Protein Engineering Through *in vivo* Incorporation of Phenylalanine Analogs

Thesis by
Isaac Sheridan Carrico

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Chemistry

California Institute of Technology, Pasadena, CA 91125
2004
(defended September 2, 2003)

For Lizzy and my family

ACKNOWLEDGMENTS

My time here at Caltech has been very interesting and rewarding and it will be hard for me to capture my gratitude for those who have shared my life here. I think I will start with my last two advisors, David Tirrell and Barbara Imperiali. Incredibly different, but both wonderful people who were crucial for my development. Barbara brought an immense amount of energy and personality to the lab, but perhaps more importantly she immediately treated me as an equal, someone of value. Dave makes a good run at being the perfect advisor. The inability to convincingly complain about him is the biggest fault that I have been able to find. Dave's breadth of knowledge is staggering, but it is the patience he displays in attempting to impart this knowledge that is truly impressive. My circuitous path through the labs of Caltech has caused most of my friends to become labmates at one point or another, so I will forgo the traditional format. Soojin Kwon, Mike Farwell, Doan Nguyen, Jason Belitsky, Peter Hackley, Carlos Bosques, Rob Dempski and Niki Zacharias comprise the classmates whom I met upon arriving at Caltech and hope never to grow far apart from. I want to thank Akif Tezcan and Derrick Debe, housemates and teammates who had a wonderful enthusiasm for science and perhaps only surpassed by their eagerness for cutting loose. Gabriel Brandt provided a quiet, calming, centering force, both on the field and in life. I also need to thank all of the wonderful people with whom I shared the fields with at Caltech; I lament that I may never find such a friendly atmosphere for sweating away the frustrations of life. I must thank Pin Wang and Kent Kirshenbaum for their positive outlook and perceptive manners that always kept us moving in the right direction. The rest of the Tirrell group has been wonderful to

work with; particularly I need express my gratitude towards David Flanagan for his lack of pretension and sharp wit, Yi Tang for the friendly aggressiveness, Jill Sakata for her generosity of spirit, Marissa Mock always a positive correcting force, and Sarah Heilshorn who was always willing to help.

My family cannot be left out. My parents have always been unwavering in their support and love. I will never forget when, upon arriving at a new school, my mother demanded that I be put in the most advanced classes, her love for her boys has always been so strong as to preclude reason. My father also provided a continuous source of encouragement and love, thrilled that I would choose to go into the sciences for which he is probably better suited. Trey's encouragements started when I was a little boy in Austin and have never faltered. Zach who is following in science, but will without a doubt soon pass me. I also want to thank the new additions to my family, the Boons, who have welcomed me into their lives and were always kind about asking "so…when are you going to finish?" Lastly, but most importantly, I need to thank my wife who is my best friend and made every day at Caltech a joy. She is my constant companion and keeps my feet moving in the right direction. Without her I would be truly lost.

ABSTRACT

Proteins mediate the bulk of biochemical functions within the cell. These biopolymers control processes utilizing specific arrangements of the natural twenty amino acids. Expanding the set of amino acids available could both aid in the study of these macromolecules as well as significantly increase their functional capabilities. A set of enzymes known as aminoacyl tRNA-synthetases lies at the heart of the fidelity of translation, the process by which genetic information gets decoded into proteins. These synthetases accurately charge a specific tRNA with its cognate amino acid in the presence of the other nineteen natural amino acids. Interestingly these enzymes demonstrate a much higher level of promiscuity with unnatural amino acids. However, acceptable amino acids are limited to those that bear steric and electronic resemblance to the natural analog.

Our efforts to expand the substrate set of phenylalanyl-tRNA (PheRS) synthetase are described in Chapters 2-4. We redesigned the catalytic site of PheRS computationally. These results combined with an already known mutant allowed us to rationally create a third mutant. All three mutants were characterized for their ability to activate a large panel of unnatural amino acids *in vitro*. Further, we were able to confirm the *in vivo* incorporation of a number of these analogs. *In vitro* and *in vivo* results were consistent and defined an expanded substrate set for the described mutants. This substrate set includes a number of analogs that are dramatically different from phenylalanine both sterically and electronically, as well as a number which contain chemical moieties valuable to protein engineering efforts.

One example is *para*-azidophenylalanine (pN_3Phe), which provides access to photochemistry as well as modified Staudinger ligations and copper mediated electrocyclizations. In Chapter 5 we describe utilization pN_3Phe as a photocrosslinking reagent. Our aim was to create photochemically crosslinkable artificial extracellular matrix proteins for the production of synthetic vascular grafts. These proteins, produced in *E. coli*, were diblocks of endothelial cell binding domains and structural domains including the pN_3Phe site. Photochemical crosslinking of these constructs provided moduli well within the range presented by the natural vascular wall. Chapter 6 describes our ability to photopattern films composed of the above protein. Photopatterning provided a means to spatially array endothelial cells, based upon a number of controllable processing parameters of such films.

