AZIDE-BEARING AMINO ACIDS IN PROTEIN ENGINEERING AND PROTEOMIC PROFILING

Thesis by

Aaron James Link

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2006

(Defended September 9, 2005)

Acknowledgements

I first wish to acknowledge my advisor Dave Tirrell. Dave has allowed me great freedom in developing projects to work on in his lab and has always been supportive of the collaborations that led to much of my work. The diverse, interdisciplinary nature of Dave's research program is one of the things that drew me to his group in the first place, and I have learned much by being in such an environment. First and foremost among collaborators are Daniela Dieterich and Erin Schuman, with whom we collaborated on the work described in Chapter 5 of this thesis. Daniela has been a joy to work with, and I look forward to many future interactions with her. Tim Chan, MG Finn, Valery Fokin, and Barry Sharpless of the Scripps Research Institute provided materials and intellectual capital that allowed us to incorporate click chemistry into our experiments. Nick Agard, Jenn Prescher, and Carolyn Bertozzi from UC Berkeley also provided reagents that proved to be the key to the success of the work described in Chapter 4 of this thesis. Paul Bessette and Patrick Daugherty of UC Santa Barbara provided excellent advice on library construction and flow cytometry.

The Tirrell group has been an excellent place to grow as a scientist. Yi Tang, Pin Wang, and Isaac Carrico were always willing to help me in the lab in my first few years, and the scientific discussions we had shaped the way I read and evaluate the literature to this day. I have enjoyed working closely with Mandy Vink, Rebecca Connor, and high-schooler Drew Klein, as well as having many discussions with Jin Montclare, Mike Diehl, Inchan Kwon, Kechun Zhang, Tae Hyeon Yoo, and Jim Van Deventer. One advantage to the location of Dave's lab is our proximity to Frances Arnold's group, and

I've had many fruitful interactions with members of that group past and present including Pat Cirino, Joff Silberg, Alex Tobias, and Allan Drummond.

I would also like to thank the members of my committee, Frances Arnold, Steve Mayo, and Shu-ou Shan for their support and willingness to read and evaluate my work. I also gratefully thank Dave, Frances, and Barry for their help with my job search.

Finally, I wish to thank my parents Richard and Melody, my brother Jason, and my sisters Annie and Abby for their constant encouragement as I worked on my PhD. Despite my decision to move all the way across the country for graduate school, they have been supportive throughout the entire process.

Abstract

The utility of non-canonical amino acids in protein engineering has grown substantially over the past decade. Proteins containing these unnatural building blocks often have radically different biochemical or spectral characteristics than their wild-type counterparts. Furthermore, proteins may be endowed with chemical reactivity not found in the natural proteome upon the introduction of non-canonical amino acids. Successful incorporation of a non-canonical amino acid into recombinant proteins in *E. coli* is often dependent on engineering of the aminoacyl-tRNA synthetase (aaRS) activity of the cell. The bulk of the work described herein has focused on developing a system to rapidly screen libraries of mutant aaRS to identify clones capable of efficiently incorporating novel reactive non-canonical amino acids. The system is based on the display of reactive amino acid side chains on the surface of *E. coli* cells upon metabolic incorporation of the amino acid into recombinant outer membrane protein C (OmpC) and the subsequent covalent biotinylation of the reactive side chains. The cells are then stained with fluorescent avidin, thus rendering the cells incorporating the amino acid fluorescent and readily identifiable and sortable by flow cytometry.

The feasibility of such a system was proven by incorporating the methionine surrogate azidohomoalanine (AHA) into OmpC and subsequently biotinylating the reactive azide groups via copper-catalyzed azide-alkyne ligation. Using an improved copper catalyst, low levels of incorporation of translationally inefficient amino acids azidoalanine, azidonorvaline, and azidonorleucine into OmpC were also detected. A saturation mutagenesis library of the methionyl-tRNA synthetase (MetRS) was designed, and cells transformed with this library were screened for the ability to incorporate the

long chain amino acid azidonorleucine into recombinant proteins efficiently. Several MetRS mutants were identified with such activity using the cell surface display system. MetRS containing a single amino acid mutation, leucine 13 to glycine (L13G) that occurs in each of the three mutants discovered in the screen, is very efficient at incorporating azidonorleucine into proteins.

In the last part of the work described in this thesis, azidohomoalanine was used to tag newly-synthesized proteins in mammalian cells, thus endowing the newly-synthesized proteins with unique bioorthogonal chemical reactivity. Following covalent biotinylation via the azide-alkyne ligation, these proteins could be selectively enriched for by avidin chromatography and identified using shotgun proteomic approaches. Nearly 200 newly-synthesized proteins were identified unequivocally in just a two-hour window. This technique promises to develop into a highly useful tool for the examination of proteome dynamics.

Table of Contents

List of Tables **Page**

List of Figures and Schemes **Page**