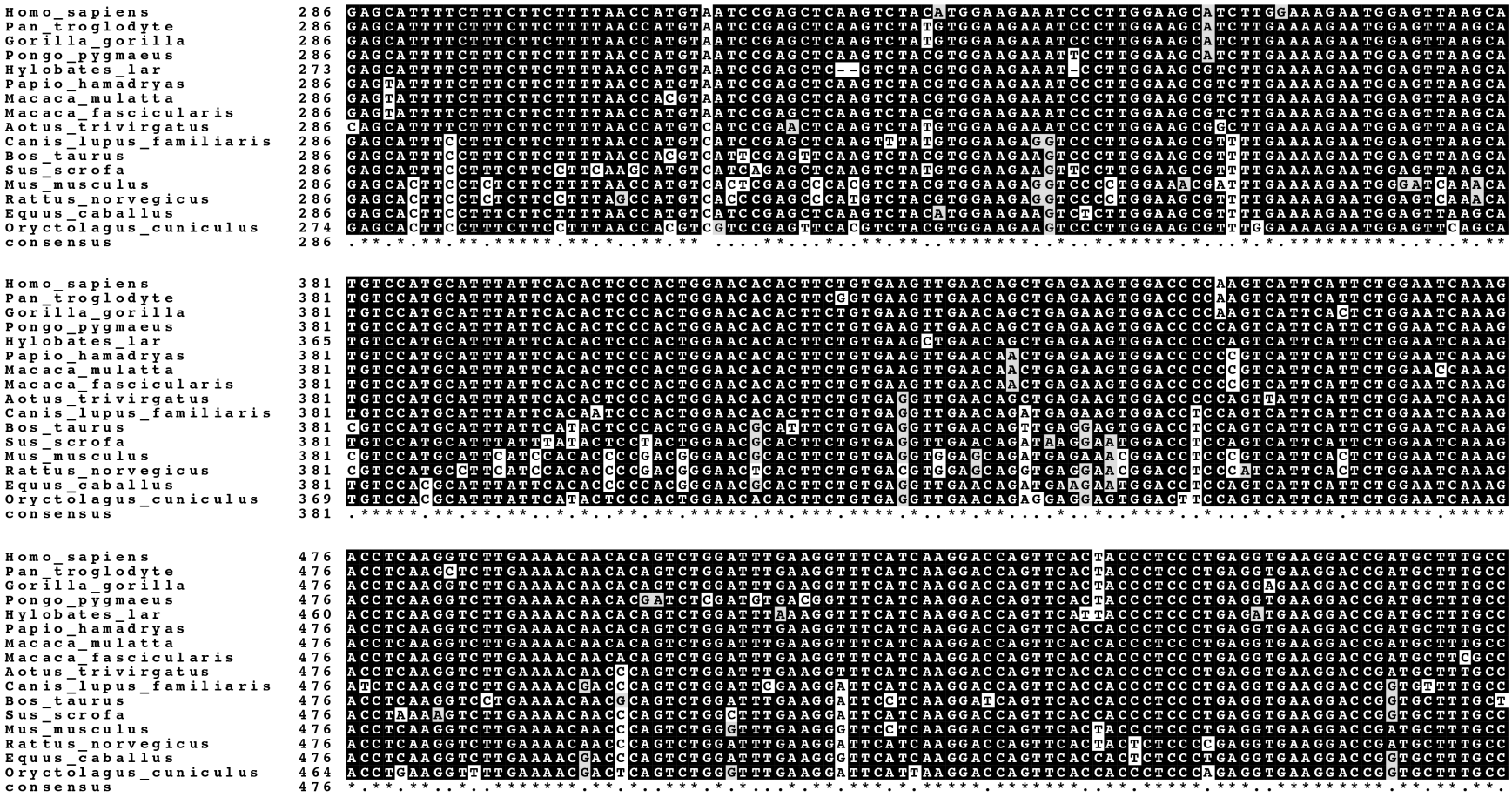
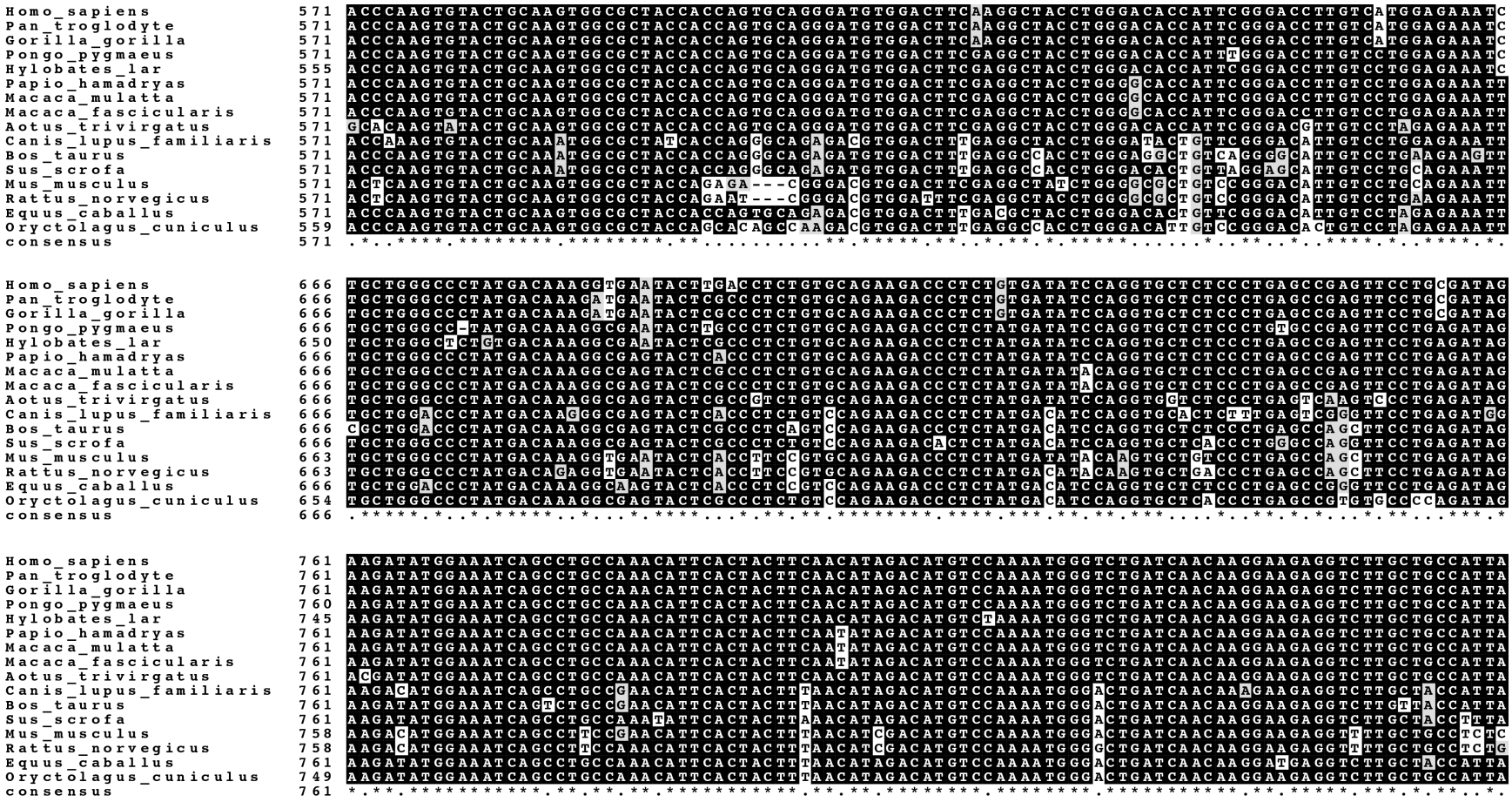


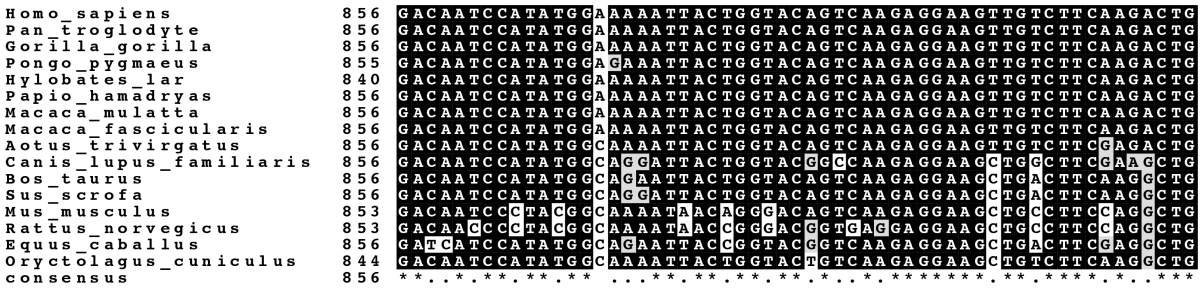
Figure S1 DNA multiple sequence alignment generated for ancestral sequence reconstruction. (MSA continues on following pages.)



(**Figure S1**, *page 2)*

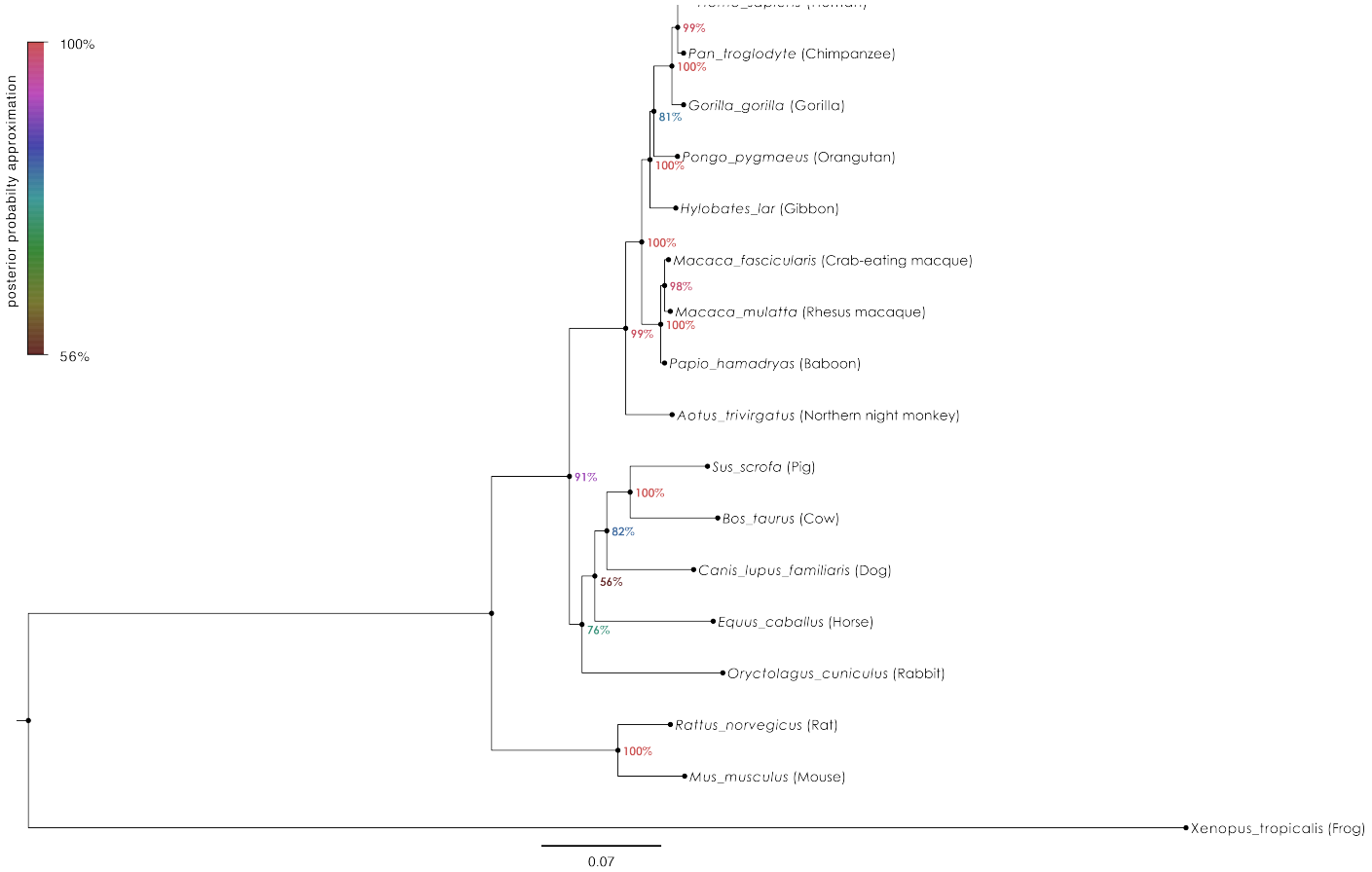


(**Figure S1**, *page 3)*



*(**Figure S1**, last page)*
The alignment was generated using ClustalX and manually edited. Shading was performed using Boxshade[4] Background shading is as follows: black are conserved sites, grey are transitions, and white are transversions.

4. The Boxshade server is available at http://www.ch.embnet.org/software/BOXC_form.html



(**Figure S2 A robust phylogeny was generated with MrBayes**. Shown is the phylogram of the consensus tree generated from 1,000,000 generations of a MCMCMC searching tree-space. The branch length scale shown at bottom is in units of changes per codon. The internal branching nodes are labeled with an approximation of the posterior probability (i.e. statistical confidence) of the branching events in the topology. The frog (Xenopus_tropicalis) was a suitable out-group sequence because of the long genetic distance (i.e. branch length) relative to the mammalian sequences.

Sequence S1 Primers used to query cDNA library for hUox

Forward primer exon 1 - UOX_Hs_F1
Sequence: 5' ATG GCC CAC TAC CAT AAC AAC 3'

Forward primer exon 2 - UOX_Hs_F2
Sequence: 5' GAT GAG GTG GAG TTT GTC CGA ACT G 3'

Forward primer exon 3 - UOX_Hs_F3
Sequence: 5'GCA TAG AAG CCT TTG GTG TG 3'

Forward primer exon 4 - UOX_Hs_F4
Sequence: 5' GGA GTT AAG CAT GTC CAT GC 3'

Forward primer exon 5 - UOX_Hs_F5
Sequence: 5' GGA CCC CAA GTC ATT CAT TCT GG 3'

Forward primer exon 6 - UOX_Hs_F6
Sequence: 5' GGA CAC CAT TCG GGA CCT TG 3'

Forward primer exon 7 - UOX_Hs_F7
Sequence 5' GAT ATG GAA ATC AGC CTG CC 3'

Forward primer exon 8 - UOX_Hs_F8
Sequence: 5' GTC TTG CTG CCA TTA GAC AAT CC 3'

Reverse primer inside of exon 8 - UOX_Hs_R1
Sequence: 5' GAG GAA GTT GTC TTC AAG ACT GTG A 3'

Reverse complement (primer): 5' TCA CAG TCT TGA AGA CAA CTT CCT C 3'

Reverse primer outside of exon 8 - UOX_Hs_R2
Sequence: 5' GCT GAG ATT GGA CTC CTA TTG 3'
Note: these primers and cDNA can be found in "XTINA box #3, cDNA uricase, in lanes A1-A9, B1-B9, C1-C9, D1-D9, and E9 and F9.

Sequence S2 Protein sequences of functional uricases used for ASR. Seventeen modern day uricase encoding sequences were retrieved from databases. Shown below are the encoded amino acid sequences used to generate the multiples sequence alignment for the computational component of the Ancestral Sequence Reconstruction (ASR) Approach).

>Papio_hamadryas (Baboon)
MADYHNNYKKNDELEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKSIEAFGVNICEYFLSSFNHVIRAQVYVEEIPWKRLEKNGVKHVHA
FIHTPTGTHFCEVEQLRSGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQCRDVDFEATWGTIRDLVLEKFAGPYDKGEYSPSVQKTLYDIQVLSLSRVPEIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>Macaca_mulatta (Rhesus Monkey)
MADYHNNYKKNDELEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKSIEAFGVNICEYFLSSFNHVIRAQVYVEEIPWKRLEKNGVKHVHA
FIHTPTGTHFCEVEQLRSGPPVIHSGTKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQCRDVDFEATWGTIRDLVLEKFAGPYDKGEYSPSVQKTLYDIQVLSLSRVPEIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>Macaca_fascicularis (Rhesus Monkey)
MADYHNNYKKNDELEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKSIEAFGVNICEYFLSSFNHVIRAQVYVEEIPWKRLEKNGVKHVHA
FIHTPTGTHFCEVEQLRSGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQCRDVDFEATWGTIRDLVLEKFAGPYDKGEYSPSVQKTLYDIQVLSLSRVPEIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>Aotus_trivirgatus (Owl Monkey)
MAHYHNDYKKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDI
IPTDTIKNTVHALAKFKGIKSIEAFAVNICQHFLSSFNHVIRTQVYVEEIPWKRLEKNGVKHVHA
FIHTPTGTHFCEVEQLRSGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFAAQVYC
KWRYHQCRDVDFEATWDTIRDVVLEKFAGPYDKGEYSPSVQKTLYDIQVVSLSQVPEIDDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>Canis_lupus_familiaris (Dog)
MAHYHNDYKKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYVYGDNSDI
IPTDTIKNTVHVLAKFKGIKSIETFAMNICEHFLSSFNHVIRAQVYVEEVPWKRFEKNGVKHVHA
FIHNPTGTHFCEVEQMRSGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATKVYC
KWRYHQGRDVDFEATWDTVRDIVLEKFAGPYDKGEYSPSVQKTLYDIQVHSLSRVPEMEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGRITGTAKRKLASKL

>Bos_taurus (Cow)
MAHYHNDYQKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLNSRREYLHGDNSDI
IPTDTIKNTVQVLAKFKGIKSIETFAMNICEHFLSSFNHVIRVQVYVEEVPWKRFEKNGVKHVHA
FIHTPTGTHFCEVEQLRSGPPVIHSGIKDLKVLKTTQSGFEGFLKDQFTTLPEVKDRCFATQVYC
KWRYHQGRDVDFEATWEAVRGIVLKKFAGPYDKGEYSPSVQKTLYDIQVLSLSQLPEIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGRITGTVKRKLTSRL

>Sus_scrofa (Pig)
MAHYRNDYKKNDEVEFVRTGYGKDMIKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDV
IPTDTIKNTVNVLAKFKGIKSIETFAVTICEHFLSSFKHVIRAQVYVEEVPWKRFEKNGVKHVHA

FIYTPTGTHFCEVEQIRNGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQGRDVDFEATWDTVRSIVLQKFAGPYDKGEYSPSVQKTLYDIQVLTLGQVPEIEDMEISL
PNIHYLNIDMSKMGLINKEEVLLPLDNPYGRITGTVKRKLTSRL

>Mus_musculus (Mouse)
MAHYHDNYGKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLRSKKDYLHGDNSDI
IPTDTIKNTVHVLAKLRGIRNIETFAMNICEHFLSSFNHVTRAHVYVEEVPWKRFEKNGIKHVHA
FIHTPTGTHFCEVEQMRNGPPVIHSGIKDLKVLKTTQSGFEGFLKDQFTTLPEVKDRCFATQVYC
KWRYQRRDVDFEAIWGAVRDIVLQKFAGPYDKGEYSPSVQKTLYDIQVLSLSQLPEIEDMEISLP
NIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLPSRL

>Rattus_norvegicus (Rat)
MAHYHDDYGKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLRSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKSIETFAMNICEHFLSSFSHVTRAHVYVEEVPWKRFEKNGVKHVHA
FIHTPTGTHFCDVEQVRNGPPIIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYQNRDVDFEATWGAVRDIVLKKFAGPYDRGEYSPSVQKTLYDIQVLTLSQLPEIEDMEISLP
NIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVRRKLPSRL

>Equus_caballus (Horse)
LVSKWLLCNQNDEVEFVRTGYGKDMIKLLHIQRDGKYHSIKEVAASVQLTLSSKKEYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKSIEAFAMSICEHFLSSFNHVIRAQVYMEEVSWKRFEKNGVKHVHA
FIHTPTGTHFCEVEQMKNGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQCRDVDFDATWDTVRDIVLEKFAGPYDKGKYSPSVQKTLYDIQVLSLSRVPEIEDMEISL
PNIHYFNIDMSKMGLINKDEVLLPLDHPYGRITGTVKRKLTSRL

>Oryctolagus_cuniculus (Rabbit)
MATTKKNEDVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKQDYVYGDNSDIIPTD
TIKNTVHVLAKFKGIKSIEVFAMNICEHFLSSFNHVVRVHVYVEEVPWKRLEKNGVQHVHAFIHT
PTGTHFCEVEQRRSGLPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYCKWRY
QHSQDVDFEATWDIVRDTVLEKFAGPYDKGEYSPSVQKTLYDIQVLTLSRVPQIEDMEISLPNIH
YFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>Xenopus_tropicalis (Frog)
MAQYHGRLSKDSDVEFAHTAYGKNAVKVLQIKRNGKQHFIKEIEVSVQLTLKSKKDYLEGDNSDI
IPTDTIKNTIYALTKLKGIQTIEEFSVEIARHFLTSFNHVTEVKVFINEAPWRRMEKNGMSHVHA
FIYSPEGVHFCELQQKRGGQPAIFSGIKELRILKTTQSGFEGFIKDRFTTLPEVKDRCFSTIVNC
KWKYGTSKAVDYDAVWKTILETILDTFAGPYDKGEYSPSVQKTLYDIQVLSLRKVPEIEEIEIIL
PNKHYFTIDMSKMGLTNQDEVLMPTDIPYGNIAGTLRRNPSSKL

Sequence S3 Inferred ancestral uricase protein sequences.

>An18
MAHYHGHLTKNAEVEFVRTGYGKDVVKVLHIQRDGKHHIIKEVATSVQLTLNSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKTIEAFAMNIGKHFLSSFNHVIRAQVYVEEVPWKRFEKNGVNHVHA
FIHTPTGTHFCEVEQKRGGPPVIHSGIKDLKVLKTTQSGFEGFIKDRFTTLPEVKDRCFATQVYC
KWRYDQSRAVDFEAIWDTVLDIVLEKFAGPYDKGEYSPSVQKTLYDIQVLSLSRVPEIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>An19/22
MAHYHNDYKKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKSIEAFAMNICEHFLSSFNHVIRAQVYVEEVPWKRFEKNGVKHVHA
FIHTPTGTHFCEVEQMRSGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQGRDVDFEATWDTVRDIVLEKFAGPYDKGEYSPSVQKTLYDIQVLSLSRVPEIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>An26
MAHYHNDYKKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKSIEAFAVNICEHFLSSFNHVIRAQVYVEEIPWKRLEKNGVKHVHA
FIHTPTGTHFCEVEQLRSGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQCRDVDFEATWDTIRDLVLEKFAGPYDKGEYSPSVQKTLYDIQVLSLSRVPEIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>An27
MAHYHNNYKKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKSIEAFGVNICEHFLSSFNHVIRAQVYVEEIPWKRLEKNGVKHVHA
FIHTPTGTHFCEVEQLRSGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQCRDVDFEATWDTIRDLVLEKFAGPYDKGEYSPSVQKTLYDIQVLSLSRVPEIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>An30
MAHYHNNYKKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKSIEAFGVNICEHFLSSFNHVIRAQVYVEEIPWKRLEKNGVKHVHA
FIHTPTGTHFCEVEQLRSGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQCRDVDFEATWDTIRDLVLEKSAGPYDKGEYSPSVQKTLYDIQVLSLSRVPEIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>An31
MAHYHNNYKKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKGIKSIEAFGVNICEHFLSSFNHVIRAQVYVEEIPWKHLEKNGVKHVHA
FIHTPTGTHFCEVEQLRSGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQCRDVDFEATWDTIRDLVLEKSAGPYDKGEYSPSVQKTLYDIQVLSLSRVPEIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>An32
MAHYHNNYKKNDEVEFVRTGYGKDMVKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDI
IPTDTIKNTVHVLAKFKEIKSIEAFGVNICEHFLSSFNHVIRAQVYVEEIPWKHLEKNGVKHVHA
FIHTPTGTHFCEVEQLRSGPQVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQCRDVDFKATWDTIRDLVMEKSAGPYDKDEYSPSVQKTLCDIQVLSLSRVPAIEDMEISL
PNIHYFNIDMSKMGLINKEEVLLPLDNPYGKITGTVKRKLSSRL

>Pig
MAHYRNDYKKNDEVEFVRTGYGKDMIKVLHIQRDGKYHSIKEVATSVQLTLSSKKDYLHGDNSDV
IPTDTIKNTVNVLAKFKGIKSIETFAVTICEHFLSSFKHVIRAQVYVEEVPWKRFEKNGVKHVHA
FIYTPTGTHFCEVEQIRNGPPVIHSGIKDLKVLKTTQSGFEGFIKDQFTTLPEVKDRCFATQVYC
KWRYHQGRDVDFEATWDTVRSIVLQKFAGPYDKGEYSPSVQKTLYDIQVLTLGQVPEIEDMEISL
PNIHYLNIDMSKMGLINKEEVLLPLDNPYGRITGTVKRKLTSRL

PUBLICATIONS

Gaucher, E.A., Kratzer, J.T., R.N Randall (2010). *Deep phylogeny-how a tree can help characterize early life on Earth*. Cold Spring Harb Perspect Biol. 2(1): a002238.

Kratzer, J. T., Cole, M. F., and E. A. Gaucher (2012). *Protein engineering guided by natural diversity*. Protein Engineering Handbook, 3rd vol.; Lutz, S. and U. T. Bornscheuer Eds.; Wiley-VCH: Weinheim, Germany, 2013. In press.

Cacan, E.*, Kratzer, J. T.*, Cole, M. F., and E. A. Gaucher. Interchanging functionality among homologous elongation factors using signatures of heterotachy. Journal of molecular evolution, 2013. 76(1-2): p. 4-12.

Kratzer, J. T., Lanaspa, M. G., Johnson, R. J. & Gaucher, E. A. Management of intracellular triglyceride levels induced by uric acid using ancient uricase enzymes. In preparation.

Kratzer, J.T., Murphy, M. N., Ortlund, E. A. & Gaucher, E. A. *Evolutionary history of modern and ancient mammalian uricases*. In preparation.

Kratzer, J. T., Gaucher, E. A. *Use of PEGylated ancestral uricases*. US PATENT APPLICATION. In preparation.

REFERENCES

1. Dobzhansky, T., *Biology, molecular and organismic* Am. Zool., 1964. **4**: p. 443-452.

2. Elias, Y. and R.H. Huang, *Biochemical and structural studies of A-to-I editing by tRNA:A34 deaminases at the wobble position of transfer RNA.* Biochemistry, 2005. **44**(36): p. 12057-12065.

3. Keebaugh, A.C. and J.W. Thomas, *The evolutionary fate of the genes encoding the purine catabolic enzymes in hominoids, birds, and reptiles*, in *Mol. Biol. Evol.* 2010: United States. p. 1359-1369.

4. Hayashi, S., S. Fujiwara, and T. Noguchi, *Evolution of urate-degrading enzymes in animal peroxisomes.* Cell Biochem. Biophys., 2000. **32**: p. 123-129.

5. Iwata, H., S. Nishio, M. Yokoyama, A. Matsumoto, and M. Takeuchi, *Solubility of uric acid and supersaturation of monosodium urate: why is uric acid so highly soluble in urine?* J. Urol., 1989. **142**(4): p. 1095-1098.

6. Gabison, L., M. Chiadmi, N. Colloc'h, B. Castro, M. El Hajji, and T. Prange, *Recapture of S -allantoin, the product of the two-step degradation of uric acid, by urate oxidase.* FEBS Lett., 2006. **580**(8): p. 2087-2091.

7. Tipton, P.A., *Urate to allantoin, specifically (S)-allantoin.* Nat. Chem. Biol., 2006. **2**(3): p. 124-125.

8. Ramazzina, I., C. Folli, A. Secchi, R. Berni, and R. Percudani, *Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes.* Nat. Chem. Biol., 2006. **2**(3): p. 144-148.

9. Nuki, G. and P.A. Simkin, *A concise history of gout and hyperuricemia and their treatment.* Arthritis Res. Ther., 2006. **8**: p. 5.

10. Burns, C.M. and R.L. Wortmann, *Gout therapeutics: new drugs for an old disease.* Lancet, 2011. **377**(9760): p. 165-177.

11. Adams, J.U., *New relief for gout.* Nat. Biotechnol., 2009. **27**(4): p. 309-311.

12. Campion, E.W., R.J. Glynn, and L.O. DeLabry, *Asymptomatic hyperuricemia. Risks and consequences in the normative aging study.* Am. J. Med., 1987. **82**(3): p. 421-426.

13. Teng, G.G., R. Nair, and K.G. Saag, *Pathophysiology, clinical presentation and treatment of gout.* Drugs, 2006. **66**(12): p. 1547-1563.

14. Schumacher, H.R., Jr., *The pathogenesis of gout.* Clevel. Clin. J. Med., 2008. **75 Suppl 5**: p. S2-4.

15. Cardona, F., F.J. Tinahones, E. Collantes, A. Escudero, E. Garcia-Fuentes, and F.J. Soriguer, *The elevated prevalence of apolipoprotein E2 in patients with gout is associated with reduced renal excretion of urates.* Rheumatology (Oxf.), 2003. **42**(3): p. 468-472.

16. de Oliveira, E.P. and R.C. Burini, *High plasma uric acid concentration: causes and consequences.* Diabetol. Metab. Syndr., 2012. **4**: p. 12.

17. Johnson, R.J., L.G. Sanchez-Lozada, and T. Nakagawa, *The effect of fructose on renal biology and disease.* J. Am. Soc. Nephrol., 2010. **21**(12): p. 2036-2039.

18. Tiu, R.V., S.E. Mountantonakis, A.J. Dunbar, and M.J. Schreiber, *Tumor lysis syndrome.* Semin. Thromb. Hemost., 2007. **33**(4): p. 397-407.

19. Davidson, M.B., S. Thakkar, J.K. Hix, N.D. Bhandarkar, A. Wong, and M.J. Schreiber, *Pathophysiology, clinical consequences, and treatment of tumor lysis syndrome.* Am. J. Med., 2004. **116**(8): p. 546-554.

20. Juraschek, S.P., L.C. Kovell, E.R. Miller, 3rd, and A.C. Gelber, *Association of kidney disease with prevalent gout in the United States in 1988-1994 and 2007-2010.* Semin. Arthritis Rheum., 2013.

21. Kottgen, A., E. Albrecht, A. Teumer, V. Vitart, J. Krumsiek, C. Hundertmark, G. Pistis, D. Ruggiero, C.M. O'Seaghdha, T. Haller, et al., *Genome-wide association analyses identify 18 new loci associated with serum urate concentrations.* Nat. Genet., 2013. **45**(2): p. 145-154.

22. Ali, S. and E.V. Lally, *Treatment failure gout.* Med. Health R. I., 2009. **92**(11): p. 369-371.

23. Edwards, N.L., *Treatment-failure gout: a moving target.* Arthritis Rheum., 2008. **58**(9): p. 2587-2590.

24. Bardin, T., *Current management of gout in patients unresponsive or allergic to allopurinol.* Joint Bone Spine, 2004. **71**(6): p. 481-485.

25. Stamp, L.K., J.L. O'Donnell, and P.T. Chapman, *Emerging therapies in the long-term management of hyperuricaemia and gout.* Intern. Med. J., 2007. **37**(4): p. 258-266.

26. Riedel, A.A., M. Nelson, N. Joseph-Ridge, K. Wallace, P. MacDonald, and M. Becker, *Compliance with allopurinol therapy among managed care enrollees with gout: a retrospective analysis of administrative claims.* J. Rheumatol., 2004. **31**(8): p. 1575-1581.

27. Terkeltaub, R., *Gout. Novel therapies for treatment of gout and hyperuricemia.* Arthritis Res. Ther., 2009. **11**(4): p. 236.

28. Fels, E. and J.S. Sundy, *Refractory gout: what is it and what to do about it?* Curr. Opin. Rheumatol., 2008. **20**(2): p. 198-202.

29. Sundy, J.S. and M.S. Hershfield, *Uricase and other novel agents for the management of patients with treatment-failure gout.* Curr. Rheumatol. Rep., 2007. **9**(3): p. 258-264.

30. Perry, M.E. and R. Madhok, *Treatment failure gout: failure to treat?* Rheumatology (Oxf.), 2010. **49**(12): p. 2233-2234.

31. Luk, A.J. and P.A. Simkin, *Epidemiology of hyperuricemia and gout.* Am. J. Manag. Care, 2005. **11**(15 Suppl): p. S435-442; quiz S465-438.

32. Przylecki, S.J., *Uricase and its action.* Biochem. J., 1930. **24**(1): p. 81-81.

33. Keilin, D. and E.F. Hartree, *Uricase, amino acid oxidase, and xanthine oxidase.* Proc. R. Soc. B-Biol. Sci., 1936. **119**(813): p. 114-140.

34. Bentley, R. and A. Neuberger, *The mechanism of the action of uricase.* Biochem. J., 1952. **52**(5): p. 694-699.

35. Hruban, Z. and H. Swift, *Uricase - localization in hepatic microbodies.* Science, 1964. **146**(364): p. 1316-&.

36. Pitts, O.M. and D.G. Priest, *Uricse reaction intermediate - mechanism of borate and hydroxide ion catalysis.* Biochemistry, 1973. **12**(7): p. 1358-1363.

37. Salleh, A.B. and W.M. Ledingham, *Some kinetic-studies on immobilized uricase.* Int. J. Biochem., 1981. **13**(10): p. 1113-1118.

38. Suzuki, H. and D.P.S. Verma, *Soybean nodule-specific uricase (Nodulin-35) is expressed and assembled into a functional tetrmeric holoenzyme in Eschericia-coli.* Plant Physiol., 1991. **95**(2): p. 384-389.

39. Imhoff, R.D., N.P. Power, M.J. Borrok, and P.A. Tipton, *General base catalysis in the urate oxidase reaction: evidence for a novel Thr-Lys catalytic diad.* Biochemistry, 2003. **42**(14): p. 4094-4100.

40. Modric, N., A.E. Derome, S.J.H. Ashcroft, and M. Poje, *Tracing and identification of uricase reaction intermediates - A direct C-13-NMR isotope-labeling evidence*. Tetrahedron Lett., 1992. **33**(44): p. 6691-6694.

41. Kahn, K. and P.A. Tipton, *Kinetic mechanism and cofactor content of soybean root nodule urate oxidase*. Biochemistry, 1997. **36**(16): p. 4731-4738.

42. Gabison, L., C. Chopard, N. Colloc'h, F. Peyrot, B. Castro, M. El Hajji, M. Altarsha, G. Monard, M. Chiadmi, and T. Prange, *X-ray, ESR, and quantum mechanics studies unravel a spin well in the cofactor-less urate oxidase*. Proteins, 2011. **79**(6): p. 1964-1976.

43. Mahler, H.R., G. Hubscher, and R. Baum, *Studies on uricase. I. Preparation, purification, and properties of a cuproprotein.* J. Biol. Chem., 1955. **216**(2): p. 625-641.

44. Voet, D. and J.G. Voet, *Biochemistry*. 3rd ed. 2004, New York: Wiley. xvii, 1223 p.

45. Colloc'h, N., *Crystal Structure of the protein drug urate oxidase-inhibitor complex at 2.05 A Resolution.* Nat. Struct. Biol., 1997. **4**(11): p. 947-952.

46. Oda, M., Y. Satta, O. Takenaka, and N. Takahata, *Loss of urate oxidase activity in hominoids and its evolutionary implications.* Mol. Biol. Evol., 2002. **19**(5): p. 640-653.

47. Varelaechavarria, A., R.M. Deocaluna, and H.A. Barrerasaldana, *Uricase protein sequences - conserved during vertebrate evolution but absent in humans.* Faseb J., 1988. **2**(15): p. 3092-3096.

48. Logan, D.C., D.E. Wilson, C.M. Flowers, P.J. Sparks, and F.H. Tyler, *Uric acid catabolism in the woolly monkey.* Metabolism., 1976. **25**(5): p. 517-522.

49. Usuda, N., M.K. Reddy, T. Hashimoto, M.S. Rao, and J.K. Reddy, *Tissue specificity and species differences in the distribution of urate oxidase in peroxisomes.* Lab. Invest., 1988. **58**(1): p. 100-111.

50. Johnson, R.J., E.A. Gaucher, Y.Y. Sautin, G.N. Henderson, A.J. Angerhofer, and S.A. Benner, *The planetary biology of ascorbate and uric acid and their relationship with the epidemic of obesity and cardiovascular disease*. Med. Hypotheses, 2008. **71**(1): p. 22-31.

51. Mayne, N., S. Keady, and M. Thacker, *Rasburicase in the prevention and treatment of tumour lysis syndrome*. Intensive Crit. Care Nurs., 2008. **24**(1): p. 59-62.

52. Bomalaski, J.S., F.W. Holtsberg, C.M. Ensor, and M.A. Clark, *Uricase formulated with polyethylene glycol (uricase-PEG 20): Biochemical rationale and preclinical studies.* J. Rheumatol., 2002. **29**(9): p. 1942-1949.

53. Goldman, S.C., J.S. Holcenberg, J.Z. Finklestein, R. Hutchinson, S. Kreissman, F.L. Johnson, C. Tou, E. Harvey, E. Morris, and M.S. Cairo, *A randomized comparison between rasburicase and allopurinol in children with lymphoma or leukemia at high risk for tumor lysis.* Blood, 2001. **97**(10): p. 2998-3003.

