The Development of a New Cloning Strategy for the Biosynthetic Production of Brush-Forming Poly(Amino Acids)

Douglas B. Henderson

(ABSTRACT)

The design and discovery of new surface-active polymers that self-assemble on solid substrates to form brush layers will have a major impact on numerous applications. Through recombinant DNA technology, there exists the potential to harness a cell's protein synthesis machinery to produce a brush-forming poly(amino acid) (or PAA) with an exactly specified amino acid sequence, thus controlling the polymer's composition at a level unequaled by conventional organic polymer synthesis. The presented work demonstrates the cloning, expression, purification and characterization of *de novo*-designed PAA's designed to form brush layers on alumina surfaces. Using conventional recombinant DNA methods, the feasibility of producing a PAA consisting of a poly-glutamate block and a poly-proline block was demonstrated. However, the PAA design was limited by the inherent limitations of conventional cloning techniques.

We introduce here the development of a simple and versatile strategy for producing *de novo*-designed, high molecular weight PAA's using recombinant DNA technology. The basis of this strategy is that small DNA modules encoding for short PAA blocks can be easily inserted directly into a commercially available and unmodified expression vector. The insertions can be made repeatedly until the gene encodes for a polymer of desired molecular weight and composition. Thus, sequential modifications can be made to the PAA without having to re-start the gene assembly process from the beginning, thereby allowing for quick determination of how these changes affect polymer structure and function. The feasibility and simplicity of this method was shown during the production of a PAA, consisting of a long zwitterionic tail block and a short acidic anchor block, designed to form optimal brush layers on alumina surfaces. The success and flexibility of this method indicates that it can be applied for production of *de novo*-

designed polypeptides in general. It is hoped that this method will contribute towards the rapid development of bio-inspired protein-based polymers for a variety of applications.

This dissertation also contains research that aimed to use phage display technology to develop a new liposome-based immunoassay against biological toxins. This work was part of a collaboration effort with the U.S. Department of Defense and Luna Innovations.

Author's Acknowledgements

The author would like to thank his co-advisors, Dr. Kevin E. Van Cott and Dr. Richey M. Davis for the support and guidance that led to the completion of this work. In addition, the author would like to thank each member of his research committee: Dr. William A. Ducker, Dr. Kim Forsten Williams, Dr. Aaron S. Goldstein and Dr. Ravi Saraf.

The author would like to acknowledge that the funding support for a majority of this project was provided by National Science Foundation Grant BES-0086876. Additional support for the work presented in Appendix A was provided by the U.S. Department of Defense and Luna Innovations.

The author would like to thank the faculty and staff of the Chemical Engineering Department.

The author would like to gratefully thank the following persons for training him in the various molecular biology techniques used in this project and for general assistance in the laboratory:

- Kristin Knight
- Tanya O'Sickey
- Roger Van Tassell (of Luna Innovations)
- Myles Lindsay

The author gratefully acknowledges the following persons for their contributions to this project:

- Jody L. Krsmanovic – for the preliminary homopolymer adsorption experiments
- Myles Lindsay – for the initial cloning work of the Pro_mGlu_n genes
- Beth Caba – for the dynamic light scattering experiments
- Kevin E. Van Cott – for his assistance with the HPLC and MS experiments

The author also acknowledges the help of the late Dr. Christopher G. Russell, formerly of Invitrogen, who consulted and aided in the cloning of the initial Pro_mGlu_n genes.

The author gratefully acknowledges Aysen Tulpar, William A. Ducker and his research group for conducting the various polymer adsorption studies presented in Chapter 4. The author would like to thank them for their major role in authoring the included journal article: "Unnatural proteins for the control of surface forces".

The technical assistance of Dr. Ron Cerny of the University of Nebraska Center for Mass Spectrometry is gratefully acknowledged, as well as the services provided by Cleveland Genomics, Inc., Commonwealth Biotechnologies, Inc. and the Biomolecular Research Facility at the University of Virginia.

Finally, and most importantly, the author would like to thank his wife, parents, brother, sisters, family and friends for their love and support. He is truly grateful for each and every one of them.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

a.a. – amino acid
bp – base pair
CIP – alkaline phosphatase, calf intestinal
CV – column volume
DEAE – diethylaminoethyl
DLVO Theory - Derjaguin, Landau, Verwey and Overbeek Theory
DNA – deoxyribonucleic acid
HPLC – high performance liquid chromatography
IDA – iminodiacetic acid
IEP (pI) – isoelectric point
IEC – ion exchange chromatography
IMAC – immobilized metal affinity chromatography
IPTG – isopropyl-B-D-thiogalactopyranoside
kDa – kilo-Daltons
MALDI-TOF MS – matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
MW – molecular weight
NTA – nitrilotriacetic acid
OD – optical density
PAA – poly(amino acids)
PCR – polymerase chain reaction
RE – restriction enzyme
RP – reversed-phase
SDS-PAGE – sodium dodecyl sulfate-polyacrylamide gel electrophoresis
ssDNA – single-stranded deoxyribonucleic acid
T4 PNK – T4 polynucleotide kinase
$\mathbf{T_A}$ – annealing temperature
$\mathbf{T_M}$ – melting temperature
Trx – HP-Thioredoxin

Notes:

- both the three-letter and one-letter amino acid abbreviations are used throughout this work; all amino acid sequences are written in order of amino (N-) terminus → carboxy (C-) terminus
- all DNA sequences are written in order of 5' → 3'

Base	Single-letter Symbol
Adenine	a
Thymine	t
Guanine	g
Cytosine	c

The Genetic Code

Amino Acid	Symbols		Standard Genetic Code
	One-Letter	Three-Letter	
Alanine	A	Ala	gct
			gcc
			gca
			gcg
Cysteine	C	Cys	tgt
			tgc
Aspartic Acid	D	Asp	gat
			gac
Glutamic Acid	E	Glu	gaa
			gag
Phenylalanine	F	Phe	ttt
			ttc
Glycine	G	Gly	ggt
			ggc
			gga
			ggg
Histidine	H	His	cat
			cac
Isoleucine	I	Ile	att
			atc
			ata
Lysine	K	Lys	aaa
			aag
Leucine	L	Leu	tta
			ttg
			ctt
			ctc
			cta
			ctg
Methionine	M	Met	atg
Asparagine	N	Asp	aat
			aac
Proline	P	Pro	cct
			ccc
			cca
			ccg
Glutamine	Q	Gln	caa
			cac

Amino Acid	Symbols		Standard Genetic Code
	One-Letter	Three-Letter	
Arginine	R	Arg	cgt
			cgc
			cga
			cgg
			aga
			agg
Serine	S	Ser	agt
			agc
			tct
			tcc
			tca
			tcg
Threonine	T	Thr	act
			acc
			aca
			acg
Valine	V	Val	gtt
			gtc
			gta
			gtg
Tryptophan	W	Trp	tgg
Tyrosine	Y	Tyr	tat
			tac
STOP	-	-	taa
			tag
			tga