Expanding the Biosynthetic Capacity of the Aminoacyl-tRNA Synthetases

Thesis by

Pin Wang

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

California Institute of Technology

Pasadena, California, USA

2003

(Defended on June 3rd, 2003)

To Lili

To My Parents and My Sister

ACKNOWLEDGEMENTS

Before I started to work on my thesis, I thought acknowledgements section would be the easiest one. In reality I found that it was far more difficult than I expected to put together this part of thesis, primarily due to the fact that the bulk of this work would not be completed without the influence of many people in my life. I thank everyone who encouraged me and supported me in many aspects of my life. It is almost impossible to list all names herein and I am sure I have missed some names, for which I apologize.

Foremost, I would like to acknowledge my thesis advisor, David Tirrell, for having guided me into this fascinating area of research, for his constant support and encouragement, for sharing with me his ethological perspectives on science, and for teaching me to follow people's opinion as long as "the opinion does not kill you." I admire his energy, his enthusiasm for science and his good memory. Dave allowed me to work on several projects and to pursue collaborations with different groups. The opportunities I have been afforded in his group were enormous. Dave is extremely supportive of my future career development, has given me confidence to look forward, and helps me in every possible aspect. I am looking forward to working with him again in the near future.

I also would like to thank other members of my committee, Prof. Mark Davis, Prof. Bill Goddard, Prof. Steve Mayo, for their contributions to this work. Their patience and insights have helped move forward much of this thesis work. I appreciate their time dedicated to reviewing my thesis results. In particular, the collaborative efforts on computational design of the aminoacyl-tRNA synthetases with Bill and Steve constitute the critical components of this thesis. Their invaluable suggestions came at a different

angle, which helped me a lot on design of experiments and taught me the computational knowledge. The interactions with Vaidehi Nagarajan from Bill's group and Deepshikha Datta from Steve's group are very enjoyable and valuable. Vaidehi mentored me on computational methods and was patient to answer every of my naïve questions.

I would like to take this opportunity to thank Prof. Zhen-Gang Wang for recruiting me to Caltech. My life at Caltech began with his firm trust and encouragements. Prof. Julie Kornfield was kind enough to allow me to work in her lab for the first summer while Dave was still at UMass. I am grateful for the support from Bill, Steve, Prof. Francis Arnold and Prof. Julie Kornfield during my academic job hunting.

The experience with the entire Tirrell group was wonderful and unforgettable. My progress and accomplishments are always associated with their assistance and ideas. When I first joined the group, Jeff, Kristi, Wendy, Karen, Dave, Jill, Kathy and Nandita from UMass had me settle down in the laboratory. I am very fortunate to have worked with other former members that passed through Caltech during the last five years, particularly Yi, Kent, Ralf, Kevin, Jens, Thierry, Hanna, Youngback, Toshi, Angelika. I especially appreciate the time I spent working with my good colleague and friend Yi Tang. He taught me the essential skills to carry out most of the thesis work. The experience with him was stimulating and inspiring. His never satisfactions of the accomplished work really influenced me and motivated me. His advice for my academic job search was extremely helpful. I am looking for the opportunity to work with him in the near future. I thank Kent for the rigorous and deep discussion on science and encouragements regarding job search. I also would like to thank the current members of

the group, including Isaac, Sarah, Marrisa, Paul, Jinsang, Wei, Jamie, Rebecca, Charles, Julie, Mike, Soojin, Chris, Jin, Kimberley and many others for their friendliness and help. Isaac provided excellent advice on my organic chemistry. Sarah and Jamie helped me with writing English. Jinsang shared with me a lot of his experience in science and pushed me with my job application. Paul made our office more enjoyable. Qualifying exam with Wei was unforgettable. Working with Soojin was also wonderful. I am very fortunate to have this group of colleagues and am looking forward to working with them again in the future.

Life outside the laboratory at Caltech is truly wonderful. The experience with Yi and Dave when we won the 10K business competition is one of the most exciting moments at Caltech. I am really lucky to have interacted with Prof. John Baldeschwieler and taken his E102 class. His enthusiasm about entrepreneurship made our hypothetical company, Duragene, almost real. Together with Chunhui Mo, Joe Andrieu, Girish N. Aakalu, Lili Yang, Ying Go, we founded Caltech Consulting Club. Knowledge gleaned from this club and experiences obtained from two summer projects with two different companies are precious.

