# PROTEIN MODIFICATION THROUGH *IN VIVO* INCORPORATION OF NONCANONICAL AMINO ACIDS

Thesis by

Marissa Lee Mock

In Partial Fulfillment of the Requirements for the

degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2005

(Defended October 21, 2005)

© 2005

Marissa Mock

All Rights Reserved

# ACKNOWLEDGEMENTS

I wish to thank my advisor, Professor David Tirrell, for his advice and support. I especially appreciate the laboratory atmosphere of intellectual independence with friendly collaboration that he fostered. I also wish to thank every member of the Tirrell Lab, past and present, who made my time here such a pleasure. In particular, this research could not have been accomplished without the support, both intellectual and personal, of Sarah Heilshorn, David Flanagan, Paul Nowatzki, Julie Liu, Kimberly Beatty, Kristi Kiick, Isaac Carrico, Kathy DiZio, and Rebecca Connor.

I thank my committee, Professors Dennis Dougherty, Carl Parker, and Christina Smolke for their advice and assistance.

I cannot adequately convey here how grateful I am to Sarah Heilshorn and Andrew Spakowitz, my roommates for four years, for simply everything.

I earnestly thank Lars Cremean for being there for me.

Finally, my deepest gratitude goes to my family, especially my parents, who never let me imagine that there was anything I couldn't do if I tried. They were willing to sacrifice to give me everything, and any credit for an accomplishment of mine is entirely theirs.

# ABSTRACT

Traditional techniques of polymer synthesis produce macromolecules with statistical distributions of chain length, composition, stereochemistry, and sequence. Nature has evolved a complex system for polypeptide synthesis that gives essentially complete control of chain length and monomer sequence. Using the natural protein biosynthesis machinery to produce protein polymers provides not only a unique opportunity to study the effects of such molecular characteristics on material properties, but also the possibility of readily incorporating bioactive domains into protein-based materials.

The objective of this thesis work was to expand upon the set of amino acids available for incorporation into proteins *in vivo* and to explore applications of the novel chemistries and physical properties provided by the new analogs.

Chapter 2 describes the incorporation of new unsaturated analogues of isoleucine, the alkene 2-amino-3-methyl-4-pentenoic acid and the alkyne 2-amino-3-methyl-4-pentynoic acid, by the wild type *E. coli* biosynthetic apparatus. Incorporation was found to be sensitive to side chain stereochemistry in the case of the alkene analog; the translational activity of the pairs of enantiomers (SS, RR and SR, RS) were markedly different. We concluded that, although the SS-isomer is a good analogue, the SR-isomer is not incorporated into proteins by this expression host.

Chapter 3 focuses on the incorporation of a fluorine-containing noncanonical amino acid, 5,5,5-trifluoroisoleucine, into artificial extracellular matrix proteins. The fluorinated proteins displayed altered solubility phase behavior and were more resistant to degradation by the physiologically relevant protease elastase, yet retained the ability to adhere endothelial cells in a sequence specific manner.

Chapter 4 describes the incorporation of the photoreactive noncanonical analog *p*-azidophenylalanine into artificial extracellular matrix proteins. Films of the azide-containing proteins were crosslinked upon short exposure to ultraviolet radiation. Using simple patterned masks, we demonstrated the ability to pattern protein films by only exposing certain regions. When protein patters were produced on a non-adhesive background, endothelial cells selectively adhered to the protein regions to create stable cell patterns.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ABBREVIATIONS

| | |
|---|---|
| **2xYT** | two times yeast/tryptone medium |
| **5TFI** | 5,5,5-trifluoroisoleucine |
| **aaRS** | aminoacyl-tRNA synthetase (*) |
| **aECM** | artificial extracellular matrix protein |
| **ATP-PP$_i$** | adenosine triphosphate-pyrophosphate |
| **βME** | β-mercaptoethanol |
| **BSA** | bovine serum albumin |
| **CAI** | cell adhesion index |
| **Cbz** | *N*-benzyloxycarbonyl |
| **CD** | circular dichroism |
| **Cy2** | cyanine dye 2 |
| **DAPI** | 4′, 6 diamindine-2-phenyl indole |
| **DMSO** | dimethylsulfoxide |
| **dpi** | dots per inch |
| **DPN** | dip-pen nanolithography |
| **E** | elastic modulus |
| **ECM** | extracellular matrix |
| **EDTA** | ethylenediaminetetraacetic acid |
| **E-Ile** | 2-amino-3-methyl-4-pentenoic acid |
| ***E. coli*** | *Escherichia coli* |
| **FTIR** | Fourier transform infrared spectroscopy |
| **HLE** | human leukocyte elastase |
| **$^1$H-NMR** | proton nuclear magnetic resonance (spectroscopy) |
| **HUVEC** | human umbilical vein endothelial cells |
| **IleRS** | isoleucyl-tRNA synthetase |
| **IPTG** | isopropyl-β-D-thiogalactopyranoside |
| **L-allo-Ile** | (2S, 3R) 2-amino-3-methyl-4-pentanoic acid |

**LCST** lower critical solution temperature

**M9 or M9AA** minimal medium

**MALDI-MS** matrix-assisted laser desorption ionization-mass spectrometry

**mDHFR** murine dihydrofolate reductase

**NEB** New England Biolabs

**OD** optical density

**PBS** phosphate-buffered saline

**PEG** poly(ethylene glycol)

**PheRS** phenylalanyl-tRNA synthetase

**PMSF** phenylmethylsulfonyl fluoride

***p*N$_3$Phe** *para*-azidophenylalanine

**PP$_i$** sodium pyrophosphate

**SDS-PAGE** sodium dodecyl sulfate - polyacrylamide gel electrophoresis

**Y-Ile** 2-amino-3-methyl-4-pentynoic acid

**Common abbreviations for the twenty canonical amino acids**

| amino acid | 3-letter | 1-letter | amino acid | 3-letter | 1-letter |
|---|---|---|---|---|---|
| alanine | Ala | A | leucine | Leu | L |
| arginine | Arg | R | lysine | Lys | K |
| aspartic acid | Asp | D | methionine | Met | M |
| asparagine | Asn | N | phenylalanine | Phe | F |
| cysteine | Cys | C | proline | Pro | P |
| glutamic acid | Glu | E | serine | Ser | S |
| glutamine | Gln | Q | threonine | Thr | T |
| glycine | Gly | G | tryptophan | Trp | W |
| histidine | His | H | tyrosine | Tyr | Y |
| isoleucine | Ile | I | valine | Val | V |

**The abbreviation for an aminoacyl-tRNA synthetase specific to an amino acid is formed by placing the appropriate three-letter amino acid abbreviation before the letters RS. For example, the isoleucyl-tRNA synthetase is abbreviated IleRS.*