The final chapter details the utilization of incorporated unnatural amino acids, particularly *para*-iodophenylalanine, *para*-acetylphenylalanine and homopropargylglycine, to access Pd(0) catalyzed cross-coupling chemistry. We demonstrated this chemistry exhibits the characteristics necessary for chemoselective ligations. Futher, we demonstrated the selective modification of proteins incorporating all of the above analogs.

TABLE OF CONTENTS

CHAPTER 1. Unnatural Amino Acids in Biomaterials and Protein Engineering

Introduction 2

References 16

CHAPTER 2. Biosynthesis of Proteins Incorporating a Versatile Set of Phenylalanine Analogs

Introduction 24

Materials and Methods 26

Results and Discussion 29

PAGE analysis of the effects PheRS* on analog incorporation into mDHFR 29

Quantitative analysis of analog incorporation by amino acid analysis 29

Confirmation of analog incorporation by MALDI-TOF analysis 29

Large-scale expression of mDHFR containing unnatural analogs 32

Analysis of mDFHR containing analogs demonstrates new UV signatures 32

Conclusion 36

References 37

CHAPTER 3. A Designed Phenylalanyl-tRNA Synthetase Variant Allows Efficient *in vivo* Incorporation of Aryl Ketone Functionality into Proteins

Introduction 42

Materials and Methods 45

Results and Discussion 51

Mutational predictions based on ORBIT calculations 51

PAGE analysis of mutant synthetase effects on incorporation of **2** into mDHFR 51

MALDI-TOF analysis of incorporation of **2** 55

Selective modification of DHFR-**2** with biotin hydrazide 55
Conclusion 59
References 60

CHAPTER 4. Engineering Relaxed Substrate Specificity into *E. coli* Phenylalanyl-tRNA Synthetase to Incorporate a Diverse Set of Non-natural Amino Acids
Introduction 64
Materials and Methods 69
Results and Discussion 74
Expression and purification of PheRS variants 74
Activation of analogs by variant enzymes *in vitro* 74
In vivo evaluation by DHFR tryptic peptide analysis 78
Conclusion 85
References 86

CHAPTER 5. Efficient Photocrosslinking of an Artificial Extracellular Matrix Protein via *in vivo* Incorporation of Arylazide Functionality
Introduction 93
Materials and Methods 97
Results and Discussion 100
Protein expression and purification 100
Infrared spectroscopy of CS5-ELF-N_3 100
Film production 105
Mechanical testing 105
Conclusion 109
References 76

CHAPTER 6. Patterning and Cell Binding Properties of a Protein Photoresist Produced in *E. coli*
Introduction 115

Materials and Methods 119
Results and Discussion 124
Phase contrast imaging of protein patterns 124
Infrared detection of azide photolysis kinetics 126
Cell attachment to CS5-ELF-N3 constructs 126
Cell patterning based on cell preference to non-irradiated CS5-ELF-N3 surfaces 130
Cell patterning on stripped surfaces 130
Conclusion 134
References 136

CHAPTER 7. Chemoselective Ligations via Pd(0) Chemistry on Unnatural Amino Acids Incorporated into Proteins
Introduction 144
Materials and Methods 149
Results and Discussion 155
Optimization of Heck and Sonagashira couplings in aqueous conditions 155
Demonstration of tolerance to protein functionality 157
Modification of mDHFR (pCCHPhe, HAG and HPG) with pIF-FLAG tag 157
Selective fluorescent modification of Barstar-pIF with lissamine rhodamine propargylsulfonamide 158
Conclusion 165
References 166

LIST OF FIGURES AND TABLES

CHAPTER 1. Unnatural Amino Acids in Biomaterials and Protein Engineering

1.1. Simplified schematic overview of transcription and translation 3
1.2. Schematic representation of *in vitro* amber suppression 5
1.3. Schematic overview of *in vivo* amber suppression technology used with *Xenopus* oocytes 6
1.4. Two-step aminoacylation of tRNA catalyzed by aaRS 7
1.5. Ribbon representation of the portion of catalytic □-subunit of PheRS from *T. thermophilus* 9
1.6. Stabilization of coiled coil peptides as a result of introduction of fluorinated leucine analogs 11
1.7. ß-lamellar structure exhibits surface properties defined by exposed amino acid 12
1.8. Subset of analogs incorporated *in vivo* via "media shift" method 13

CHAPTER 2. Biosynthesis of Proteins Incorporating a Versatile Set of Phenylalanine Analogs

