Enzyme promiscuity and the origins of cellular innovations

A dissertation presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy
in
Biochemistry

at Massey University, Albany, New Zealand.

Valerie Wooi Chee Soo

2012

ABSTRACT

Biochemistry textbooks define enzymes as being efficient and highly specific. However, these characteristics are usually associated with a lack of versatility, and therefore, an inability to evolve new functions. In spite of this, it is known that new enzymes can arise rapidly (such as when bacteria evolve antibiotic resistance). One hypothesis proposes that enzymes are actually promiscuous (Jensen, 1976); that is, they are able to carry out secondary reactions, in addition to the one they evolved to catalyze. The goal of this research was to explore the role that promiscuity plays in the origins and evolution of enzyme functions, using *Escherichia coli* as a model organism.

In the first part of this thesis, I report the discovery of two enzymes (alanine racemase and cystathionine β-lyase) that are reciprocally promiscuous, and are dependent on the cofactor pyridoxal 5'-phosphate (PLP) for activity. *In vivo*, the cofactor-mediated promiscuous activities of alanine racemase and cystathionine β-lyase were each successfully improved to near wildtype levels using directed evolution experiments. These results extend Jensen's hypothesis, and led me to propose that PLP played a significant role in the evolution of new enzymes, in the primordial world.

In the second part of the thesis, I developed a comprehensive library-on-library screen to search for *E. coli* proteins that could mediate improved growth in environments containing either a foreign nutrient or a toxin. Proteins were over-expressed in an attempt to increase their weak, promiscuous activities, and to mimic the common genetic phenomenon of gene amplification. Over-expression of individual proteins conferred improved growth to the host cell in 35% of ~2,000 environments. The findings have important implications for understanding bacterial adaptation to new environments, such as when antibiotic resistance emerges. The ability of promiscuous proteins to drive the emergence of new phenotypes, when their expression is increased, validates the feasibility of the Innovation, Amplification and Divergence (IAD) model for the evolution of new genes (Bergthorsson *et al*., 2007).

Overall, the work described in this thesis demonstrates that protein promiscuity is common, though difficult to predict *a priori*. My experimental results are consistent with the work of others, in suggesting that promiscuous activities are evolvable. Together, the high frequency and evolvability of promiscuous proteins appear to underpin many different cellular innovations.

[Blank page]

ACKNOWLEDGEMENTS

Biochemistry was one of my least liked subjects when I was an undergraduate.

But this has changed in less than four years. For introducing me to the coolness of Biochemistry, I thank Dr. Wayne Patrick. As a PhD supervisor, he has done everything he could to make sure I didn't become a burnt pancake — as they say, the first one is usually a tricky one. Above all, Wayne has been a source of inspiration. I particularly appreciate his constructive feedback, which (usually) manifests in the form of red scribbles all over the papers. Perhaps the most meaningful gift I could present to him is not this thesis, but a stash of red pens (which can last until his retirement)!

I am grateful for the all-round advice from Dr. Monica Gerth. Her patience for me is amazing, despite my endless questions about science, life, the universe… She just seems to have a knack of setting me straight, and telling me not to sway when in doubt.

My secondary supervisor, Prof. Paul Rainey, has been generous with his time and evolutionary insights whenever I needed them. Dr. Xue-Xian Zhang has been kind to share his knowledge on microbial genetics. Both have my thanks.

I extend my gratitude to Prof. Dan Andersson (Uppsala University) for his helpful discussions, which have partly shaped the work presented in Chapters 2 and 3. Dan also paid for my meals and lodging during my short trip to Uppsala in 2010. Special thanks to Prof. John Roth (University of California, Davis) for his invaluable comments on the work described in Chapters 4 and 5.

I also wish to thank Dr. Gill Norris for her help in interpreting the mass spec data; Dr. Christopher Squire (University of Auckland) for taking an interest in solving the structure of MetC(P113S); Dr. Tim Clausen (Research Institute of Molecular Pathology, Vienna) for his willingness to share one of his old articles and the structure coordinates of an inhibitor-bound MetC.

Financial support for my PhD came from the Institute of Molecular Biosciences/Institute of Natural Sciences (INS) Doctoral Scholarship, and the New Zealand (NZ) Institute of Advanced Study/INS Postgraduate Research Excellence Award. For sorting out my financial (and other administrative) issues, I thank the Institute

Administrator, Muharram Khoussainova. Travel grants from the NZ Society for Biochemistry and Molecular Biology, and the Federation of European Biochemical Societies, allowed me to attend conferences without financial worries.

A big thank-you goes to the people in the Patrick Lab, especially Paulina Hanson-Manful, Ilana Gerber, Susan Morton, Laura Nigon and Natasha le Roux. In addition to their technical support, the girls have also been excellent lunch buddies. I can't thank Matteo Ferla enough for his help with Illustrator and Photoshop. Mack Saraswat has also been great in taking my mind off research during our chats over coffee.

Colleagues and friends from Building 11 have been very supportive, especially towards the end of my PhD. Moral support from Saumya Agrawal and Martina Dautel was very much appreciated on days when I seemed to be at the end of my tether. Ralph Grand, Jarod Young and Matilda Newton have been offering heavy doses of encouragement and optimism.

I am also indebted to the *ad hoc* proofreading team: Monica, Ralph, Saumya and Matilda.

A large chunk of this thesis was written at home. For keeping the noise volume to a minimum in our house, I thank my flatmates: Kataraina Hohua, Nurul Syaza Abdul Latif and Noriza Ahmad.

I am thankful to have Jingjing Khoo as a best friend. We have supported each other when our experiments didn't go well. I hope she will overcome her last hurdle soon in obtaining a PhD in Biochemistry (from Monash University).

To Mum, Dad and Sis, thank you for believing in me, and my dreams. Your sheer love is unparalleled. 感謝家人一直給慧琪的愛護，支持，鼓勵和包容。 爹哋（蘇芳先生），媽咪（彭水珍女士）和小妹（慧琳），謝謝你們！

Last but not least, I owe various *E. coli* strains (see Section I.3 in Appendix I) great thanks. Without them, this thesis would turn out to be very different.

TABLE OF CONTENTS

ABBREVIATIONS

Alr	Alanine racemase
ASKA	A complete set of *E. coli* K-12 ORF Archive
A_{xxx}	Absorbance measured at xxx nm
cfu	Colony-forming unit
CHES	*N*-cyclohexyl-2-aminoethanesulfonic acid
DAAO	D-amino acid oxidase
DMSO	Dimethyl sulfoxide
DTNB	5,5′-dithiobis(2-nitrobenzoic acid)
EC	Enzyme Commission number
EDTA	Ethylenediaminetetraacetic acid
epPCR	Error-prone PCR
GFP	Green Fluorescent Protein
$(His)_6$	Hexa-histidine
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
lacZ	β-galactosidase
LB	Lysogeny broth
LDH	Lactate dehydrogenase
m/*z*	Mass-to-charge ratio
MetC	Cystathionine β-lyase
MIC	Minimum inhibitory concentration
NAD^+	Oxidized form of β-nicotinamide adenine dinucleotide
NADH	Reduced form of β-nicotinamide adenine dinucleotide
NoIns (or NoInsert)	The negative control plasmid, pCA24N-NoIns (or *E. coli* harbouring the pCA24N-NoIns plasmid)

OD_{600}	Optical density measured at 600 nm
ORF	Open reading frame
PDB	The RCSB Protein Data Bank
PLP	Pyridoxal 5′-phosphate
PM	Phenotype Microarray
RMSD	Root Mean Square Deviation
S.E.M.	Standard error of mean
SOB	Super Optimal Broth
SOC	SOB with Catabolite repression
Tris	Tris(hydroxymethyl)aminomethane
W	Relative fitness
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside