Discovery of Aminoacyl-tRNA Synthetase Mutants for the Incorporation of Noncanonical Amino Acids into Proteins

Thesis by
İsmet Çağlar Tanrıkulu

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
2009

(Defended October 8, 2008)

Acknowledgements

As it is for many, graduate school has been a long, convoluted journey for me as well. Now that this part of my life is coming to a close, I realize more than ever the importance of so many people who helped me reach where I am right now.

First and foremost I would like to thank my “experimental” advisor, Prof. David Tirrell, for his guidance and support for the past six years. His patience in letting me carve my own research path, his willingness to learn about the details of my work, listen to the problems and discuss solutions, his genuine interest in understanding the principles behind the data on the screen all allowed me to learn a great deal on how to do science and how to have fun doing it. My “computational” advisor, Prof. Bill Goddard, was extremely patient with me during the many years of failed binding-site designs, and always had countless ideas for the next step. If it wasn’t for his encouragement for me to move on, I might have still been testing “dead” PheRS mutants until this day.

I would like to thank the current and past members of my thesis committee, Profs. Harry Gray, Linda Hsieh-Wilson, Rich Roberts, and Vaidehi Nagarajan. Our meetings occurred at critical times during my training. Their interest in my work and questions they posed helped me reevaluate my research path in new light and greatly influenced the direction of research I pursued. I very much appreciate their guidance.

Perhaps the most difficult period in my PhD was after my first candidacy exam, which ended less than satisfactorily for most involved. Prof. Judy Campbell, who was the BMB option head at the time, and Alison Ross, our option secretary, were both very supportive during this time and encouraged me in my quest to find a new research lab. My conversations with Prof. Doug Rees, the chair of my first committee, helped me put all events in perspective. Without their help I most likely would not be here today.

My first introduction to both molecular biology and computer programming happened while I was working in the Pierce and Mayo labs. I learned many fundamental molecular biology techniques from Rhonda Diguisto, who is a very

patient and organized teacher. Shannon Marshall introduced me to circular dichroism, and Possu Huang was helpful with anything that went wrong in the lab, which, as some may know, is the norm rather than the exception. As someone who did not have much programming experience I received a lot of help deciphering, modifying, and using the ORBIT code from Geoff Hom, Possu Huang, Dan Bolon, Robert Dirks, Shannon Marshall, Premal Shah, Deepshikha Datta, Eun Jung Choi, Jessica Mao, Kyle Lassila, and Marie Ary.

The Tirrell lab, where I completed the experimental part of my PhD work, has provided me with a wonderful environment, both scientifically and socially. I really doubt I will ever encounter such strong camaraderie again in the future. A diverse set of skills are required for success in the Tirrell lab and I got introduced to many of these through Marissa Mock, Sarah Heilshorn, Jin Montclare, and Jamie Link. However, I became confident with my lab skills only after receiving tremendous amounts of help from Inchan Kwon, with whom I shared both my lab bay and my office. In addition to letting me use his endless knowledge on all types of procedures, Inchan also taught me how to design and pursue experiments, how to judge the factors affecting them, and above all, how to be consistent. I deeply appreciate his help. I was also fortunate enough to have many people around me to troubleshoot problems, to exchange ideas, or just to have a talk with in the lab. Among these I have bothered Rebecca Connor, Tae Hyeon Yoo, Nick (J.D.) Fisk, John Ngo, Julie Champion, Ayae Sugawara, Makoto Ouchi, Soojin Son and Jin Montclare a countless times, and I thank them for their support and friendship. Jamie Link was the person who first developed the screening strategy I describe and use in Chapter 2, and his assistance was very valuable to me as I was starting up. Nick Fisk, Jamie Link, John Ngo and Mark van Eldijk allowed me to use reagents they had synthesized. Inchan Kwon and Tae Hyeon Yoo offered me advice and assistance on flow cytometry and library screening in general, and, with Marissa Mock, helped me with the $^{32}PP_i$-exchange assays. The MetRS-SLL x-ray crystal structure was obtained through a collaboration initiated by Tae Hyeon Yoo, with Yves Mechulam and Emmanuelle Schmitt and made many of the analyses presented here possible.

Finally I would like to thank John Ngo for turning the results of what was originally an intellectual exercise into an exciting tool for cell-specific metabolic labeling.

I performed the computational part of my PhD work in the Goddard lab, which contains an immense amount of expertise. The two people in the group who had the biggest impact on me were Vaidehi Nagarajan and Ravi Abrol. Vaidehi taught me the basics of all the computational work I present here. Her patience while working with me on the difficult problems and her encouragement kept me moving forward even when chances of success seemed bleak. I also learned a lot from Ravi, who has the amazing ability divide large problems into logical, manageable chunks. His guidance enabled me to obtain the results I now present in this work. I was lucky to have access to such great mentors. My work in the group involved the use of many newly introduced computational techniques, which required me to work closely with a number of people. David Zhang, who created the COP strategy for binding site design, helped me greatly as I was getting started in the group. Victor Kam's side chain placement procedure (SCREAM) was integral to my design strategy from the start, and he was always available for my questions and requests. My meetings with Pete Kekenes-Huskey were always intellectually stimulating, whether they were about rotamer generation or binding calculations. I really hope his torsional sampling code finds the audience it deserves. Jonas Oxgaard and Julius Su helped me out whenever my QM calculations got stuck. Andres Jaramillo-Botero was very understanding with my time constraints when I had to work on our ligand docking code and my thesis at the same time. Finally, it has always been very gratifying to see my code being used by the group, and therefore, I also would like to thank the users of various programs I have written through the years.

At Caltech, we are lucky to have an amazing staff to help us get through the day. Alison Ross (BMB), Carol Mann and Anne Horman (Tirrell lab), and Shirley Feng (Goddard lab) have worked with me through many scheduling and administrative problems that came up. Suresh Ghupta (ChE) and Darryl Willick (Goddard lab) have provided solutions to my computer related problems over and over and over again. Haick Issaian, with Karen Baumgartner and other

EHSS staff, answered countless questions on radiation use and helped me locate reagents and working instruments when I needed them. Mona Shahgholi trained me, helped me run MALDI and LC-MS measurements, and helped troubleshoot problems. In addition to research-related staff, I also would like to thank all the people on campus, from Jim Endrizzi and Athena Trentin (ISP) to Tess Legaspi (Registrar's office), from the health center staff to the pizza guys at Chandler, for keeping life at Caltech easy for us.

It would have been impossible for me to live through years of failed experiments if it wasn't for the love and support I received from my friends and family. In addition to ones I've already mentioned, the Turkish community at Caltech, especially Ozgun Konca, Sirin Caliskan, Can Aydiner, Baris Bingol, Arkadas Ozakin, Ercan Gurses, Omer Durak, Ersen Bilgin and Grace Song, and Muruvvet Buyukboyaci have been a great source of support for me, day in and day out. I thank Sarina Mohanty for constantly checking on me during the preparation of my thesis, and to Jose Navarette and Ozgun Konca for playing music with me from time to time. Although I am yet to determine whether she was a blessing or a curse, Eun Jung Choi was there for me during most of my stay at Caltech and deserves credit for her assistance and support.

My parents have always encouraged me to ask questions about anything and everything since I started talking, even if this would cause problems once in a while. They did everything for me to get a good education, and it seems to have worked. I have been away from home for the most part in the last two decades, and during this time they always made sure I knew they were there for me. In this past year, during the preparation of this thesis, I felt their presence even stronger. I am forever grateful for their love, care and support.

At last, I would like to thank Julie Champion, John Ngo, Ravi Abrol, Shelly Tzlil, Grace Song and Prof. David Tirrell for all the time they spent reviewing and editing various parts of this thesis.

Abstract

Efficient *in vivo* incorporation of a noncanonical amino acid into proteins often requires engineering new aminoacyl-tRNA synthetase (AARS) activity into the cell, usually by modifying a natural aaRS. Although experimental methods, relying on mutagenesis and library screening, have identified many successful mutant aaRS-substrate pairs in the recent years, computational approaches have reported only a few successes. Here we compare the results of computational and experimental screens of an *E. coli* methionyl-tRNA synthetase (MetRS) saturation-mutagenesis library for binding (*in silico*), and activation and cell-surface display (*in vivo*) of azidonorleucine (Anl).

Three positions (L13, Y260, and H301) in the methionine binding pocket of *E. coli* MetRS were randomized, and the resulting library was screened for MetRS activity toward Anl, based on a screening strategy previously established in our group. This strategy relies on the introduction of reactive side chains into surface-exposed sites on outer-membrane protein C (OmpC), and their subsequent labeling with reactive, fluorescent probes. We have discovered a large diversity of MetRS mutants that allow the incorporation of Anl into proteins *in vivo*. The extent of OmpC expression and the amount of available Anl during the screen have substantial effects on the outcome of the screens. In addition to displaying improved activities toward Anl, identified mutants also show an improved discrimination against Met. We have shown that the degree of cell-surface labeling *in vivo* correlates well with the measured rates of Anl activation *in vitro*, which reflects the success of the screen design.

Computational analysis of the experimentally identified mutants revealed a good agreement between computed binding energies and *in vitro* activation data. To better test the computation model, we performed an *in silico* screen for Anl binding on a saturation-mutagenesis library comparable to the experimental library. Computational screen predominantly selects mutants that interact with Anl through hydrogen bonds, whereas the hydrophobic residues are selected more often by the experimental screen. We identify that experimental mutants try to optimize packing at the Y260 and H301 sites, but not at the L13 site. We

discuss possible explanations for these results. Combined results from computation and experiments suggest the importance of various factors in ligand recognition and *in vivo* selection of MetRS mutants. We explore the implications of these factors to the future efforts in the engineering of new MetRS activities.

Table of Contents

Page

List of Figures

List of Tables