54. Hershfield, M. and S.J. Kelly. Urate oxidase. 7056713 USPTO application. 2006.

55. Sundy, J.S., H.S. Baraf, R.A. Yood, N.L. Edwards, S.R. Gutierrez-Urena, E.L. Treadwell, J. Vazquez-Mellado, W.B. White, P.E. Lipsky, Z. Horowitz, et al., *Efficacy and tolerability of pegloticase for the treatment of chronic gout in patients refractory to conventional treatment: two randomized controlled trials.* JAMA, 2011. **306**(7): p. 711-720.

56. Lyseng-Williamson, K.A., *Pegloticase: in treatment-refractory chronic gout.* Drugs, 2011. **71**(16): p. 2179-2192.

57. Li, W.-H., *Molecular evolution*. 1997, Sunderland, Mass.: Sinauer Associates. xv, 487 p.

58. Vanin, E.F., *Processed pseudogenes: characteristics and evolution.* Annu. Rev. Genet., 1985. **19**: p. 253-272.

59. Tutar, Y., *Pseudogenes.* Comp. Funct. Genomics, 2012.

60. Rouchka, E.C. and I.E. Cha, *Current Trends in Pseudogene Detection and Characterization.* Curr. Bioinf., 2009. **4**(2): p. 112-119.

61. Wu, X.W., C.C. Lee, D.M. Muzny, and C.T. Caskey, *Urate oxidase: primary structure and evolutionary implications.* Proc. Natl. Acad. Sci. U. S. A., 1989. **86**(23): p. 9412-9416.

62. Yeldandi, A.V., V. Yeldandi, S. Kumar, C.V. Murthy, X.D. Wang, K. Alvares, M.S. Rao, and J.K. Reddy, *Molecular evolution of the urate oxidase-encoding gene in hominoid primates: nonsense mutations.* Gene, 1991. **109**(2): p. 281-284.

63. Lai, H.M., Y.Y. Chiang, C.C. Hsu, and F. Wu, *A recognition machine for CpG-islands based on Boltzmann model.* J Med Biol Eng, 2008. **28**(1): p. 23-30.

64. Benner, S.A., S.O. Sassi, and E.A. Gaucher, *Molecular Paleoscience: Systems Biology from the Past*, in *Advances in Enzymology*, E.J. Toone, Editor. 2010, John Wiley & Sons, Inc. p. 1-132.

65. Bomalaski, J.S. and M.A. Clark, *Serum uric acid-lowering therapies: where are we heading in management of hyperuricemia and the potential role of uricase*. Curr. Rheumatol. Rep., 2004. **6**(3): p. 240-247.

66. Vogels, G.D. and C. Van der Drift, *Degradation of purines and pyrimidines by microorganisms.* Bacteriol. Rev., 1976. **40**(2): p. 403-468.

67. Needham, J., *Chemical embryology*. 1963, New York,: Hafner Pub. Co.

68. Gersch, C., S.P. Palii, W. Imaram, K.M. Kim, S.A. Karumanchi, A. Angerhofer, R.J. Johnson, and G.N. Henderson, *Reactions of peroxynitrite with uric acid: formation of reactive intermediates, alkylated products and triuret, and in vivo production of triuret under conditions of oxidative stress.* Nucleosides, 2009. **28**(2): p. 118-149.

69. Gersch, C., S.P. Palii, K.M. Kim, A. Angerhofer, R.J. Johnson, and G.N. Henderson, *Inactivation of nitric oxide by uric acid.* Nucleosides, 2008. **27**(8): p. 967-978.

70. Waring, W.S., A. Convery, V. Mishra, A. Shenkin, D.J. Webb, and S.R. Maxwell, *Uric acid reduces exercise-induced oxidative stress in healthy adults.* Clin. Sci., 2003. **105**(4): p. 425-430.

71. Ames, B.N., R. Cathcart, E. Schwiers, and P. Hochstein, *Uric acid provides an antioxidant defense in humans against oxidant- and radical-caused aging and cancer: a hypothesis.* Proc. Natl. Acad. Sci. U. S. A., 1981. **78**(11): p. 6858-6862.

72. Watanabe, S., D.H. Kang, L.L. Feng, T. Nakagawa, J. Kanellis, H. Lan, M. Mazzali, and R.J. Johnson, *Uric acid, hominoid evolution, and the pathogenesis of salt-sensitivity.* Hypertension, 2002. **40**(3): p. 355-360.

73. Lanaspa, M.A., L.G. Sanchez-Lozada, Y.J. Choi, C. Cicerchi, M. Kanbay, C.A. Roncal-Jimenez, T. Ishimoto, N. Li, G. Marek, M. Duranay, et al., *Uric acid induces hepatic steatosis by generation of mitochondrial oxidative stress: potential role in fructose-dependent and -independent fatty liver.* J. Biol. Chem., 2012. **287**(48): p. 40732-40744.

74. Lanaspa, M., L. Sanchez-Lozada, C. Cicerchi, N. Li, C. Roncal-Jimenez, T. Ishimoto, M. Le, G. Garcia, J. Thomas, C. Rivard, et al., *Uric acid stimulates fructokinase and accelerates fructose metabolism in the development of fatty liver.* PLoS One, 2012. **7**(10).

75. Jungreis, I., M.F. Lin, R. Spokony, C.S. Chan, N. Negre, A. Victorsen, K.P. White, and M. Kellis, *Evidence of abundant stop codon readthrough in Drosophila and other metazoa.* Genome Res., 2011. **21**(12): p. 2096-2113.

76. Wu, X.W., D.M. Muzny, C.C. Lee, and C.T. Caskey, *Two independent mutational events in the loss of urate oxidase during hominoid evolution.* J. Mol. Evol., 1992. **34**(1): p. 78-84.

77. Sogaard, T.M., C.G. Jakobsen, and J. Justesen, *A sensitive assay of translational fidelity (readthrough and termination) in eukaryotic cells.* Biochemistry (Mosc), 1999. **64**(12): p. 1408-1417.

78. Manuvakhova, M., K. Keeling, and D.M. Bedwell, *Aminoglycoside antibiotics mediate context-dependent suppression of termination codons in a mammalian translation system.* RNA, 2000. **6**(7): p. 1044-1055.

79. Kratzer, J.T., Cole, M.F., and E.A. Gaucher, *Protein engineering guided by natural diversity*, in *Protein Engineering Handbook*, S. Lutz and U.T. Bornscheuer, Editors. 2013, Wiley-VCH: Weinheim, Germany.

80. Pauling, L. and E. Zuckerkandl, *Chemical Paleogenetics Molecular Restoration Studies of Extinct Forms of Life.* Acta Chem. Scand., 1963. **17**: p. S9-S16.

81. Stackhouse, J., S.R. Presnell, G.M. Mcgeehan, K.P. Nambiar, and S.A. Benner, *The Ribonuclease from an extinct bovid ruminant.* FEBS Lett., 1990. **262**(1): p. 104-106.

82. Yokoyama, S. and Y.S. Shi, *Molecular analysis of the evolutionary significance of ultraviolet vision in vertebrates.* Proc. Natl. Acad. Sci. U. S. A., 2003. **100**(14): p. 8308-8313.

83. Thornton, J.W., E. Need, and D. Crews, *Resurrecting the ancestral steroid receptor: Ancient origin of estrogen signaling.* Science, 2003. **301**(5640): p. 1714-1717.

84. Gaucher, E.A., S. Govindarajan, and O.K. Ganesh, *Palaeotemperature trend for Precambrian life inferred from resurrected proteins.* Nature, 2008. **451**(7179): p. 704-707.

85. Chen, F., E.A. Gaucher, N.A. Leal, D. Hutter, S.A. Havemann, S. Govindarajan, E.A. Ortlund, and S.A. Benner, *Reconstructed evolutionary adaptive paths give polymerases accepting reversible terminators for sequencing and SNP detection.* Proc. Natl. Acad. Sci. U. S. A., 2010. **107**(5): p. 1948-1953.

86. Thornton, J.W., *Resurrecting ancient genes: experimental analysis of extinct molecules.* Nat. Rev. Genet., 2004. **5**(5): p. 366-375.

87. Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-2948.

88. Notredame, C., D.G. Higgins, and J. Heringa, *T-Coffee: A novel method for fast and accurate multiple sequence alignment.* J. Mol. Biol., 2000. **302**(1): p. 205-217.

89. Mount, D.M., *Bioinformatics: Sequence and Genome Analysis* 2nd ed. 2004, Cold Spring Harbor, NY.: Cold Spring Harbor Laboratory Press.

90. Fitch, W.M., *Toward defining the course of evolution: minimum change for a specific tree topology.* Syst. Zool., 1971. **20**(4): p. 406-416.

91. Yang, Z., S. Kumar, and M. Nei, *A new method of inference of ancestral nucleotide and amino acid sequences.* Genetics, 1995. **141**(4): p. 1641-1650.

92. Huelsenbeck, J.P., F. Ronquist, R. Nielsen, and J.P. Bollback, *Bayesian inference of phylogeny and its impact on evolutionary biology.* Science, 2001. **294**(5550): p. 2310-2314.

93. Di Giulio, M., *The universal ancestor was a thermophile or a hyperthermophile: Tests and further evidence.* J Theor Biol, 2003. **221**(3): p. 425-436.

94. Woese, C.R., *Bacterial evolution.* Microbiol. Rev., 1987. **51**(2): p. 221-271.

95. Tawfik, D.S., S. Bershtein, and K. Goldin, *Intense neutral drifts yield robust and evolvable consensus proteins.* J. Mol. Biol., 2008. **379**(5): p. 1029-1044.

96. Tawfik, D.S. and N. Tokuriki, *Stability effects of mutations and protein evolvability.* Curr Opin Struc Biol, 2009. **19**(5): p. 596-604.

97. Perez-Jimenez, R., A. Ingles-Prieto, Z.M. Zhao, I. Sanchez-Romero, J. Alegre-Cebollada, P. Kosuri, S. Garcia-Manyes, T.J. Kappock, M. Tanokura, A. Holmgren, et al., *Single-molecule paleoenzymology probes the chemistry of resurrected enzymes.* Nat. Struct. Mol. Biol., 2011. **18**(5): p. 592-596.

98. Gaucher, E.A., J.M. Thomson, M.F. Burgan, and S.A. Benner, *Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins.* Nature, 2003. **425**(6955): p. 285-288.

99. Joudrier, P., M.F. Gautier, F. de Lamotte, and K. Kobrehel, *The thioredoxin h system: potential applications.* Biotechnol. Adv., 2005. **23**(1): p. 81-85.

100. Tavare, S., *Some probabilistic and statistical problems in the analysis of DNA sequences.* Lectures Math. Life Sci., 1986. **17**: p. 57-86.

101. Lio, P. and N. Goldman, *Models of molecular evolution and phylogeny.* Genome Res., 1998. **8**(12): p. 1233-1244.

102. Yang, Z., *PAML 4: phylogenetic analysis by maximum likelihood.* Mol. Biol. Evol., 2007. **24**(8): p. 1586-1591.

103. Jukes, T.H. and C.R. Cantor, *Evolution of Protein Molecules*. 1969, New York: Academic Press.

104. Arnold, K., L. Bordoli, J. Kopp, and T. Schwede, *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.* Bioinformatics, 2006. **22**(2): p. 195-201.

105. Schrodinger, LLC, *The PyMOL Molecular Graphics System, Version 1.3r1*, 2010.

106. London, M. and P.B. Hudson, *Uricolytic activity of purified uricase in two human beings.* Science, 1957. **125**(3254): p. 937-938.

107. Conley, T.G. and D.G. Priest, *Purification of uricase from mammalian tissue.* Prep. Biochem., 1979. **9**(2): p. 197-203.

108. Delsuc, F., G. Tsagkogeorga, N. Lartillot, and H. Philippe, *Additional molecular support for the new chordate phylogeny.* Genesis, 2008. **46**(11): p. 592-604.

109. Hedges, S.B., J. Dudley, and S. Kumar, *TimeTree: a public knowledge-base of divergence times among organisms.* Bioinformatics, 2006. **22**(23): p. 2971-2972.

110. Kelly, S.J., M. Delnomdedieu, M.I. Oliverio, L.D. Williams, M.G.P. Saifer, M.R. Sherman, T.M. Coffman, G.A. Johnson, and M.S. Hershfield, *Diabetes insipidus in uricase-deficient mice: A model for evaluating therapy with poly(ethylene glycol)-modified uricase.* J. Am. Soc. Nephrol., 2001. **12**(5): p. 1001-1009.

111. Koyama, Y. and T. Ichikawa. Mutant uricase, a mutant uricase gene, a novel recombinant DNA, and a process for producing mutant uricase. US 5376545 US 5700674, USPTO application. 1997.

112. Yoshiaki, N., T. Astsushi, K. Takahide, and H. Takao. Method for improving stability of uricase and modified uricase having improved stability. 198289 JPO application. 2006.

113. Reumers, J., F. Rousseau, and J. Schymkowitz, *Multiple evolutionary mechanisms reduce protein aggregation.* Open Biol., 2009(2): p. 176-184.

114. Weber, C.A., P.J. Mehta, M. Ardito, L. Moise, B. Martin, and A.S. De Groot, *T cell epitope: friend or foe? Immunogenicity of biologics in context.* Adv. Drug Delivery Rev., 2009. **61**(11): p. 965-976.

115. De Groot, A.S. and D.W. Scott, *Immunogenicity of protein therapeutics.* Trends Immunol., 2007. **28**(11): p. 482-490.

116. Sherman, M.R., M.G.P. Saifer, L.D. Williams, M.S. Hershfield, and S.J. Kelly. Aggregate-free urate oxidase for preparation of non-immunogenic polymer conjugates. 7927852 USPTO application. 2011.

117. Zhang, C., K. Fan, X. Ma, and D. Wei, *Impact of large aggregated uricases and PEG diol on accelerated blood clearance of PEGylated canine uricase.* PLoS One, 2012. **7**(6).

118. Fraczkiewicz, R. and W. Braun, *Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules.* J. Comput. Chem., 1997. **19**(3): p. 319-333.

119. Zalipsky, S. and J.M. Harris. *Introduction to chemistry and biological applications of poly(ethylene glycol)*. in *Poly(ethylene glycol) Chemistry and Biological Applications: ACS Symposium Series 680*. 1997. Washington, DC: American Chemical Society.

120. Roberts, M.J., M.D. Bentley, and J.M. Harris, *Chemistry for peptide and protein PEGylation.* Adv. Drug Delivery Rev., 2002. **54**(4): p. 459-476.

121. Sherman, M.R., D.W. Williams, M.G. Saifer, J.A. French, L.W. Kwak, and J.J. Oppenheim. *Conjugation of high-molecular weight poly(ethylene glycol) to cytokines: granulocyte-macrophage colony-stimulating factors as model substrates*. in *Poly(ethylene glycol) Chemistry and Biological Applications*. 1997. San Francisco, CA: American Chemical Society.

122. Conan, J.F. and M.V.A. James, *PEG-proteins: Reaction engineering and separation issues.* Chem. Eng. Sci., 2006. **61**.

123. Veronese, F.M. and G. Pasut, *PEGylation, successful approach to drug delivery.* Drug Discov. Today, 2005. **10**(21): p. 1451-1458.

124. Veronese, F.M. and A. Mero, *The impact of PEGylation on biological therapies.* Biodrugs, 2008. **22**(5): p. 315-329.

125. Fee, C.J. and J.A. Van Alstine, *PEG-proteins: Reaction engineering and separation issues.* Chem. Eng. Sci., 2006. **61**(3): p. 924-939.

126. Sezer, A.D. and A.F. Yagci, *Overview of peptide and protein PEGylation: properties and general strategies.* Acta Pharm. (Zagreb, Croatia), 2010. **52**: p. 377-389.

127. Kunitani, M., G. Dollinger, D. Johnson, and L. Kresin, *On-line characterization of polyethylene glycol-modified proteins* J. Chromatogr., 1991(588): p. 125-137.

128. Fee, C.J. and J.M. Van Alstine, *Prediction of the viscosity radius and the size exclusion chromatography behavior of PEGylated proteins.* Bioconjugate Chem., 2004. **15**(6): p. 1304-1313.

129. Williams, D.L., M.S. Hershfield, S.J. Kelly, M.G.P. Saifer, and M.R. Sherman. PEG-urate oxidase conjugates and use thereof. 8067553 USPTO application. 2011.

130. Pasut, G. and F. Veronese, *State of the art in PEGylation: the great versatility achieved after forty years of research.* J. Controlled Release, 2012. **161**(2): p. 461-472.

131. Chen, R.H., A. Abuchowski, T. Van Es, N.C. Palczuk, and F.F. Davis, *Properties of two urate oxidases modified by the covalent attachment of poly(ethylene glycol)*. Biochim. Biophys. Acta, 1981. **660**(2): p. 293-298.

132. Hartman, J. and S. Mendelovitz. Variant forms of urate oxidase and use thereof. 8188224 2012.

133. Williams, D.L., M.S. Hershfield, S.J. Kelly, M.G.P. Saifer, and M.R. Sherman. PEG-urate oxidase conjugates and use thereof. 6576235 2003.

134. Sherman, M.R., M.G.P. Saifer, D.L. Williams, M.S. Hershfield, and S.J. Kelly. Aggregate-free urate oxidase for preparation of non-immunogenic polymer conjugates. 7927852 2011.

135. Hartman, J.H., IL), Mendelovitz, Simona (Ramat Aviv, IL, US), Rehrig, Claudia D. (Plainsboro, NJ, US), Huang, William (Florham Park, NJ, US), Hershfield, Michael (Durham, NC, US). Methods for lowering elevated uric acid levels using intravenous injections of PEG-uricase. 8148123 2012.

136. Bomalaski, J.S., D.H. Goddard, D. Grezlak, M.A. Lopatin, F.W. Holtsberg, C.M. Ensor, and M.A. Clark, *Phase I study of uricase formulated with polyethylene glycol 3 (Uricase-PEG 20)*. Arthritis Rheum., 2002. **46**(9): p. S141-S141.

137. Zhang, C., K. Fan, H. Luo, X. Ma, R. Liu, L. Yang, C. Hu, Z. Chen, Z. Min, and D. Wei, *Characterization, efficacy, pharmacokinetics, and biodistribution of 5kDa mPEG modified tetrameric canine uricase variant.* Int J Pharm, 2012. **430**(1-2): p. 307-317.

138. Zhang, C., K. Fan, W. Zhang, R. Zhu, L. Zhang, and D. Wei, *Structure-based characterization of canine-human chimeric uricases and its evolutionary implications.* Biochimie, 2012. **94**(6): p. 1412-1420.

139. Morpurgo, M. and F.M. Veronese, *Conjugates of peptides and proteins to polyethylene glycols*, in *Bioconjugation Protocols: Strategies and Methods*, C.M. Niemeyer, Editor. 2004, Humana Press, Inc: Totowa, NJ. p. 45-70.

140. Mero, A., C. Clementi, F.M. Veronese, and G. Pasut, *Covalent conjugation of poly(ehtylene glycol) to proteins and peptides: Strategies and methods*, in *Bioconjugation protocols: Strategies and Meethods*, S.S. Mark, Editor. 2011, Springer Science and Business Media. p. 95-129.

141. Zheng, C., C. Zheng, G. Ma, and Z. Su, *Native PAGE eliminates the problem of PEG-SDS interaction in SDS-PAGE and provides an alternative to HPLC in characterization of protein PEGylation.* Electrophoresis, 2007. **28**(16): p. 2801-2807.

142. Artimo, P., M. Jonnalagedda, K. Arnold, D. Baratin, G. Csardi, E. de Castro, S. Duvaud, V. Flegel, A. Fortier, E. Gasteiger, et al., *ExPASy: SIB bioinformatics resource portal.* Nucleic Acids Res., 2012. **40**: p. 603.

143. Zhang, C., K. Fan, X.F. Ma, and D.Z. Wei, *Impact of Large Aggregated Uricases and PEG Diol on Accelerated Blood Clearance of PEGylated Canine Uricase.* PLoS One, 2012. **7**(6): p. 11.

144. Sherman, M.R., M.G. Saifer, and F. Perez-Ruiz, *PEG-uricase in the management of treatment-resistant gout and hyperuricemia.* Adv. Drug Delivery Rev., 2008. **60**(1): p. 59-68.

145. Alconcel, S.N.S., A.S. Baas, and H.D. Maynard, *FDA-approved poly(ethylene glycol)-protein conjugate drugs.* Polym. Chem., 2011. **2**: p. 1442-1448.

146. Armstrong, J.K., G. Hempel, S. Koling, L.S. Chan, T. Fisher, H.J. Meiselman, and G. Garratty, *Antibody against poly(ethylene glycol) adversely affects PEG-asparaginase therapy in acute lymphoblastic leukemia patients.* Cancer, 2007. **110**(1): p. 103-111.

147. Garay, R.P., R. El-Gewely, J.K. Armstrong, G. Garratty, and P. Richette, *Antibodies against polyethylene glycol in healthy subjects and in patients treated with PEG-conjugated agents.* Expert Opin. Drug Deliv., 2012. **9**(11): p. 1319-1323.

148. Sherman, M.R., L.D. Williams, M.A. Sobczyk, S.J. Michaels, and M.G. Saifer, *Role of the methoxy group in immune responses to mPEG-protein conjugates.* Bioconjug. Chem., 2012. **23**(3): p. 485-499.

149. Tsuji, J., K. Hirose, E. Kasahara, M. Naitoh, and I. Yamamoto, *Studies on antigenicity of the polyethylene glycol (PEG)-modified uricase.* Int. J. Immunopharmacol., 1985. **7**(5): p. 725-730.

150. Caliceti, P., O. Schiavon, and F.M. Veronese, *Biopharmaceutical properties of uricase conjugated to neutral and amphiphilic polymers.* Bioconjug. Chem., 1999. **10**(4): p. 638-646.

151. Caliceti, P., O. Schiavon, and F.M. Veronese, *Immunological properties of uricase conjugated to neutral soluble polymers.* Bioconjug. Chem., 2001. **12**(4): p. 515-522.

152. Stevenson, W.S., C.D. Hyland, J.G. Zhang, P.O. Morgan, T.A. Willson, A. Gill, A.A. Hilton, E.M. Viney, M. Bahlo, S.L. Masters, et al., *Deficiency of 5-hydroxyisourate hydrolase causes hepatomegaly and hepatocellular carcinoma in mice.* Proc. Natl. Acad. Sci. U. S. A., 2010. **107**(38): p. 16625-16630.

VITA

James T. Kratzer

James was born in Sacramento CA. He attended Brookwood High School in Snellville, Georgia and received a B.S. in Biochemistry from Berry College in Rome, Georgia in 2006. He came to the Georgia Institute of Technology in 2006 to pursue a doctorate in Biochemistry where he has been shown the wonderful world of evolutionary synthetic biology. He had been lucky in pursuing his passion with translational research. Outside of his scientific endeavors, Mr. Kratzer enjoys spending time with his wife, Stacia and the rest of his family.