2.1. Analogs used to probe the fidelity of the A294G mutant of PheRS 25
2.2. SDS-PAGE of cell lysates of AF-IQ[pQE-15] and AF-IQ[pQE-FS] 30
Table 2.1. Extent of substitution of phe by analogs **2-9** in DHFR 31
2.3. MALDI-TOF mass spectra of tryptic peptides derived from DHFR expressed in media supplemented with phe (a) or analog **4** (b) 33
Table 2.2. MALDI-TOF data for tryptic peptides derived from DHFR expressed in media supplemented with phe or analogs **2-9** 34
2.4. UV spectra of purified DHFR expressed in media supplemented with phe or with one of the analogs **2-5** or **7** 35

CHAPTER 3. A Designed Phenylalanyl-tRNA Synthetase Variant Allows Efficient *in vivo* Incorporation of Aryl Ketone Functionality into Proteins

3.1. Ribbon representation of the portion of catalytic β-subunit of PheRS from *T. thermophilus* 44

Table 3.1. ORBIT Calculation for *para*-acetylphenylalanine (**2**) binding into *T. Thermophilus* PheRS 52

3.2. Active site of *T. Thermophilus* PheRS/Phe 53

3.3. SDS-PAGE of cell lysates of 4 hr post-induction with 1 mM IPTG 54

3.4. MALDI-TOF mass spectra of tryptic peptides derived from DHFR 56

3.5. Western blot showing chemoselective modification of ketone functionality in mDHFR 58

CHAPTER 4. Engineering Relaxed Substrate Specificity into *E. coli* Phenylalanyl-tRNA Synthetase to Incorporate a Diverse Set of Non-natural Amino Acids

4.1. Ribbon representation of the portion of catalytic β-subunit of PheRS from *T. thermophilus* 66

4.2. Chemical structures of the amino acids involved in this study 68

Table 4.1. ATP-PPi exchange kinetics 75

Table 4.2. ATP-PPi exchange kinetics 76

4.3. Mass data for peptide Fragment 1 79

4.4. Amino acid sequence of target protein mDHFR 80

4.5. MALDI-MS of tryptic peptide fragment 2 81

CHAPTER 5. Efficient Photocrosslinking of an Artificial Extracellular Matrix Protein via *in vivo* Incorporation of Arylazide Functionality

5.1 Extracellular matrix construct 94

5.2 Photodecomposition of arylazide 96

5.3 Tris-Tricine SDS-PAGE (9%) analysis of expression results 101

Table 5.1. Percent of phenylalanine replaced by pN_3Phe as a function of pN_3Phe in the growth medium 102

5.4 Expanded 600MHz ^1H NMR of CS5-ELF-F and CS5-ELF-N_3 in DMSO-d_6 103

5.5 Full 600 MHz ^1H NMR of CS5-ELF-N_3 (53%) and CS5-ELF-F 104

5.6 Transmission and ATR infrared studies of azide decomposition 106
5.7 Representative stress vs. strain curves for photocrosslinked CS5-ELF-N_3 films 108
Table 5.2. Shear and elastic modulus, MW between crosslinks and photocrosslinking efficiency as a function of percent Phe replacement by pN_3Phe 108

CHAPTER 6. Patterning and Cell Binding Properties of a Protein Photoresist Produced in *E. coli*

6.1 Schematic Illustration of negative photoresist 116
6.2 Photodecomposition of arylazide 118
6.3 Extracellular matrix construct 120
6.4 Phase contrast images of photopatterned CS5-ELF-N_3 (53%) 125
6.5 Photolysis of incorporated pN_3Phe followed by FT-IR 127
6.6 Representative phase contrast images of HUVECs 2 hours after plated on CS5-ELF-N_3 containing different levels of azide 128
6.7 Representative phase contrast images of HUVECs on CS5-ELF-N_3 (53%) 129
6.8 Representative phase contrast images of HUVECs plated upon photopatterned CS5-ELF-N_3 (53%) 131
6.9 HUVEC patterning 6 hours after seeding on "protein resist" CS5-ELF-N_3 (53%) patterned on polyethylene glycol coated slides 132
6.10 Fluorescent HUVEC array images 133

CHAPTER 7. Chemoselective ligations via Pd(0) chemistry on unnatural amino acids incorporated into proteins

7.1 Chemoselective ligation reactions 145
7.2 Simplified Sonagashira and Heck catalytic cycles 147
7.3 Partially optimizing conditions for Heck and Sonagashira couplings 156
7.4 Pd (0) mediated modification of DHFR with pIF-FLAG 159

7.5 Crystal structure of Barstar demonstrated as stick model 161
7.6 Schematic depicting the selective modification of Barstar-pIF with lissamine rhodamine propargyl sulfonamide 162
7.7 Modification of Barstar with lissamine rhodamine propargylsulfonamide 163
7.8 Selective modification of Barstar with lissamine rhodamine propargyl sulfonamide 164