I have been very fortunate to gain many lasting friendships over the last five years at Caltech. These friends have been an important part of my graduate school experience not only on a scientific basis but on a personal level as well. I would like to thank the gang like Chunhui, Hongyu, Qiang, Ying, Lan, Huazhang, Xianglei, Wei, Weijun, Lin, Junhua, and Jay for all the good times. Thanks also go to my friends in Boston like Ying, Minghao, Jian and old friends like Xianzhi, Lixin. It is Minghao who encouraged me to transfer to Caltech.

I would like to thank my parents for their love and care. Also I would like to thank my uncle Y. C. Yang and his family for helping me study and live in the US.

Last but definitely not least, I thank my wonderful girlfriend Lili Yang. My biggest achievement at Caltech is having her in my life. Her unconditional support, kindness, and love completely changed my life in the last three years. She made my life colorful and fruitful. I am truly thanking her for everything she has done for me.

The research in this thesis was funded by NIH Grant R01-GM62523 and by the NSF Center for the Science and Engineering of Materials at Caltech.

ABSTRACT

Incorporation of non-natural amino acids into proteins *in vivo* can provide biological materials with new chemical functions and improved physical properties. Examples include new posttranslational modification chemistry by introducing azide and ketone moieties into recombinant proteins, and novel strategies for engineering hyper-stable proteins by incorporating fluorinated side chains. Implementing such methods requires manipulation of protein biosynthesis to specifically alter the genetic code. The rules of the genetic code are established by theaminoacylation reaction, where the aminoacyl-tRNA synthetases (aaRS) catalyze the attachment of the amino acids to theircognate tRNAs. Thus manipulation of cellular aminoacylation reactions could potentially expand the available set of amino acid building blocks for protein engineering and biomaterials engineering.

By simple depletion of the cellular pool of isoleucine and utilization of isoleucine auxotrophic hosts, we were able to force the endogenous isoleucyl-tRNA synthetase to join 5,5,5-trifluoroisoleucine (5TFI) to $tRNA^{Ile}$ and assign it to isoleucine codons *in vivo*. Murine interleukin 2 containing 5TFI retains its biological activity. We showed that engineering of bacterial expression hosts can allow a single RNA message to be read in different ways, depending on the relative rates of competing aminoacylation reactions. Specifically, we showed that the *2S,3R*-form of 4,4,4-trifluorovaline can be assigned either to isoleucine or to valine codons, depending on whether the bacterial host overexpresses the isoleucyl- or the valyl-tRNA synthetase. When anamino acid analog of interest is not recognizedby the corresponding wild-type aaRS, we can either identify the appropriate modification of the amino acid as a promising ligand or design new

synthetase activity. We describe an attempt to develop a virtual ligand screening method to find non-natural amino acids that can serve as ligands for the phenylalanyl-tRNA synthetase and our computational results correlate well with experimental results *in vitro* and *in vivo*. We also present a computational method for identifying the sites of mutations to relax the substrate specificity of the *E. coli* phenylalanyl-tRNA synthetase (*e*PheRS). One designed variant of *e*PheRS allows the efficient *in vivo* incorporation of aryl ketone functionality into proteins *in vivo*. Proteins outfitted with ketone functionality can be chemoselectively ligated with hydrazide reagents under mild conditions. Three designed mutants of *e*PheRS were subjected to extensive examination, and a broad activation profile toward many non-natural aromatic amino acids was observed. *E. coli* host strains were established to over-express these mutant *e*PheRSs, enabling the re-assignment of the Phe codons to many non-natural amino acids. By rational attenuation of the editing function of a leucyl-tRNA synthetase, oxonorvaline was incorporated into a recombinant protein in *Escherichia coli*.

The work described above addresses the multi-site incorporation of new amino acids into proteins *in vivo*, which can be utilized to engineer the overall properties of biomacromolecules such as protein stability. The second component of this thesis focused on the site-specific incorporation of novel amino acids into proteins *in vivo*, which can be applied to problems that require local change of protein behavior. We have refined a previously described system, where we introduce a mutant form of yeast PheRS co-transformed with a cognate suppressor tRNA, allowing incorporation of several aromatic amino acids into proteins in response to an amber codon. The results firmly

demonstrate the general strategy of importing an exogenous synthetase/tRNA pair to achieve site-specific incorporation of non-natural amino acids into proteins *in vivo*.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES