

**THE QUEST FOR A GENERAL CO-CRYSTALLIZATION STRATEGY
FOR MACROMOLECULES: LESSONS ON THE USE OF CHAPERONES
FOR MEMBRANE PROTEIN CRYSTALLIZATION**

A Dissertation
Presented to
The Academic Faculty

by

Jennifer Leigh Johnson

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry and Biochemistry

Georgia Institute of Technology

August 2015

COPYRIGHT © JENNIFER JOHNSON 2015

**THE QUEST FOR A GENERAL CO-CRYSTALLIZATION STRATEGY
FOR MACROMOLECULES: LESSONS ON THE USE OF CHAPERONES
FOR MEMBRANE PROTEIN CRYSTALLIZATION**

Approved by:

Dr. Raquel Lieberman, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Adegboyega Oyelere
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Wendy Kelly
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Ingeborg Schmidt-Krey
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Nael McCarty
Department of Pediatrics
Emory University

Date Approved: [May 14, 2015]

To Dad, Mom, and Sam. Words cannot express how much you three mean to me.

ACKNOWLEDGEMENTS

First I would like to thank my advisor, Dr. Raquel Lieberman for the support, guidance, and advice through the years. It was not always an easy road, but we made it though.

Next, my success in this endeavor would not have been possible without each one of the lab members, past and present, that I had the opportunity to work with. Jeff Culver, Katherine Turnage, Dr. Susan Orwig, Dr. Rebecca Donegan, Dr. Jason Drury, Dr. Derrick Watkins, Sibel Kalyoncu, Swe-Htet Naing, Chiamaka Ukachukwu, Dr. Tanay Dasai, Lindsay Porter, Dr. Shannon Hill, Elaine Nguyen, Casey Bethel, Dustin Huard, you guys rock! To undergraduate students David Heaner, Ivan Morales, Dana Freeman, Aly Sheppard, James Rives, Rachel Wills, Michelle Kwon, the lab definitely would not have been the same without you (I hope I'm not forgetting anyone).

I would like to thank my committee members, Dr. Ingeborg Schmidt-Krey, Dr. Wendy Kelly, Dr. Adegboyega Oyelere, and Dr. Nael McCarty for their guidance and direction. I also collaborated with numerous labs and researchers across the world, and I would like to thank each one of you for your contributions to my work and allowing me the opportunity to contribute to yours.

Finally, I would like to thank my family, for without each member, I never would have made it. Especially Mom, Dad, and Sam, you guys know.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	IV
LIST OF TABLES	XI
LIST OF FIGURES	XII
SUMMARY	XVIII
CHAPTER 1: INTRODUCTION TO MEMBRANE PROTEIN CRYSTALLIZATION	1
1.1 Membrane Proteins and the Importance of Determining Their High Resolution Structure 1	
1.1.1 The prevalence of membrane proteins in biology and pharmaceuticals.....	1
1.1.2 X-ray crystallography as a tool for studying proteins	2
1.1.3 Membrane protein structures are severely underrepresented	2
1.2 Molecular Biology and Protein Expression	3
1.2.1 Prokaryotic orthologs	3
1.2.2 Truncation or removal of regions by molecular biology	4
1.2.3 Surface entropy reduction	4
1.2.4 Synthetic symmetrization.....	5
1.2.5 Covalent crystallization chaperones	5
1.2.6 Stabilizing mutations and directed evolution	6
1.2.7 Expression platform	7
1.3 Membrane Protein Purification.....	7
1.3.1 Detergent selection for membrane solubilization and purification.....	8
1.4 Membrane Protein Crystallization.....	8
1.4.1 Limited proteolysis.....	9
1.4.2 Lipid cubic phase (LCP)	9
1.4.3 Ligand or other binding partner	10
1.4.4 Non-covalent crystallization chaperones.....	10
1.5 An Example: The Structure of Human β2 Adrenergic Receptor	12
1.6 Thesis Objectives	13

1.6.1	Development of an antibody fragment and monomeric streptavidin as crystallization chaperones for tagged membrane proteins.....	14
1.6.2	Molecular basis of novel 1-deoxygalactonojirimycin arylthiourea pharmacological chaperones binding to α -galactosidase A for the treatment of Fabry disease.....	15

CHAPTER 2: EXPRESSION, PURIFICATION, AND DETERGENT OPTIMIZATION OF SIGNAL PEPTIDE PEPTIDASE: LESSONS FOR PREPARING MEMBRANE PROTEINS FOR STRUCTURAL CHARACTERIZATION 17

2.1	Introduction	17
2.1.1	Biology of SPP.....	17
2.1.2	Chapter Overview	19
2.1.3	Individual contributions to the work	19
2.1.4	Publication resulting from this work	20
2.2	Methods/Protocol.....	20
2.2.1	Signal peptide peptidase ortholog selection	20
2.2.2	Molecular biology for signal peptide peptidase orthologs	20
2.2.3	TEV protease expression and purification	21
2.2.4	Signal peptide peptidase expression.....	21
2.2.5	Membrane isolation from harvested cells	23
2.2.6	Protein solubilization from membrane and purification by Ni ²⁺ -affinity chromatography.....	23
2.2.7	Protein stability in different detergents assessed by size exclusion chromatography ...	24
2.2.8	Protein stability in different detergents by CD.....	25
2.2.9	Detergent quantification by TLC	27
2.3	Results.....	28
2.3.1	Molecular biology for target membrane protein	28
2.3.2	Membrane protein expression	31
2.3.3	Membrane isolation and protein purification	31
2.3.4	Protein stability in different detergents	31
2.3.5	Protein stability in different detergents by circular dichroism	33
2.3.6	Detergent quantification.....	34
2.4	Discussion and Lessons Learned	36

CHAPTER 3: CHARACTERIZATION OF EPITOPE-SPECIFIC FAB ANTIBODY FRAGMENT FOR USE AS A GENERAL CRYSTALLIZATION CHAPERONE.....	42
3.1 Introduction	42
3.1.1 3D5 scFv	43
3.1.2 Conversion of 3D5 scFv to anti-EE scFv	44
3.1.3 Chapter overview	45
3.1.4 Individual contributions to the work	45
3.1.5 Publications resulting from this work	46
3.2 Methods	46
3.2.1 Molecular biology, expression and purification of Fab/EE.....	46
3.2.2 Biophysical characterization of Fab/EE	47
3.2.3 Protein crystallization, data collection, structure determination and refinement.....	48
3.2.4 Computational analysis of Fab/EE crystal contacts	49
3.2.5 Molecular biology, expression and purification of soluble test proteins presenting the EE epitope	50
3.2.6 Binding assays and complex formation.....	50
3.2.7 Size exclusion chromatography	51
3.3 Results.....	51
3.3.1 Fab/EE molecular biology, expression and purification	51
3.3.2 Fab/EE structural characterization	51
3.3.3 Fab/EE forms stable complexes with soluble EE-tagged client proteins	58
3.4 Discussion	60
3.4.1 Fab/EE rationale, cloning, and biophysical properties	60
3.4.2 Fab/EE structural characterization	61
3.4.3 Fab/EE forms stable complexes with soluble EE-tagged client proteins	61
 CHAPTER 4: PROGRESS TOWARD CO-CRYSTALLIZATION WITH EPITOPE-SPECIFIC FAB ANTIBODY FRAGMENT AND TAGGED MEMBRANE PROTEINS.....	 63
4.1 Introduction	63
4.1.1 Individual contributions to the work	64
4.1.2 Publications resulting from this work	64
4.2 Methods	65

4.2.1	A ₂ aR expression and purification.....	65
4.2.2	SPR with Fab/EE and A ₂ aR-GFP-EE	67
4.2.3	ELISA of Fab/EE and A ₂ aR-GFP-EE	67
4.2.4	SPP molecular biology, expression and purification.....	68
4.2.5	Fab/EE complexation with SPP-EE variants over size exclusion chromatography	73
4.2.6	Intimin molecular biology, expression, and purification.....	73
4.2.7	Complexation of Fab/EE with intimin-EE7, -EE8.....	76
4.2.8	Molecular dynamic (MD) simulations	76
4.3	Results.....	77
4.3.1	EE peptide insertion into the extracellular loop of membrane proteins	77
4.3.2	Fab/EE forms complexes with EE-tagged membrane proteins	79
4.3.3	MD simulations of Fab/EE-intimin-EE complexes.....	81
4.3.4	Fab/EE forms a solution complex with SPP constructs with truncated TM6-TM7 loop 84	
4.4	Discussion	86
4.4.1	Conclusions: implications for chaperone-mediated crystallization of MPs	86

CHAPTER 5: MONOMERIC STREPTAVIDIN AS A CRYSTALLIZATION CHAPERONE FOR A BIOTINYLATED SOLUBLE TEST PROTEIN 88

5.1	Introduction	88
5.1.1	Individual contributions to the work	92
5.2	Methods	92
5.2.1	Protein expression and purification.....	92
5.2.2	Enzymatic biotinylation of MBPavi using BirA and purification of biotinylated MBPavi	98
5.2.3	Complexation of bMBPavi with mSA	99
5.2.4	Crystallization trials of bMBP:mSA2 complex.....	100
5.3	Results.....	100
5.3.1	Insertion of the Avi-tag into MBP.....	100
5.3.2	Purification of components	101
5.3.3	Complexation of biotinylated MBPavi with monomeric streptavidin.....	102
5.3.4	Progress in the crystallization of bMBP:mSA2 complex	103
5.4	Discussion	104
5.4.1	Concluding remarks and future directions	105

**CHAPTER 6: MOLECULAR BASIS OF NOVEL 1-
DEOXYGALACTONOJIRIMYCIN ARYLTHIOUREA
PHARMACOLOGICAL CHAPERONES BINDING TO ALPHA
GALACTOSIDASE A FOR THE TREATMENT OF FABRY DISEASE.... 107**

6.1	Introduction	107
6.1.1	Individual contributions to the work	109
6.1.2	Publication resulting from this work	110
6.2	Methods	110
6.2.1	Synthesis of DGJ-arylthioureas and other DGJ Derivatives	110
6.2.2	Cell culture, transfection, and pharmacological chaperone test	111
6.2.3	Measurement of enzyme activities in the lysosome	112
6.2.4	Inhibition and stabilization of α -Gal A <i>in vitro</i>	112
6.2.5	Crystallization and X-ray data collection of α -Gal A in complex with DGJ- <i>p</i> FPhT ..	113
6.2.6	Immunofluorescence staining	115
6.2.7	Immunoblotting.....	115
6.3	Results.....	115
6.3.1	Comparison of α -Gal A crystal structures in complex with DGJ and DGJ- <i>p</i> FPhT.....	115
6.3.2	Effects of DGJ arylthiourea compounds on α -Gal A <i>in vitro</i>	116
6.3.3	The activity of α -Gal A is increased in normal and FD fibroblast cells after treatment with DGJ-ArTs.....	117
6.3.4	The effect of DGJ- <i>p</i> MeOPhT on mutants of α -Gal A.....	121
6.3.5	Autophagic proteins are upregulated in FD cells and restored by DGJ- <i>p</i> MeOPhT and DGJ-FPhT	121
6.3.6	Effects of proteostasis regulators on α -Gal A activity are additive	124
6.4	Discussion	125

**CHAPTER 7: PERSPECTIVES AND FUTURE DIRECTIONS ON THE USE OF
EPILOPE-SPECIFIC CRYSTALLIZATION CHAPERONES 129**

7.1	Perspectives on Membrane Protein Crystallography	129
7.2	Perspectives on the Use of Epitope-Specific Crystallization Chaperones for Membrane Proteins	131
7.3	Conclusions.....	132

REFERENCES.....	133
VITA.....	159

LIST OF TABLES

Table 1 Primers used for the insertion of a TEV cleavage site in <i>M. mar</i> SPP	23
Table 2 Melting temperatures for <i>H. mar</i> SPP in 6 detergents.....	34
Table 3 Data table for TLC detergent test	36
Table 4 Biophysical characteristics of 3D5 compared to scFv variants (192).....	44
Table 5 Fab/EE data collection and refinement statistics	49
Table 6 Biophysical characteristics of EE peptide-binding antibody fragments	53
Table 7 Fab/EE heavy chains (chain H and A) Kabat numbering and modelled residues	53
Table 8 Fab/EE light chains (chain L and B) Kabat numbering and modelled residues	54
Table 9 PISA contact comparison between scFv/EE and Fab/EE.....	56
Table 10 Characterization of EE-tagged protein binding kinetics of Fab/EE by SPR.....	59
Table 11 Primers for A2aR-GFP-EE SDM, colony screening, and sequencing.....	65
Table 12 <i>Mm</i> SPP constructs.....	71
Table 13 Primers used to make <i>Mm</i> SPP constructs outlined in Table 12	72
Table 14 Design for intimin-EE constructs	74
Table 15 Primers used to make intimin-EE3, -EE4, -EE7, and -EE8 using SDM	75
Table 16 Characterization of EE-tagged protein binding kinetics of Fab/EE by SPR.....	79
Table 17 Primers used for both rounds of SDM to make MBPavi.....	94
Table 18 Data collection and refinement statistics for α -Gal A in complex with DGJ- <i>p</i> FPhT pharmacological chaperone	114
Table 19 IC ₅₀ values of DGJ derivative compounds against Gal A in vitro	120

LIST OF FIGURES

Figure 1 The structure of β 2AR complexed with its G protein as an example of techniques used to crystallize membrane proteins	13
Figure 2 Cleavage action of SPP in membrane of the ER	18
Figure 3 Alignment of human SPP with 3 archaeal SPP orthologs	22
Figure 4 Workflow to determine protein stability in different detergents using circular dichroism.....	25
Figure 5 Structures of detergents used for SEC and CD experiments	26
Figure 6 Workflow to determine optimal filter MWCO for concentrating protein samples	28
Figure 7 Signal-3L and SignalP signal sequence prediction for <i>M. mar</i> SPP	29
Figure 8 pET-22b(+) DNA and amino acid sequence in the area where the ortholog DNA is inserted.....	30
Figure 9 SDS-PAGE of <i>M. mar</i> SPP samples during each purification step.....	32
Figure 10 Ni^{2+} affinity chromatographs from purification of two SPP orthologs	32
Figure 11 Gel filtration chromatograms of SPP orthologs in 6 detergents	33
Figure 12 Normalized CD thermal unfolding experiment of <i>H. mar</i> SPP in 6 detergents.	34
Figure 13 Detergent quantification of SPP sample using TLC	35
Figure 14 AcrB crystals using visible and UV light.....	40
Figure 15 Anti-his scFv 3D5 crystal lattice.....	43
Figure 16 Fab/EE purification.....	52
Figure 17 Structure of Fab/EE.....	55
Figure 18 Interactions and final $2F_o-F_c$ electron density for crystal contact ID2.....	57
Figure 19 Interactions and final $2F_o-F_c$ electron density for crystal contact ID5	58
Figure 20 Complexation data for Fab/EE binding to soluble EE peptide-containing proteins.....	59
Figure 21 Design and purification of A_2aR -GFP-EE	66
Figure 22 Alignment of MmSPP-EE variants showing EE tag in three locations.....	70
Figure 23 Intimin structure showing location of designed intimin-EE constructs	75
Figure 24 WT intimin, intimin-EE7 and intimin-EE8 SDS-PAGE.....	76
Figure 25 Fab/EE binding to EE-tagged A_2aR -GFP	79

Figure 26 Elution profile of purified Fab/EE incubated with intimin WT, intimin-EE7 or intimin-EE8.....	80
Figure 27 Fab/EE complexes with HmSPP-EE over SEC	81
Figure 28 Molecular dynamics analysis of intimin-EE1 and intimin-EE2 with and without Fab/EE	83
Figure 29 Analysis of intimin L4 interactions	84
Figure 30 Fab/EE complexes with MmSPP-EE2S constructs	85
Figure 31 Structure of WT streptavidin.....	88
Figure 32 Design of monomeric streptavidin based on rhizavidin dimer	91
Figure 33 Alignment of WT MBP and MBPavi after each round of SDM	94
Figure 34 Purification of BirA, MBPavi, and MBP-mSA using amylose resin	97
Figure 35 Cleavage of mSA and gel filtration over Superdex 75 column	98
Figure 36 Purification of biotinylated MBPavi using avidin resin	99
Figure 37 MBP structure showing placement of Avi-tag	101
Figure 38 Gel filtration and SDS-PAGE of the bMBP:mSA complex.....	103
Figure 39 Structures of 3 partial or full Avi-tags from the PDB	105
Figure 40 Streptavidin in complex with 13 amino acid miniprotein ligand	106
Figure 41 Design criteria for the novel FD PCs, DGJ-ArTs	109
Figure 42 Chemical structures of DGJ-ArTs and DGJ derivatives	111
Figure 43 Synthesis of DGJ-ArT pharmacological chaperones and bicyclic DGJ derivative	111
Figure 44 Crystal structure of α -Gal A in complex with DGJ-pFPhT.....	118
Figure 45 Effects of DGJ-ArTs and derivatives on human normal fibroblast cells in vitro	119
Figure 46 Stabilization activity of DGJ and DGJ derivatives on α -Gal A in skin fibroblast cells.	120
Figure 47 Gal-A activity is enhanced by DGJ-ArTs in normal and FD fibroblasts	122
Figure 48 Activity of DGJ-pMeOPhT on normal and FD fibroblasts, and COS7 cells expressing WT and mutant α -Gal A.	123
Figure 49 Effects of DGJ-pMeOPhT and DGJ-pFPhT in the impairment of autophagy of FD fibroblasts	124
Figure 50 Effects of DGJ-pMeOPhT with proteostasis regulators on α -Gal A activity	125

LIST OF SYMBOLS AND ABBREVIATIONS

Abbreviation or symbol	Definition
°	Degrees
μ	Micro
3D	3 Dimensional
4-MU	4-Methylumbelliferone
4-PBA	4-Phenylbutyrate
Å	Angstrom
A ₂ aR-GFP-EE	EE-tagged human adenosine A ₂ a G protein-coupled receptor
AA	Amino acid
AcrB	Acriflavine resistance protein from <i>E. coli</i>
ACS	American Chemical Society
AD	Alzheimer's disease
APS	Advanced Photon Source
ArT	Arylthiourea
ATP	Adenosine triphosphate
Avi-tag	Amino acid sequence GLNDIFEAQKIEWHE
BirA	Biotin ligase enzyme
BLASTp	Basic Local Alignment Search Tool for proteins
BRIL	Apocytochrome b ₅₆₂ RIL
BSA	Bovine serum albumin
C8E4	Octyltetraoxyethylene
CD	Circular dichroism
cDNA	Complementary DNA
CDR	Complementarity determining region
CFTR	Cystic fibrosis transmembrane conductance regulator
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CHAPSO	3-[(3-Cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate
CHS	Cholesteryl hemisuccinate
CMC	Critical micelle concentration
CPM	<i>N</i> -[4-(7-Diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide
CV	Column volume
Cy-5	5-Cyclohexyl-1-pentyl-β-D-maltoside
Da	Dalton
DAPI	4',6-Diamidino-2-phenylindole
DARPinS	Designed ankyrin repeat proteins

DDM	<i>n</i> -dodecyl- β -D-maltopyranoside
DELTA	Domain Enhanced Lookup Time Accelerated
DGJ	1-Deoxygalactonojirimycin
DGJ-BnT	DGJ <i>N</i> ² -benzylthiourea
DGJ-NphT	<i>N</i> ² -(1-naphthyl) DGJ
DGJ-pFPhT	<i>N</i> ² -(<i>p</i> -Fluorophenyl)
DGJ-pMeOPhT	<i>N</i> ² -(<i>p</i> -Methoxyphenyl) DGJ
DGJ-pMeSPhT	<i>N</i> ² -(<i>p</i> -Methylthiophenyl) DGJ
DM	<i>n</i> -Decyl- β -D -maltopyranoside
DMEM	Dulbecco's Modified Eagle's Medium
DMPC	1,2-Dimyristoyl- <i>sn</i> -glycero-3-phosphocholine
DNA	Deoxyribonucleic acid
DPC	<i>n</i> -Dodecylphosphocholine
<i>E. coli</i>	<i>Escherichia coli</i>
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EE	Amino acid sequence EYMPME
EGTA	Ethylene glycol tetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
ERT	Enzyme replacement therapy
Fab/EE	Anti-EE Fab antibody fragment
FACS	Fluorescence-activated cell sorting
FC12	Fos-choline 12
FD	Fabry disease
FPLC	Fast protein liquid chromatography
FSEC	Fluorescent size exclusion chromatography
Gb3	Globotriasylceramide
GC-MS	Gas chromatography coupled with mass spectrometry
GFP	Green fluorescent protein
GPCR	G protein-coupled receptor
<i>H. mar</i>	<i>Haloarcula marismortui</i>
<i>H. sal</i>	<i>Halobacterium salinarum</i>
HBS	HEPES buffered saline, 50 mM HEPES pH 7.5, 150 mM NaCl
HCDR	Complementarity determining region from the heavy chain
HEK	Human embryonic kidney
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
Hex	Hexosaminidase
<i>Hm</i> SPP-EE	EE-tagged SPP from <i>H. marismortui</i>

HRP	Horseradish peroxidase
IAP	Intramembrane aspartyl protease
IC50	Half the maximal inhibitory concentration
ICL3	Intracellular loop 3
intimin-EE	EE-tagged <i>E. coli</i> intimin
IPTG	Isopropyl β -D-1-thiogalactopyranoside
K _D	Dissociation constant
kDa	Kilodalton
LB	Luria-Burtani
LCP	lipidic cubic phase
LDAO	<i>N,N</i> -Dimethyldodecylamine- <i>N</i> -oxide
LDAO	2,2-Didecylpropane-1,3-bis- β -D-maltopyranoside or lauryl maltose neopentyl glycol
LDH	Lactate dehydrogenase
LSD	Lysosomal storage disorder
M	Molar
<i>M. mar</i>	<i>Methanoculleus marisnigri</i>
MBP	Maltose-binding protein
MBPavi	MBP harboring the Avi-tag
MBP-EE	Maltose binding protein with C-terminal EE tag
MBP-His ₆	Maltose binding protein with C-terminal hexahistidine tag
MBP-KEE	Maltose binding protein with internal EE-tag
MBP-mSA2	MBP-fused mSA2
MD	Molecular dynamic
<i>MmSPP</i> -EE	EE-tagged SPP from <i>M. marisnigri</i>
mSA	Monomeric streptavidin
mSA2	Monomeric streptavidin harboring L25H mutation
MWCO	Molecular weight cut-off
NMR	Nuclear magnetic resonance
NTR1	Neurotensin receptor 1
OD ₆₀₀	Optical density measure at 600 nm
OG	<i>n</i> -Octyl- β -d-glucoopyranoside
PBS	Phosphate buffered saline
PC	Pharmacological chaperone
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PDC	Protein detergent complex
PEG	Polyethylene glycol
pFPhIM-DGJ	5 <i>N</i> ,6 <i>S</i> -(<i>p</i> -Fluorophenylimino methylydene)-6-thio-1-deoxynojirimycin
PISA	PDBe Protein Interfaces, Surfaces and Assemblies

PMSF	Phenylmethylsulfonyl fluoride
PSI	Pounds per square inch
REU	Research Experience for Undergraduates
RF	Restriction free
RMSD	Root mean squared deviation
RNA	Ribonucleic acid
RPM	Revolutions per minute
RU	Response units
scFv	Single-chain variable fragment
scFv/EE	Anti-EE scFv
scFv-EE ₁	Single chain antibody fragment with EE-tag in the linker region
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SE-HPLC	Size exclusion high pressure liquid chromatography
SEM	Standard error from the mean
SER	Surface entropy reduction
SER-CAT	Southeast Regional Collaborative Access Team
SPP	Signal peptide peptidase
SPR	Surface plasmon resonance
SUNY	State University of New York
T4L	T4 lysozyme
TB	Terrific broth
TCA	Trichloroacetic acid
TEV	Tobacco etch virus
TLC	Thin layer chromatography
T _m	Melting temperature
TM	Transmembrane helix
TMB	3,3',5,5'-Tetramethylbenzidine
Tr-mStrav	An earlier, less stable version of mSA
TRX	Thioredoxin
WT	Wild-type
α -gal A	α -Galactosidase A
α -NAGA	α -N-Acetyl-D-galactosaminide for α -N-acetylgalactosaminidase
β 2AR	Human β 2 adrenergic receptor
β CD	β -Cyclodextran
β -Gal	β Galactosidase

SUMMARY

Crystallization is often a major bottleneck to macromolecular structure determination. This is particularly true for membrane proteins, which have hydrophobic surfaces that cannot readily form crystal contacts. Of the roughly 109,000 protein structures in the PDB, only about 539 represent unique membrane proteins, despite immense interest in membrane proteins from both a biological and therapeutic standpoint. Membrane protein crystallization has been facilitated by the development of new detergents, lipidic cubic phase methods, soluble protein chimeras, and non-covalent protein complexes. The design process of protein fusion constructs and non-covalent antibody fragments specific for each target membrane protein, however, is costly and time-consuming. An improved, more general method of membrane protein co-crystallization is needed. This dissertation details the development of two approaches for cost-effective non-covalent crystallization chaperones: (1) Engineered hypercrystallizable Fab antibody fragment with high affinity for EYMPME (EE epitope), which form complexes with EE-tagged soluble and membrane proteins. (2) Engineered monomeric streptavidin (mSA2) for complexation with biotinylated membrane proteins. Both methods are generalizable through insertion of a short epitope into a surface-exposed loop of a membrane protein by site directed mutagenesis. Crystallization trials of representative chaperone-membrane protein complexes and possible difficulties with the approach are discussed.

CHAPTER 1: INTRODUCTION TO MEMBRANE PROTEIN CRYSTALLIZATION

1.1 MEMBRANE PROTEINS AND THE IMPORTANCE OF DETERMINING THEIR HIGH RESOLUTION STRUCTURE

1.1.1 The prevalence of membrane proteins in biology and pharmaceuticals

Integral membrane proteins represent nearly 30% of all sequenced genomes (10), specifically, ~ 27% of the human genome (11). Membrane proteins and their corresponding functions are diverse, and include (a) receptors that mediate specific cellular responses upon ligand binding, (b) transporters and channels that control movement of substrates across membranes into the cell, (c) intramembrane enzymes that catalyze chemical reactions, (d) intramembrane structural and adhesion proteins that mediate cell-cell contacts and development of cell structure, and (e) intramembrane ligand proteins that enable communication between cells. Because of their broad functions and their location on the surface of the cell, membrane proteins represent over half of all known drug targets (11, 12). The study of membrane proteins is further motivated by the numerous associated diseases caused by mutations within membrane proteins. To name a few examples, a cause of early onset Alzheimer's disease (AD) is mutations within presenilin (13), an intramembrane aspartyl protease that cleaves amyloid precursor protein to generate the amyloid β peptide that is deposited in the brain of AD patients. Mutations in the transporter cystic fibrosis transmembrane conductance regulator (CFTR) affect regulation of chloride ions and cause viscous secretions in the lungs and pancreas resulting in cystic fibrosis (14). Finally, mutations within $\beta 2$

adrenergic receptor (β 2AR), a G protein-coupled receptor (GPCR) are thought to be associated with certain asthma phenotypes (15).

1.1.2 X-ray crystallography as a tool for studying proteins

The three-dimensional (3D) structure of crystalline molecules, including macromolecules (DNA and proteins), can be determined by taking advantage of the diffraction of X-rays by the electrons in the highly ordered crystalline sample (16, 17). The resulting electron density map and best-fit model of the atoms in the protein molecule can lead to insights into the protein mechanism which may allow researchers to better understand associated diseases. The high resolution study of ligand-binding properties of the protein of interest can lead to drug discovery and structure-based drug design. An example of the value of crystal structures as a tool for drug design for a soluble protein can be seen in Chapter 6 of this thesis and a review of structural biology and drug discovery can be found in (18) and (19).

1.1.3 Membrane protein structures are severely underrepresented

Despite the importance and prevalence of membrane proteins throughout biology and human disease, their structures are severely underrepresented in the Protein Data Bank (PDB, www.rcsb.org). Of the 108,789 structures in the PDB as of May 13, 2015, only 539 structures are of unique membrane proteins (see <http://blanco.biomol.uci.edu/mpstruc/>). One of the first steps of protein X-ray crystallography is crystal growth, and challenges of membrane protein crystallization are numerous. Methods to overcome the challenges can be broken into 3 phases: molecular biology and expression, purification, and crystallization. Each phase has its own challenges and necessary considerations. To better understand the difficulty of membrane

protein crystallization and the methods used in the following chapters, several emerging and successful strategies geared toward crystallization of membrane proteins are covered in the following sections.

1.2 MOLECULAR BIOLOGY AND PROTEIN EXPRESSION

During the molecular biology and expression phase of protein crystallography, endless modifications can be made to the protein construct and the expression protocol. Each new construct must then be tested for protein expression and often several, sometimes hundreds, of constructs are made and tested before successful crystallization. The following sections discuss some of the principles considered when designing constructs and expression protocols for protein crystallization, some of which are designed to increase protein crystallization propensity directly and some indirectly by increasing protein expression, stability, and/or solubility. Many of the following strategies have the disadvantage of solving the structure of a non-native protein, but the crystallization of the protein would likely otherwise be unsuccessful.

1.2.1 Prokaryotic orthologs

Though human membrane proteins may be of particular interest from a medical standpoint, they are especially difficult to express and crystallize. Eukaryotic proteins generally require chaperones and posttranslational modifications for proper folding, requiring the use of an expensive, time-consuming, and often low-yielding eukaryotic expression systems (20). Eukaryotic membrane proteins often have prokaryotic orthologs which can be expressed in higher yields in *Escherichia coli* and are easier to crystallize because eukaryotic proteins require posttranslational modifications and cofactors for stability (see Chapter 2). Understanding the prokaryotic structure usually leads to

significant insights into the human protein counterpart. Recent prokaryotic structures of important human proteins include vitamin C transporter from *E. coli* (21) glutamate transporter from archaeon *Pyrococcus horikoshii* (22), and the translocator protein from *Rhodobacter sphaeroides* (23). Though structures are likely very similar between the eukaryotic and prokaryotic orthologs, small differences in structure and function may limit drug discovery (24).

1.2.2 Truncation or removal of regions by molecular biology

Unstructured or flexible loop are difficult to crystallize because they lead to heterogeneity in the protein solution. Recently, Tanabe *et al.* (25) successfully crystallized human adiponectin receptors AdipoR1 and AdipoR2 by removing the N-terminal 87 and 99 amino acids from AdipoR1 and AdipoR2 respectively. This truncation improved the expression and purification properties of the protein and allowed crystal growth, without changing its ligand binding properties (26). Similarly, the first 66 amino acids of the polymodal K2P channel TREK-1 were removed (27). This same principle is often employed by the cleavage of purification tags after affinity purification, which has been shown to increase crystal diffraction. A cleavage site must be inserted into the protein sequence in the molecular biology phase (see Chapter 2).

1.2.3 Surface entropy reduction

Surface entropy reduction (SER) is a method of systematically replacing small patches of 2 or 3 surface residues that have high conformational entropy (for example, lysine, glutamate, and glutamine) with residues of low conformational entropy. Alanine is a good replacement residue, and using threonine and tyrosine as replacement residues can help mediate crystal contacts (28). This method has been successful for the crystallization

of numerous soluble proteins (for example (29, 30), a list of structures solved using SER can be found at <http://ginsberg.med.virginia.edu/Ser/>), and the technique could be utilized in soluble loops of membrane proteins. If the protein structure is unknown, the determination of residues to mutate without disrupting structure and function can be difficult, but a prediction server is available (<http://services.mbi.ucla.edu/SER/>). The addition of nonpolar amino acids can decrease protein solubility and stability, making protein expression and purification difficult (31), therefore this method can only be utilized with proteins that are stable and relatively easy to express.

1.2.4 Synthetic symmetrization

To aid in crystallization, specific mutations are introduced on the protein surface to drive symmetric self-association of the target protein. For example, a single cysteine mutation was introduced on the surface of T4 lysozyme (T4L) to form an intermolecular disulfide bond, leading to several new crystal forms of T4L (32). The same approach was used to crystallize thirans CelA, a protein of previously unknown structure (33). A similar approach involves engineering metal-binding sites on the surface of the target protein to promote metal-mediated synthetic symmetrization (34). The utility of this method has been demonstrated using maltose binding protein (MBP) and T4L (34) and ferritin (with some variants forming 24-mers) (35), but to my knowledge, not membrane proteins. It may be difficult to design mutations for synthetic symmetrization for a protein of unknown structure, but principles of the design process have been discussed (34).

1.2.5 Covalent crystallization chaperones

Soluble, easily expressed, globular proteins have been fused to proteins that are difficult to crystallize to improve crystallization propensity (31). Many well folded

soluble proteins have been used as so called covalent chaperones (or carrier proteins), including MBP (36-38), lysozyme (39), thioredoxin (TRX) (40). T4L (41-44) and apocytochrome *b₅₆₂*RIL (BRIL) (45, 46) have been especially successful for the crystallization of GPCRs, by replacing the third intracellular loop (ICL3) or the N-terminus with the covalent chaperone. This method has the further advantages that the covalent chaperone often increases protein expression level and solubility (47-50) and the fusion protein is often also used as a tag for affinity purification, streamlining the purification protocol (51, 52). A disadvantage of this approach is that the linker region between the covalent crystallization chaperone and the protein of interest is flexible and could introduce conformational heterogeneity in the crystallization drop (53). The addition of the covalent chaperone has also been shown to cause deformation in the structure of the target protein (54, 55).

1.2.6 Stabilizing mutations and directed evolution

Stabilizing mutations and directed evolution are techniques used to increase the protein expression, stability, and solubility, indirectly increasing the crystallization propensity of the protein. Stabilizing mutations are usually rationally designed point mutations. For example, seven non-conserved residues of a glutamate transporter were replaced with histidine residues which resulted in improved protein expression (22). Several other examples exist for the crystallization of membrane proteins (54, 56, 57).

Directed evolution involves generating genetic diversity by error prone polymerase chain reaction (PCR), transforming the entire DNA library into the desired expression host, expressing the protein, then selecting for the desired property (58). The selection process can vary widely depending on the protein and the desired property (for

example, higher expression, increased activity, or higher melting temperature). For GPCRs, a protocol for directed evolution has been outlined by Schlinkmann and Plückthun (59, 60). For the selection process, the authors add a fluorescent ligand of the GPCR to the cells expressing the library of GPCR mutants. The cells are sorted using fluorescence-activated cell sorting (FACS), and cells with higher fluorescence will harbor GPCR DNA of protein capable of higher expression levels. The GPCR DNA is sequenced from single colonies and the protein expression and stability is analyzed. The successful mutant GPCR can be used for subsequent structural studies. For the GPCR neurotensin receptor 1 (NTR1) this method increased the expression over WT NTR1 by an order of magnitude (60).

1.2.7 Expression platform

The selection of an appropriate expression system is important for the adequate expression of membrane proteins, but remains largely empirical (61). Prokaryotic membrane proteins can often be recombinantly expressed in sufficient yields using *E. coli* expression systems. Eukaryotic membrane proteins, however, are more difficult to express due to required lipids, chaperones, and post-translational modifications (62), and often require the use of yeast (63-66), human embryonic kidney cells (HEK) (67-69), or insect cells (25, 27, 70, 71). Different expression systems must often be tested before a system that produces suitable protein yields is determined (72). Manipulation of the DNA, however, is often more fruitful than changing expression systems because it is time-consuming and expensive to set up and optimize new systems.

1.3 MEMBRANE PROTEIN PURIFICATION

Protein purification is the second major impediment to the crystallization of membrane proteins. The same optimizations done for the purification of soluble proteins must also be considered with membrane proteins and are not covered here.

1.3.1 Detergent selection for membrane solubilization and purification

Hundreds of detergents are commercially available, but the detergents used for membrane protein studies must retain both structure and function of the protein, and a suitable detergent is generally determined by trial and error. Further considerations must be made for detergents used in crystallography, as detergents undergo their own phase transition during the vapor diffusion process. Detergents also form micelles which, depending on micelle size, could encompass the few polar residues in the membrane protein that are available for forming crystal contacts. The detergents most successful for the crystallization of membrane proteins are *N,N*-dimethyldodecylamine-*N*-oxide (LDAO), *n*-octyl- β -D-glucopyranoside (OG), octyltetraoxyethylene (C8E4), *n*-decyl- β -D-maltopyranoside (DM) and *n*-dodecyl- β -D-maltopyranoside (DDM) (73), but some proteins are unstable in the most popular detergents (73, 74). Each detergent must be screened for protein stability. One must also consider and monitor the amount of detergent in the sample, as excess detergent can lead to ligand dissociation and phase separation in the crystallization drop (75) (see Chapter 2).

1.4 MEMBRANE PROTEIN CRYSTALLIZATION

Even after the successful expression and purification of stable protein, crystallization can be quite difficult. The following sections outline some methods that have been used during the crystallization phase to obtain quality crystals of the protein of interest.

1.4.1 Limited proteolysis

Similar to the removal of disordered regions in the molecular biology phase, limited proteolysis removes disordered regions after protein purification. Cleavage trials with low levels of several proteases (e.g. trypsin, V8 protease, papain, thermolysin, and subtilisin (76)) are performed, and the samples are run on SDS-PAGE so the level of protein digestion can be analyzed. Once an appropriate protease and protease concentration are determined, the protease is added to the purified protein sample either prior to crystallization trials or within the crystallization drop. Disordered and flexible regions are then cleaved by the protease, allowing for the crystallization of core protein. Limited proteolysis was utilized to determine the core of UlaA, a vitamin C transporter (21). The full-length protein readily formed crystals, but the diffraction limit was about 10 Å. In contrast, the truncated protein, which was missing 7 amino acids from the C-terminus and the octahistidine tag, formed crystals that diffracted to 1.65 Å. Several other examples have been published (76-78).

1.4.2 Lipid cubic phase (LCP)

An alternative to the crystallization of membrane proteins in a vapor diffusion environment containing detergent is crystallization in the presence of a lipid bilayer system (24, 25). So called *in meso* crystallization employs a spontaneously formed cubic phase containing single lipid bilayers and aqueous channels that extend in three-dimensions (79). The target protein is incorporated into the lipid phase, and crystallization cocktails are added. The protein molecules diffuse through the membrane, allowing crystal nucleation to begin. Several recent examples of membrane proteins

crystallized in lipidic cubic phase have been published (70, 80-83), and excellent reviews have been written on the theory and use of LCP (79, 84-86).

1.4.3 Ligand or other binding partner

Co-crystallization of a membrane protein with a known binding partner, for example a soluble protein or a high affinity small molecule ligand, can help stabilize the protein of interest and reduce conformational heterogeneity (24), which generally aid crystallization. Numerous membrane protein structures have been solved with a bound ligand (for example, (21, 87-91)).

1.4.4 Non-covalent crystallization chaperones

The category of non-covalent crystallization chaperones includes macromolecules used to aid in the crystallization of a difficult protein by increasing available hydrophilic area available for forming crystal contacts. In addition to the methods listed below, covalent crystallization chaperones are also in this category, but must be included at the gene level (see Section 1.2.5).

1.4.4.1 Protein specific Fab, scFv, and nanobodies

To increase the hydrophilic residues and surface area available for forming crystal contacts and possibly lock the protein into a specific conformation, antibody fragments are bound to membrane proteins, usually over size exclusion chromatography, prior to crystallization (24, 92). To generate antibody fragments specific for the protein of interest, hybridoma technology (93, 94) or phage display (95, 96) is used. This technique has had much success for the crystallization of membrane proteins using Fab fragments (97-101), single chain variable fragments (scFv) (101-103), and 13 kDa fragments derived from immunized llamas devoid of the light chain (nanobodies) (104-106) as

crystallization chaperones. Though antibody fragment production has become somewhat routine, the procedure is still costly, labor-intensive, and time-consuming (24, 107, 108), and a new antibody fragment must be generated for each target protein.

1.4.4.2 Ankyrin repeat proteins

Directed evolution strategies were used by the Plückthun group to evolve Designed Ankyrin Repeat Proteins (DARPin) (109, 110). DARPins are made up of ankyrin repeats, a 33-amino acid sequence with secondary structure elements including β -turn, two antiparallel α -helices, and a loop to link connecting repeats (111-113). The repeats form a crescent-shaped tertiary structure that binds a properly folded epitope (111-114). Similar to protein specific antibody fragments, a new DARPin is created for each target protein, which can be done using ribosome or phage display. The main advantage of DARPins over protein-specific antibody fragments are the expression levels that can be achieved in *E. coli* (>100 mg DARPin per liter of cell culture). DARPins have been generated for several membrane protein targets (109, 110), and their utility for crystallization has been shown using model membrane protein multidrug efflux pump subunit AcrB (115).

1.4.4.3 Epitope specific antibody fragments

Another type of non-covalent crystallization chaperone is an epitope-specific antibody fragment. A stable, easily expressed and crystallized antibody fragment that has affinity for a short epitope is generated and can then be used for complexation with any epitope-tagged protein. The epitope is introduced into the target protein easily using site-directed mutagenesis (SDM). In contrast to protein-specific crystallization chaperones, epitope-specific antibody fragments do not need to be generated for each target protein.

To my knowledge, the only epitope-specific antibody fragment that has been successful in co-crystallization is a Fab fragment that recognizes a specific RNA motif (116, 117), though other attempts have been made. An anti-FLAG Fab fragment was made by truncating the Sigma anti-FLAG antibody and attempts of co-crystallization with a FLAG-tagged protein were unsuccessful (118). By our lab attempts of co-crystallization of the anti-EYMPME (EE) scFv (119), and the anti-EE Fab fragment ((120), see Chapter 3) have also been unsuccessful. The addition of the epitope may cause unintended conformational heterogeneity, making the complex difficult to crystallize (see Chapter 4).

1.4.4.4 Streptavidin as a crystallization chaperone

The streptavidin-biotin interaction is one of the strongest non-covalent interactions in nature, with a dissociation constant (K_D) of 10^{-13} to 10^{-15} M (121), but the tetrameric assembly of streptavidin makes it less attractive for use as a crystallization chaperone. A monomeric streptavidin has been engineered (6, 122), and could be of use as a crystallization chaperone for biotinylated membrane proteins (see Chapter 5).

1.5 AN EXAMPLE: THE STRUCTURE OF HUMAN β 2 ADRENERGIC

RECEPTOR

The Nobel Prize-winning (Brian Kobilka, Nobel Prize in Chemistry, 2012) structure of human β 2AR coupled to a G protein (105) can be used as an impressive example of overcoming the challenges to the crystallization of membrane proteins. During the molecular biology and expression phase, the N-terminus of β 2AR was truncated (see Section 1.2.2), and T4L was fused to increase solubility and crystallization propensity (see Section 1.2.5). Two surface methionine residues were replaced with threonine residues, increasing the protein expression, and N187 was mutated to glutamate

to remove the glycosylation site (see Section 1.2.6). The protein was expressed in Sf9 insect cells (see Section 1.2.7). For the protein purification phase, DDM was used for membrane solubilization and over 50 detergents were screened for the stabilization of the GPCR-G protein complex (see Section 1.3.1). For crystallization, a nanobody was generated to stabilize the GPCR-G protein interaction (see Section 1.4.4.1). A high-affinity agonist, BI-167107, was added to the protein to restrict conformational heterogeneity (see Section 1.4.3), and the complex was crystallized in lipidic cubic phase (see Section 1.4.2). **Figure 1** depicts some of the methods used. As shown by this example, it is sometimes necessary to use multiple techniques to obtain a high-resolution crystal structure of a difficult protein.

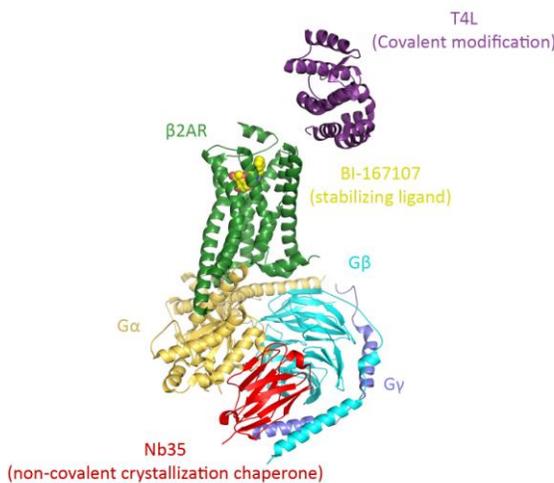


Figure 1 The structure of β 2AR complexed with its G protein as an example of techniques used to crystallize membrane proteins

β 2AR (green) was fused to T4L (purple) at the N-terminus. A nanobody (red) was used to stabilize the G protein (yellow-orange, cyan, and blue). A high-affinity ligand (yellow) was used to stabilize a specific protein conformation. PDB ID 3SN6

1.6 THESIS OBJECTIVES

1.6.1 Development of an antibody fragment and monomeric streptavidin as crystallization chaperones for tagged membrane proteins.

As noted extensively in this introduction, membrane proteins are difficult to crystallize and require special modifications and considerations. The objective of this work was to first prepare the membrane protein of interest, signal peptide peptidase (SPP), for crystallization by developing a protocol to express and purify SPP with adequate protein yield and stability (Chapter 2). The protocol encompasses the molecular biology used to construct the plasmid, protein expression, and purification. Detergents used for protein purification were screened using size exclusion chromatography and circular dichroism, and detergent concentration is measured using thin layer chromatography (TLC).

The next objective was to generate an epitope-specific Fab antibody fragment for general use as a crystallization chaperone for membrane proteins (Chapter 3). Very few examples of epitope-specific crystallization chaperones exist, and the generation of a stable, easily expressed and crystallized anti-EE Fab fragment (Fab/EE) could aid in the rapid solution of membrane protein structures. This chapter details the process of generating the anti-EE Fab fragment, the expression, and purification. EE-tagged soluble proteins were used as test proteins to show that Fab/EE binds EE-tagged proteins by surface plasmon resonance (SPR) and size exclusion chromatography (SEC).

Following the optimization of SPP and production of Fab/EE, the next objective was to obtain a co-crystal structure of SPP in complex with the Fab/EE crystallization chaperone (Chapter 4). Though I was ultimately unsuccessful with this objective, significant progress towards using Fab/EE as a crystallization chaperone for membrane

proteins was made. The ability of Fab/EE to form solution complexes with several EE-tagged membrane proteins is shown. Molecular dynamic (MD) simulations were run which highlight the difficulties and drawbacks in the use of epitope-specific crystallization chaperones. Implications for the solution of membrane protein structures are discussed.

A similar but unrelated objective to those above was to successfully use a monomeric version of streptavidin as a crystallization chaperone for enzymatically biotinylated proteins (Chapter 5). The test protein, MBP, was enzymatically biotinylated by biotin ligase (BirA) and a complex was formed with monomeric streptavidin over SEC. Though ultimately no co-crystal structure was produced, crystallization trials of the complex and lessons learned are discussed.

1.6.2 Molecular basis of novel 1-deoxygalactonojirimycin arylthiourea pharmacological chaperones binding to α -galactosidase A for the treatment of Fabry disease

Chapter 6 is unrelated to the previous chapters. The objective was to use X-ray crystallography to determine the orientation and hydrogen bonding interactions of designed pharmacological chaperones for the treatment of Fabry disease. The chapter details the design, synthesis, and testing of the novel chaperones. Current treatments for Fabry disease include enzyme replacement therapy to replace mutated α -galactosidase A (α -gal A) and the use of 1-deoxygalactonojirimycin (DGJ) as a pharmacological chaperone to stabilize the mutated protein and allow for trafficking to the lysosome. DGJ is largely hydrophilic, which limits the diffusion through the cell membrane. To increase diffusion, new pharmacological chaperones were designed with aromatic arylthiourea

substitutions to DGJ (DGJ-ArTs). The objective of Chapter 6 is to detail the design of DGJ-ArT pharmacological chaperones, test them in Fabry disease cells, and determine the structure of α -gal A in complex with one of the designed DGJ-ArTs.

CHAPTER 2: EXPRESSION, PURIFICATION, AND DETERGENT OPTIMIZATION OF SIGNAL PEPTIDE PEPTIDASE: LESSONS FOR PREPARING MEMBRANE PROTEINS FOR STRUCTURAL CHARACTERIZATION

2.1 INTRODUCTION

As noted in Chapter 1, membrane protein structures lag far behind their soluble counterparts, in large part due to difficulty in protein expression (123). Mammalian membrane proteins often need to be expressed in eukaryotic expression systems because the proteins of interest require chaperones, specific lipids, and post translational modification for proper folding (123, 124). Eukaryotic expression systems are often costly, time consuming, and relatively low yielding (20). In contrast, bacterial expression systems require simple growth medium, have rapid cell growth, and can express membrane proteins in high abundance (20). This makes the archaeal (or bacterial, if available) orthologs much more attractive to pursue for structural studies, a strategy that has been met with considerable success (See (4, 125-127), for example).

2.1.1 Biology of SPP

SPP is an intramembrane aspartyl protease (IAP) with orthologs found in human to extremophilic archaea (128). All IAP family members have 9 transmembrane helices and share a conserved, membrane-embedded signature motif, YD in transmembrane helix (TM) 6 and GXGD in TM7, where X is any amino acid (129).

In eukaryotes, following the cleavage of a the signal peptide of a nascent protein by signal peptidase, SPP uses two aspartate residues to cleave type-2 signal peptides from the endoplasmic reticulum (ER) membrane (**Figure 2, (5)**), and the remnant short peptides are then either degraded as shown for the signal peptide of the protein Crumbs in

Drosophila (130), detoxified as shown by cleavage of the signal peptide of eosinophil cationic protein by human SPP in a human carcinoma cell line (131), or act as signaling molecules for cell-cell communication, as has been demonstrated for MHC class I (132) and HLA-E molecules (133, 134). Human SPP is also responsible for the processing of the N-terminal core domain of hepatitis C virus, a step required for efficient propagation of the virus (5).

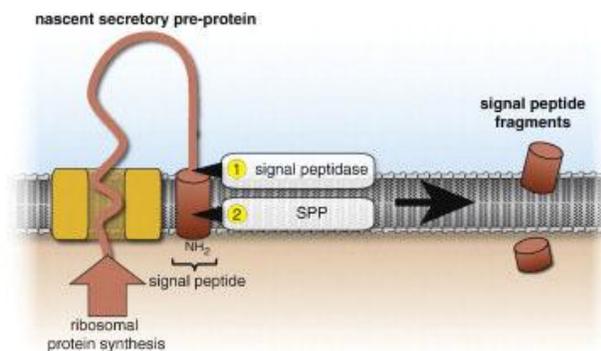


Figure 2 Cleavage action of SPP in membrane of the ER

After protein synthesis on the ribosome, the ER-targeting signal peptide of the nascent protein is cleaved by signal peptidase (1). The signal peptide remaining in the membrane is then cleaved by signal peptide peptidase. Figure was reprinted with permission from (5).

Presenilin, the protease mentioned in Chapter 1 for its role in Alzheimer's disease, is similar to SPP in structure, with opposite topology. Presenilin and SPP are inhibited by some of the same active-site directed molecules, indicating that the active sites are similar (135). Unlike SPP, presenilin requires self-cleavage into two fragments and several cofactors for activity (136), making expression, purification, and crystallization of presenilin very difficult. Because of its role in the cleavage of diverse signal peptides, maturation of hepatitis C virus, and similarity to presenilin without some of the challenges, SPP is an attractive target for structural studies.

2.1.2 Chapter Overview

In this chapter, I present the protocol developed for orthologs of SPP, which is easily adapted for other α -helical membrane proteins. Orthologs of human SPP have similar inhibition profiles (135, 137-139) and cleavage patterns (135), all which strongly suggest these enzymes share a similar structure and utilize a common chemical mechanism. Unlike human SPP, archaeal SPPs are not glycosylated and are active after overexpression and purification in heterologous bacterial hosts, though their biological function is unknown (4, 140). Our approach to the expression and purification of SPP has been to express archaeal orthologs in *E. coli* using a pET vector encoding a C-terminal hexahistidine tag, isolate membrane, solubilize in DDM detergent, purify using nickel affinity chromatography, and further polish the sample purity using size exclusion chromatography. We then screen detergents for enhanced protein stability using circular dichroism because protein stability correlates with crystallizability (141), and proteins in stabilizing solutions maintain monodispersity longer (142). By optimizing such conditions, we allow for more time to perform activity assays and crystallization trials. We are also concerned with the amount of detergent in our final purified sample, as excess detergent can lead to ligand and subunit dissociation and phase separation in the crystallization drop (75). Here, we outline the work flow developed and optimized for our lab to determine ideal buffer conditions, detergents and their concentration for structural studies. Our methods were developed with relatively limited resources and thus can be, at best, considered medium-throughput. Wherever possible, notes for troubleshooting and alternative methods have been included in the discussion.

2.1.3 Individual contributions to the work

Professor Raquel Lieberman optimized the original SPP purification protocol, including membrane solubilization and Ni²⁺-affinity chromatography. Sibel Kalyoncu and I cloned *Methanoculleus marisnigri* SPP from genomic DNA and optimized the purification protocol including membrane solubilization, Ni²⁺-affinity chromatography, and gel filtration. The design of the workflow to determine appropriate detergents and detergent concentrations was a collaborative effort between Dr. Lieberman, Sibel, and myself. All detergent screening by gel filtration and CD data were performed by either Sibel or me, as we split up the numerous constructs for testing. The TLC protocol was optimized by me, and Sibel used the TLC assay for testing several constructs to determine final detergent concentration.

2.1.4 Publication resulting from this work

This protocol will be published as a chapter in Methods in Molecular Biology volume titled Heterologous Expression of Membrane Proteins: Methods and Protocols (143).

2.2 METHODS/PROTOCOL

2.2.1 Signal peptide peptidase ortholog selection

The Domain Enhanced Lookup Time Accelerated (DELTA) feature of Basic Local Alignment Search Tool for proteins (BLASTp (144), <http://blast.ncbi.nlm.nih.gov/>) was used to identify prokaryotic SPPs by excluding eukaryotes (taxid: 2759).

2.2.2 Molecular biology for signal peptide peptidase orthologs

Haloarcula marismortui (*H. mar*), *Halobacterium salinarum* (*H. sal*), and *Methanoculleus marisnigri* (*M. mar*) genomes were purchased from ATCC (<http://www.atcc.org/>). Online signal sequence prediction programs, Signal-3L (145) and

SignalP 4.0 (146) were used to predict the presence of a signal sequence. The genes of the target proteins (**Figure 3**) with the addition of restriction sites *NcoI* and *SalI* were amplified by PCR and ligated into pET-22b(+) vector. To some constructs, a cleavage site for tobacco etch virus (TEV) protease (ENLYFQS) was inserted between the SPP gene and the hexahistidine tag using primers listed in **Table 1**. Gene sequence fidelity is confirmed by DNA sequencing (MWG Operon, www.operon.com).

2.2.3 TEV protease expression and purification

Though TEV protease can be purchased, we have found that it is quite easy to prepare it in-house. The S219V mutant TEV protease is expressed in *E. coli* as a MBP fusion with a TEV protease site in the linker region between the two proteins, and is purified as previously described (147). The protein is stored in phosphate buffered saline (PBS) at roughly 1.6 mg/mL at -80 °C until needed.

2.2.4 Signal peptide peptidase expression

After transformation into *E. coli* Rosetta 2(DE3), a single colony is added into a 200 mL Luria-Burtani (LB) supplemented with 60 µg/mL ampicillin and 34 µg/mL chloramphenicol. The starter culture is incubated for 18-20 hours at 37 °C with shaking at 225 RPM. The next day, 1 L LB containing 60 µg/mL ampicillin and 34 µg/mL chloramphenicol is inoculated with 10 mL of the starter culture. Cultures are incubated at 37 °C and 225 RPM until optical density at 600 nm (OD₆₀₀) reached 0.6-0.8, at which point the temperature is reduced to 18 °C and incubated for 1 hour before inducing protein expression by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Protein expression proceeds at 18 °C for 16-20 hours, after which cells are

harvested by centrifugation at 2500 x g for 10 minutes before flash freezing in liquid nitrogen for storage at -80 °C.

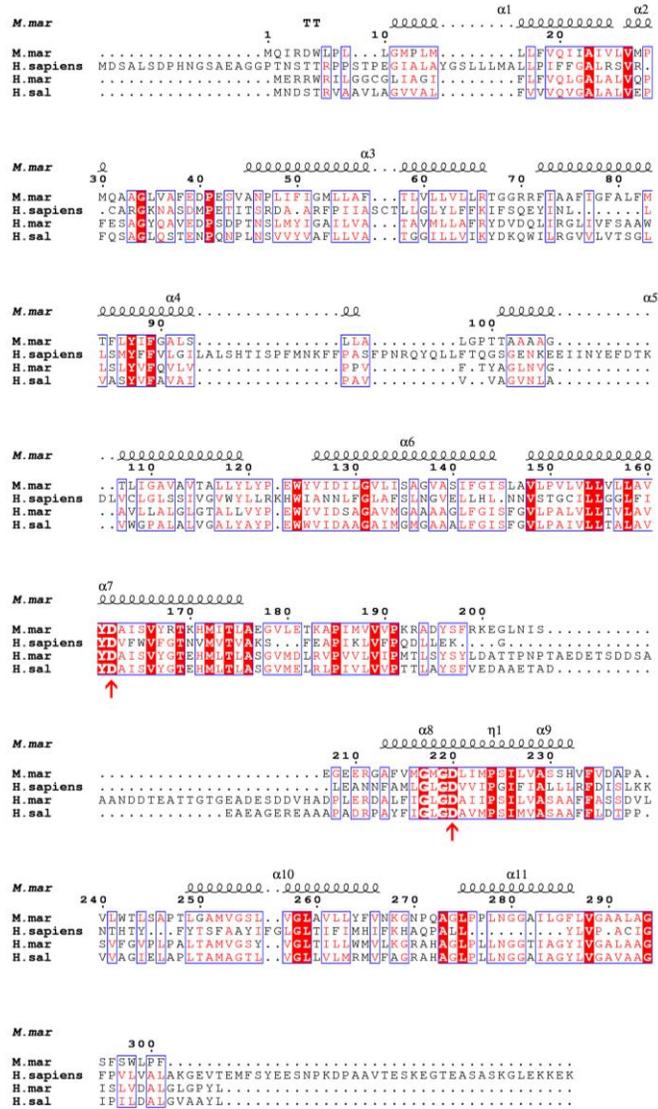


Figure 3 Alignment of human SPP with 3 archaeal SPP orthologs

Alignment was performed using Clustal Omega and rendered in ESPrpt (1). Identical residues in all four sequences are white with red background. Similar residues are in red and conserved patches are boxed in blue. The conserved motif of all IAP family members is in bold white, with the catalytic aspartate residues denoted with a red arrow. α -helices are marked by spiral along the top of each row and are based on the crystal structure of *M. mar* SPP (PDB ID 4HYC).

Table 1 Primers used for the insertion of a TEV cleavage site in *M. mar* SPP

	Forward primer	Reverse Primer
Insertion of TEV cleavage site by SDM	Round 1	Round 1
	GTTTTCGTGGCTTCCTTTCGTCGACGAG	GGTGGTGGTGGTGCTCGAGTGCGTAC
	AATCTGTACGCACTCGAGCACCACCACC	AGATTCTCGTCGACGAAAGGAAGCCA
	ACC	CGAAAAC
	Round 2	Round 2
	CGTTTTCGTGGCTTCCTTTCGTCGACGAG	GCCGGATCTCAGTGGTGGTGGTGGTG
	AATCTGTACTTCCAGTCTGCACTCGAGC	GTGCTCGAGTGCAGACTGGAAGTACA
	ACCACCACCACCACCCTGAGATCCGGC	GATTCTCGTCGACGAAAGGAAGCCAC
		GAAAACG

2.2.5 Membrane isolation from harvested cells

A 7-8 g mass of cell pellet is resuspended on ice in 25-30 mL cell lysis buffer (50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 200 mM NaCl, Roche Complete ethylenediaminetetraacetic acid (EDTA) free protease inhibitor cocktail) and lysed by French press at 18,000 PSI cell pressure. Cellular debris is pelleted by centrifugation (5000 x g for 15 minutes) at 4 °C. The supernatant is placed in a new centrifuge tube and the centrifugation step is repeated until no discernible pellet remained. The supernatant from the prior step is combined and subjected to ultracentrifugation at 120,000 x g for 45 minutes at 4 °C to pellet membrane. Membrane is then placed in 7 mL Dounce homogenizer and resuspended in 7 mL cell lysis buffer (without protease inhibitor). Sample is subjected to ultracentrifugation again, and pellet is added to a microcentrifuge tube and flash frozen in liquid nitrogen for storage at -80 °C.

2.2.6 Protein solubilization from membrane and purification by Ni²⁺-affinity chromatography

Approximately 0.3 to 1 g of wet membrane is thawed and resuspended in 7 mL membrane resuspension buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole) in a 7 mL Dounce homogenizer. Separately, a 1% solution of DDM or foscholine 12 (FC12) in membrane resuspension buffer is prepared with sufficient volume to

make a 1% (w/v) solution of solubilized membrane; e.g. 1 g DDM is dissolved in 93 mL membrane resuspension buffer, and added on ice to 1 g wet membrane resuspended in 7 mL buffer. The membrane solution is stirred at 4 °C for at least 1 hour. The sample is then centrifuged at 181,000 x g for 45 minutes at 4 °C to remove unsolubilized material. For purification, the solution is loaded into an appropriate superloop connected to an AKTA fast protein liquid chromatography (FPLC) instrument (GE Healthcare) or other similar automated purification instrument. The protein is purified by Ni²⁺-affinity chromatography using a 1 mL Ni²⁺-affinity column (GE Healthcare) equilibrated with 50 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole, and 0.1% DDM (Buffer A) and eluted with an imidazole gradient by mixing with 50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole and 0.1% DDM (Buffer B). Peak elution fractions are pooled and concentrated in a 10K molecular weight cut-off (MWCO) Amicon Ultra centrifugal filter.

Prior to SEC (see next section), and for constructs containing a cleavable hexahistidine tag, the tag is removed by incubation with TEV protease. The protein sample is exchanged into gel filtration buffer (50 mM HEPES pH 7.5, 200 mM NaCl, 0.0174% DDM) before the addition of 1:1 TEV protease:SPP mass ratio. The TEV protease/SPP mixture is incubated at 4 °C for 16 to 20 hours to complete the cleavage reaction. The sample is repurified by Ni²⁺-affinity chromatography using the same buffers, this time collecting the flowthrough material; the elution fractions contain the uncleaved protein and TEV protease.

2.2.7 Protein stability in different detergents assessed by size exclusion chromatography

An overview of our procedure for using gel filtration and circular dichroism (CD) to measure protein stability in different detergents is presented in **Figure 4**. Following Ni^{2+} -affinity chromatography described above, the sample is divided into 6 aliquots of approximately 250 μL each for testing each of 6 different detergents. In turn, each sample is injected on a Superose 12 10/300 column equilibrated with 2 column volumes of gel filtration buffer supplemented with 2X the critical micelle concentration (CMC) of a selected detergent. In the examples presented here, FC12, DDM, DM, 5-cyclohexyl-1-pentyl- β -D-maltoside (Cy-5), LDAO, 2,2-didecylpropane-1,3-bis- β -D-maltopyranoside or lauryl maltose neopentyl glycol (LMNG), were tested (see **Figure 5** for structures of detergent molecules). Initial assessment of protein stability includes inspection of elution peak intensity and shape. Protein detergent complex (PDC) size is evaluated by retention volume and purity is assessed using unboiled samples on SDS-PAGE (148).

Phase I: Test protein stability in different detergents

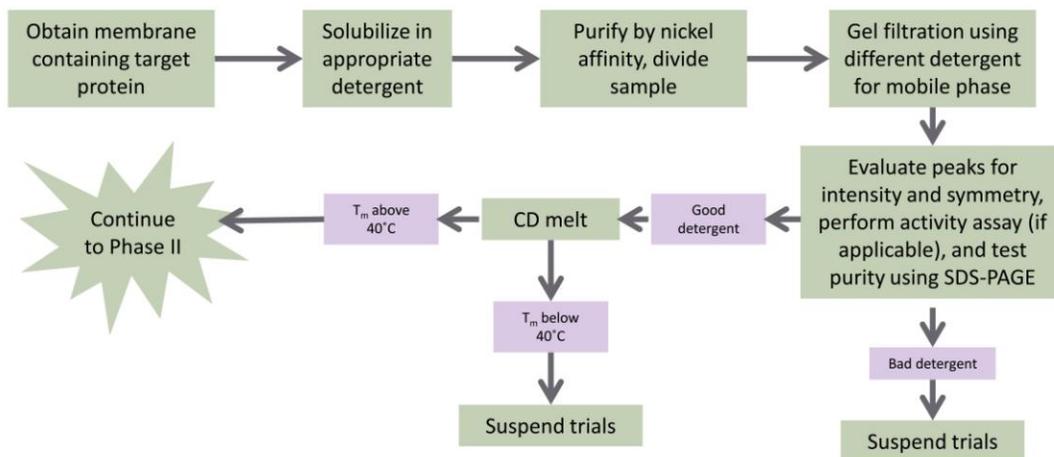


Figure 4 Workflow to determine protein stability in different detergents using circular dichroism.

See **Figure 6** for Phase II.

2.2.8 Protein stability in different detergents by CD

To assess protein thermal stability, peak elution fractions from each Superose 12 purification are concentrated to 8 μM in 10K MWCO Amicon Ultra centrifugal filter. Concentration is measured using molecular weight (MW) and extinction coefficient calculated by ExPASy ProtParam (149), and absorbance measured using a Nanodrop 2000. A CD thermal unfolding experiment is conducted on each sample using a JASCO CD spectropolarimeter and a CD cell of 0.1 cm in width. CD parameters used are as follows: sensitivity=standard, start wavelength=300 nm, end wavelength=200 nm, data pitch=1, scanning mode=continuous, scanning speed=500nm/min, response time=1 second, bandwidth=1 nm, accumulation=10, temperature increment=2 $^{\circ}\text{C}$. The temperature versus the normalized molar ellipticity at the minimum wavelength is plotted in GraphPad Prism to determine the melting temperature (T_m) for each sample using Boltzmann sigmoid equation.

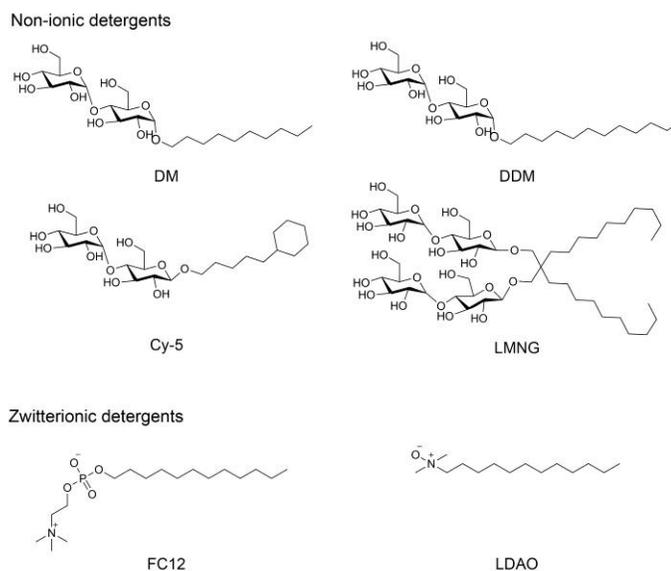


Figure 5 Structures of detergents used for SEC and CD experiments

Structures were drawn using ChemBioDraw Ultra according to the structures on www.anatrace.com.

2.2.9 Detergent quantification by TLC

An overview of our procedure for using TLC for detergent quantification is presented in Figure 6. Detergent standards in gel filtration buffer (0.25%, 0.5%, 1%, 2%, 3% detergent of interest) are prepared. After SEC in a detergent of interest, 500 μL of the sample is concentrated in an Amicon Ultra centrifugal filter with one of four different MWCOs, 10K, 30K, 50K, 100K until each filter contains less than 50 μL . The filtrate of each filter is then placed in a new 10K MWCO Amicon Ultra centrifugal filter and centrifuged at 8,500 \times g until less than 50 μL remains. The total volume and the concentration of each sample is measured using a nanodrop and the protein extinction coefficient calculated by ExPASy ProtParam (149).

Detergent standards (5 μL), concentrated samples, and the concentrated filtrate are spotted on a silica 60 TLC plate approximately 1 inch from the bottom of the plate, and the plate is allowed to dry for at least 30 minutes. The plate is then placed in a TLC chamber containing 0.5 inch TLC solvent (63:35 chloroform:methanol) until the solvent has run at least halfway up the plate. The TLC plate is dried in open air for 5 minutes, then placed into a chamber containing iodine chips until detergent samples are readily visible. The plate is imaged using a desktop scanner and analyzed using either Photoshop or ImageQuant.

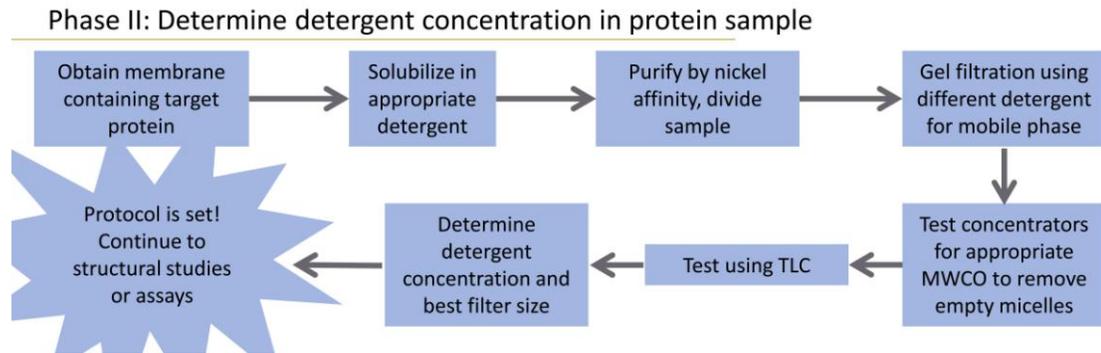


Figure 6 Workflow to determine optimal filter MWCO for concentrating protein samples

TLC is used to determine detergent concentration in the samples.

2.3 RESULTS

2.3.1 Molecular biology for target membrane protein

Figure 3 shows the final Clustal Omega (150) alignment between human SPP and our selected targets, SPP from archaeal *H. mar*, *H. sal*, and *M. mar* rendered in ESPript (1) with secondary structure information from the crystal structure of *M. mar* SPP (PDB ID 4HYC) (4).

pET-22b(+) was chosen as the expression plasmid because it contains N-terminal pelB leader sequence and C-terminal hexahistidine tag. To ensure correct protein insertion into the membrane, the protein sequence was analyzed by signal sequence prediction software, Signal-3L (145) and SignalP 4.1 (146) (**Figure 7**). Signal-3L predicted that the first 23 amino acids of *M. mar* SPP are a signal sequence, but SignalP 4.1 predicted that *M. mar* SPP contained no signal sequence. Both the truncated (missing the predicted signal sequence) and full length *M. mar* constructs were prepared in pET-22b(+) vector using *SalI* and *NcoI* restrictions sites with a TEV cleavage site inserted between the protein and the hexahistidine tag (**Figure 8**).

a --- (a) -----Prediction Results-----

According to Signal-3L engine for your selected species, the signal peptide is: 1-23

MQIRDWLP LLGMPLMLLFVQIIAIVLVMPMQAAGLVAFEDPESVANPLIFIGMLLAFTLV
LLVLLRTGGR

(b)

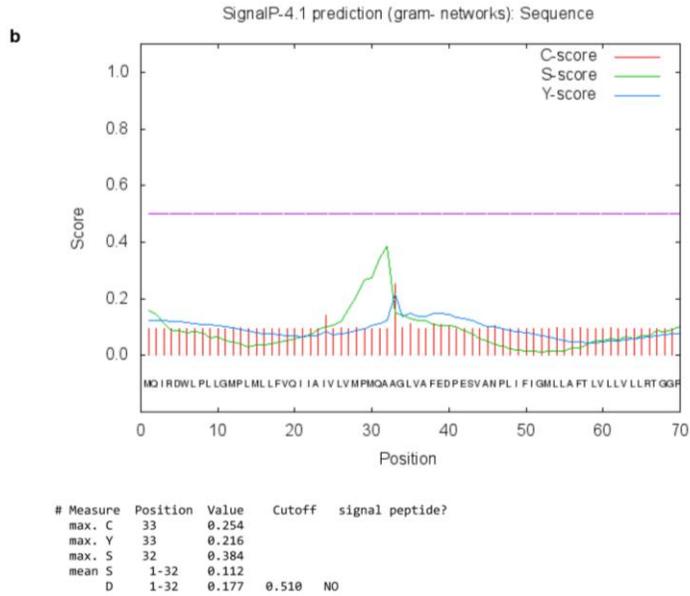


Figure 7 Signal-3L and SignalP signal sequence prediction for *M. mar* SPP

(a) Output from Signal-3L predicting that the first 27 amino acids (red) from *M. mar* SPP are a signal sequence. (b) Output from SignalP predicting the same protein sequence does not contain a signal peptide.



Figure 8 pET-22b(+) DNA and amino acid sequence in the area where the ortholog DNA is inserted.

Areas of interest are marked. The red DNA and amino acid sequence were replaced with the target DNA. The orange amino acids were removed and replaced with TEV cleavage sequence ENLYFQS. Image was adapted from that of Novagen.

2.3.2 Membrane protein expression

All SPP orthologs were successfully expressed in Rosetta 2(DE3) *E. coli* cells using a general expression protocol with reduced temperature and elongated time for protein induction and expression. Typical yield is 2.5 g wet cell mass per liter of LB culture.

2.3.3 Membrane isolation and protein purification

For all SPP orthologs, membrane was first isolated from harvested cells then solubilized using DDM or FC12 as described in Materials and Methods. SPP is one of the most abundant protein in the solubilized sample (**Figure 9**, lane 1). SPP is then purified using Ni²⁺-affinity chromatography (**Figure 10**). The protein of interest elutes with 100-165 mM imidazole, but contains low levels of several impurities (**Figure 9**, lane 2). After removal of the hexahistidine tag with TEV protease, SPP is slightly smaller, as observed by SDS-PAGE (**Figure 9**, lane 3). When not doing the detergent screen, protein collected in the flowthrough of the Ni²⁺-affinity chromatography step after TEV cleavage is concentrated in a 10K MWCO Amicon Ultra Centrifugal filter and purified using SEC, with final purified protein being in DDM being ~90% pure (**Figure 9**, lane 4).

2.3.4 Protein stability in different detergents

SPP orthologs were subjected to size exclusion chromatography in different detergents to assess the effect of the detergent on protein homogeneity and characteristics of elution volume. For both *H. sal* and *H. mar* SPP, LMNG and LDAO gave elution peaks in the void volume of the column, indicating a size > 2,000,000 Da. FC12 gave the

highest intensity peak for both SPP orthologs. Conversely, the lowest intensity and broadest peak was seen with DDM (**Figure 11**).

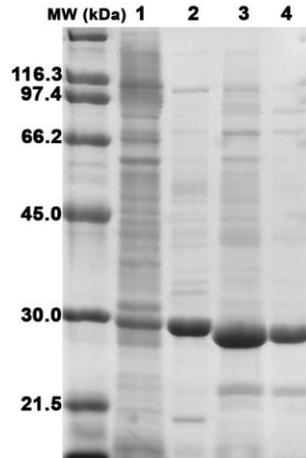


Figure 9 SDS-PAGE of *M. mar* SPP samples during each purification step.

Broad range molecular weight marker with band sizes is shown on the left. *M. mar* SPP is the prominent band just under 30 kDa in each lane. Lane 1 is the sample after membrane solubilization. Lane 2 is the protein sample after HisTrap purification. Lane 3 is the protein purified over HisTrap column a second time after TEV cleavage of the hexahistidine tag. Lane 4 is the purified protein sample after size exclusion chromatography on the Superose 12 column in Gel Filtration Buffer with 0.0174% DDM.

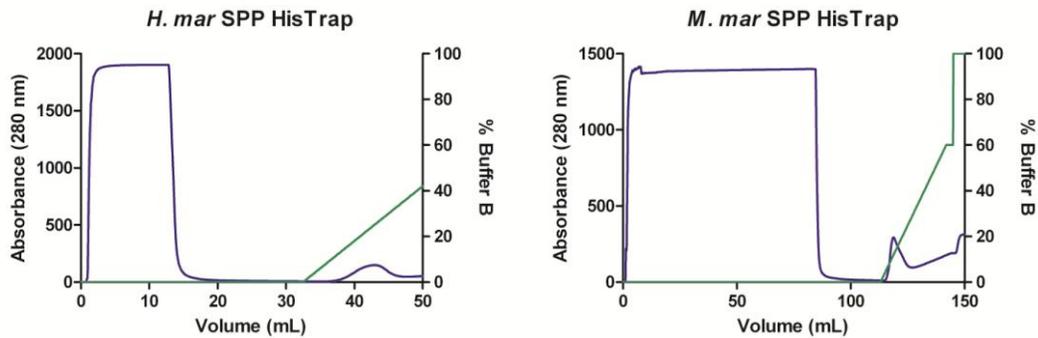


Figure 10 Ni²⁺ affinity chromatographs from purification of two SPP orthologs

Chromatograms were obtained using Unicorn software of *H. mar* SPP (left) and *M. mar* SPP (right). Blue trace is absorbance at 280 nm and green trace is % Buffer B.

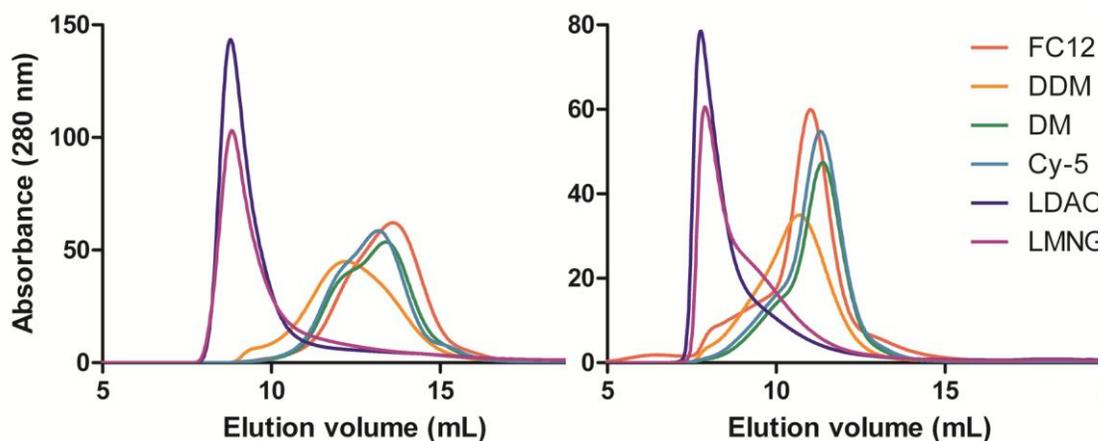


Figure 11 Gel filtration chromatograms of SPP orthologs in 6 detergents

H. sal SPP in 6 detergents (left) and *H. mar* SPP in 6 detergents (right). Detergent abbreviations are as follows. FC12: fos-choline-12, DDM: *n*-dodecyl β -D-maltoside, DM: decyl β -D-maltoside, Cy-5: 5-cyclohexyl-1-pentyl- β -D-maltoside, LDAO: *N,N*-dimethyldodecylamine-*N*-oxide, LMNG: 2,2-didecylpropane-1,3-bis- β -D-maltopyranoside or lauryl maltose neopentyl glycol.

2.3.5 Protein stability in different detergents by circular dichroism

CD thermal unfolding experiments for samples in each detergent were acquired, revealing a range of protein thermal stabilities. For *H. mar* SPP, T_m s range from 79.0° for LDAO to 55.6° for DM (**Figure 12** and **Table 2**). The T_m roughly correlated with the peak intensity from SEC, with the exception of DDM, which was the lowest intensity peak but not the lowest T_m . SPP samples in LMNG and LDAO eluted in the void volume of the SEC (**Figure 11**), and also had the highest T_m (**Table 2**), perhaps due to a stabilized oligomeric state not seen with the other detergents. The micelle size of LDAO (17-21.5 kDa (151, 152)) is significantly smaller than that of DDM (72 kDa (152)), indicating that the decrease in retention volume for the sample with LDAO was not due to micelle size. Interestingly, the T_m s for samples with DM and DDM, which are only different in structure by 2 carbon atoms (**Figure 5**), differ by ~5 °C. Two zwitterionic detergents were tested (FC12 and LDAO, **Figure 5**) and the T_m s differ by 15 °C. Though

DDM gave the lowest intensity peak on size exclusion chromatography and did not give the highest T_m , DDM was chosen for solubilization and crystallization of *M. mar* SPP because DDM is the most common detergent for membrane protein crystallography and SPP crystals grew in DDM (data not shown).

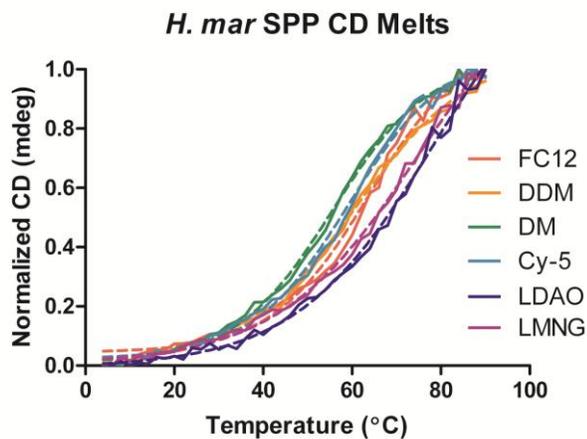


Figure 12 Normalized CD thermal unfolding experiment of *H. mar* SPP in 6 detergents.
The solid trace is the normalized CD data, the dashed trace is the Boltzmann sigmoidal fit.

Table 2 Melting temperatures for *H. mar* SPP in 6 detergents

According to Boltzmann sigmoidal fit from **Figure 12**

Detergent	T_m (°C)
LDAO	79.0
LMNG	75.6
FC12	63.9
DDM	60.7
Cy-5	59.3
DM	55.6

2.3.6 Detergent quantification

TLC was performed to quantify the detergent levels in the SPP samples after purification. The amount of detergent present in the sample decreased with increasing

Amicon Ultra centrifugal filter size. Because the micelle of DDM alone is about 72 kDa (152), the 10K MWCO filter likely concentrated all of the excess detergent while the 100K MWCO filter allowed some detergent to pass through the filter. All protein and detergent concentration values are listed in **Table 3**. Though the gel filtration buffer contained 2X CMC (0.0174%) DDM, the calculated starting detergent concentration of the sample that was concentrated using the 10K MWCO centrifugal filter was 0.0275%, indicating that more detergent was present in the sample than we expected.

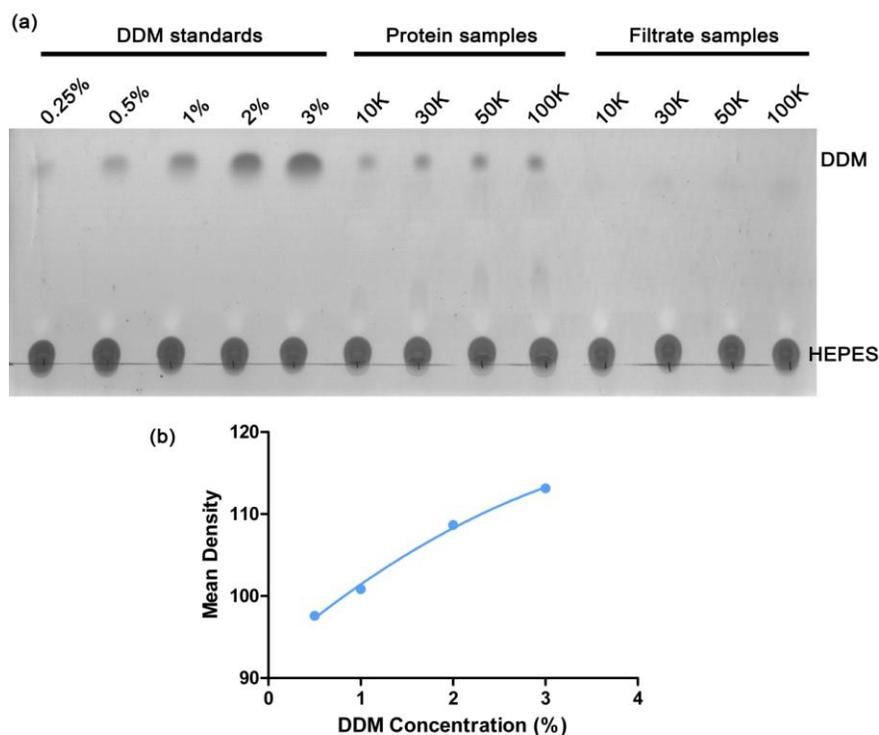


Figure 13 Detergent quantification of SPP sample using TLC

(a) TLC from detergent filter test of *M. mar* SPP in 20 mM HEPES pH 7.5, 200 mM NaCl, 0.0174% DDM. (b) Standard curve prepared using mean density values from Photoshop for each standard DDM concentration. $R^2 = 0.9960$.

Table 3 Data table for TLC detergent test

Filter MWCO	10K	30K	50K	100K
Starting concentration (mg/mL)	0.054	0.054	0.054	0.054
Volume after concentrating (μL)	26	17.5	14	9
Concentration factor ^a	19.2	28.6	35.7	55.6
Theoretical concentration (mg/mL) ^b	1.038	1.543	1.929	3.000
Measured concentration (mg/mL) ^c	0.995	1.296	1.388	1.848
Protein recovery (%)^d	95.8	84.0	72.0	61.6
Theoretical [detergent] ^e (%)	0.335	0.497	0.621	0.967
Actual [detergent] ^f (%)	0.529	0.687	0.771	0.885
Detergent recovery (%) ^g	1.58	1.38	1.24	0.92
Detergent recovery using 0.0275% as starting detergent concentration^h (%)	100.0	87.5	78.5	57.9

^a Concentration factor = 500 μL/(volume after concentrating)

^b Theoretical concentration = (Starting concentration)*(Concentration factor)

^c Measured concentration was measured using a nanodrop and the protein extinction coefficient calculated by ExPASy ProtParam (149).

^d Protein recovery = (Measured concentration)/(Theoretical concentration)

^e Theoretical detergent concentration = 0.0174*(Concentration factor)

^f Actual detergent concentration was estimated using TLC (Figure 13a) and the standard curve prepared using mean density calculation in Photoshop from DDM standards on TLC (Figure 13b)

^g Detergent recovery = (Theoretical detergent concentration)/(Actual detergent concentration)

^h It is assumed that the 10K filter concentrates all the detergent. Since Actual detergent concentration was larger than Theoretical detergent concentration in most cases, the starting detergent concentration was calculated using the Actual detergent concentration and the Concentration factor for the 10K filter. That detergent concentration (0.0275%) was used to calculate the Detergent recovery.

2.4 DISCUSSION AND LESSONS LEARNED

In this chapter, our procedure for the production of highly stable SPP orthologs for structural characterization has been outlined. We clone SPP orthologs into pET-22b(+) vector, express SPP in *E. coli*, purify by Ni²⁺-affinity chromatography and gel filtration. *E. coli* was chosen as the expression organism due to our lab having previous experience with *E. coli* expression systems, low cost, and ease of production of proteins from lower organisms. We chose to study archaeal orthologs of human SPP so that protein could be easily produced in high quantities. Then protein stability in different detergents is tested using CD and detergent concentration is quantified using TLC. Gel

filtration was used to exchange the detergent in the protein sample and peak intensity and shape were used as indicators of protein stability. T_m s were measured by CD thermal unfolding experiment in 6 detergents. The melting temperature of the sample is important as protein stability is related to crystallization propensity (153). Detergent concentration in the sample is important because too much excess detergent has been shown to dissociate ligands and negatively impact crystallization propensity (75).

We began this process with 10 SPP orthologs, and the three SPP orthologs described here have been successfully expressed and purified. The leaky pipeline reinforces the notion that the ability to express and purify membrane proteins in heterologous hosts is highly individualized. Starting with more targets gives a better chance of success in the pipeline. We also prepared the full length and truncated versions of *H. mar* and *M. mar* SPP. Both constructs of *M. mar* SPP were active, had similar expression levels, and extents of crystallizability (data not shown). By contrast, the full length construct of *H. mar* SPP could not be expressed in reasonable quantity, likely due to the addition of the *pelB* leader sequence with a natural signal sequence, which caused incorrect membrane insertion.

For protein expression, we use *E. coli* Rosetta 2(DE3) cells, which contains a plasmid to correct for codon usage differences between the ortholog organism and *E. coli* host; some proteins express well in BL21(DE3) and C43(DE3). Variables affecting protein expression are numerous, and include media, cell line, optical density at induction, induction temperature, and induction duration. For each new protein, small scale expression trials are first conducted to determine the best conditions for high expression. When the C-terminus of the membrane protein is in the cytosol (C_{in}

topology), a new tactic to rapidly assess homolog expression yields is to utilize a membrane protein-green fluorescent protein (GFP) fusion by whole-cell and in-gel fluorescence (154, 155). Other procedures for optimizing the expression of membrane proteins have been published (123, 124, 156) and could be used to increase the throughput of the overall procedure.

For membrane solubilization, several detergents have been tested. LDAO, OG, DM, DDM have been successful for the crystallization of membrane proteins (73) so they are a good starting point for optimizing the solubilization. We chose to start with DDM, as it the most popular detergent for membrane protein crystallization (157). FC12 was used for solubilization of *H. mar* SPP, as purification in DDM did not yield pure protein (not shown). FC12, however, has not been very successful in the crystallization of many membrane proteins (157). As shown by the gel filtration and CD thermal unfolding experiment of *H. mar* SPP in 6 detergents, protein in FC12 is still very stable, though all crystal trials of *H. mar* SPP were unsuccessful in our lab.

Our procedure to determine stabilizing detergents for each construct is a bit cumbersome as each protein must be purified in different detergents and then a CD thermal unfolding experiment must be performed in each detergent. Several other methods have been developed that could increase the throughput. Fluorescent size exclusion chromatography (FSEC) can be used with a GFP-fused protein without the need to purify the fusion protein before testing (154, 155, 158). A fluorescence detector must be in line with the gel filtration column to use this method, but small amounts of fusion protein can be detected. If buried cysteines are present, effects of detergents and additives on purified protein can be measured in a more high-throughput manner by a

thermal unfolding assay using *N*-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide (CPM), a dye that fluoresces upon reaction with a thiol (159-161). Denaturing SDS-PAGE and size exclusion high pressure liquid chromatography (SE-HPLC) has also been used to determine the effect of detergents, pH, additives, and lipids on the proportion of monomeric protein and can be used with minimal protein amounts (142, 162, 163).

Other methods for determining detergent concentration have also been developed. Phenol/sulphuric acid assay works to determine the concentration of sugar based detergents. A molybdate assay can be used for detergents containing phosphate. Nuclear magnetic resonance (NMR) can be used to identify and determine the concentration of detergents, though it takes a large sample (520 μ L) and proteins can interfere with the signal. Gas chromatography coupled with mass spectrometry (GC-MS) can be very sensitive (0.05-3 μ g DDM). Overall though, TLC is very simple to perform in the lab, requires only small amount of sample (5 μ L) and works for many different detergents. All methods are reviewed in Price and Jia, 2013 (164).

With Ni^{2+} -affinity chromatography, it is important to be aware of the possibility of copurification with *E. coli* membrane protein AcrB. Serendipitously, AcrB binds to the nickel resin, is too large to pass through concentration devices, remains through SEC, and is hypercrystallizable even at low levels not detected on SDS-PAGE. Several labs (165), including ours, have crystallized and solved the structure of AcrB instead of our intended membrane protein. It is best to be wary of crystals resembling AcrB (see Figure 14) and to search the PDB for unit cell dimensions of any membrane protein crystals obtained by this affinity purification method to make sure the dimensions do not match those of

AcrB. To remove copurification with AcrB, we have added a TEV cleavage site between the protein and the hexahistidine tag. After purifying SPPs over the Ni²⁺-affinity column, we remove the hexahistidine tag using TEV protease, which can be readily prepared in the laboratory in high yield (147) and is insensitive to the specialized detergents used for membrane protein solubilization and purification. The TEV itself contains a C-terminal histidine tag, so a second purification over the Ni²⁺-affinity column traps TEV, uncleaved SPP, the tags, and AcrB; cleaved SPP is collected in the flowthrough fractions.

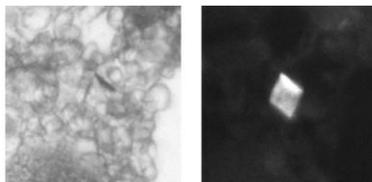


Figure 14 AcrB crystals using visible and UV light

Crystals were obtained using crystallization condition containing 20 mM sodium citrate pH 5.6, 0.1 M NaCl, and 17% polyethylene glycol 3350. The crystals formed within 6 days at room temperature. Space group is R23 and cell dimensions (Å) are a=144 b=144 c=518.36 and angles (°) of $\alpha=90$ $\beta=90$ $\gamma=120$.

The buffer exchange, cleavage, and second Ni²⁺-affinity chromatography steps can also be problematic from the standpoint of detergent concentration, which could hinder crystallization and ligand binding (75). After the first Ni²⁺-affinity chromatography step, the protein is buffer exchanged, concentrating the detergent at least 20 fold. The protein is then cleaved and repurified over Ni²⁺-affinity chromatography. Since the flowthrough is then collected, all of the detergent concentrated in the buffer exchange step passes through the column and is collected with the protein. It is then concentrated ~64 fold (16 mL flowthrough concentrated to 250 μ L for injection onto Superose 12 column), and the empty detergent micelles and the protein elution peak

overlap on the gel filtration column causing the final detergent in the purified sample to be much higher than expected (as shown in **Figure 13** and **Table 3**). In our experiments, we observed co-elution of empty DDM micelles and *M. mar* SPP. To decrease the detergent levels, other gel filtration columns should be tested to separate the detergent peak from the protein peak and different detergents should also be tested because detergent micelle size and protein elution volume (**Figure 11**) can vary greatly. Another option to remove detergent is dialysis, but sample storage stability could limit the utility of this method. For SPP, the activity has been shown to decrease after a week at 4 °C (data not shown), and dialysis could take several weeks to remove a sufficient amount of excess detergent. Our purified *M. mar* SPP does form reproducible crystals, so the excess detergent does not completely block crystallization, but may affect reproducibility of crystals in terms of their diffraction limit.

In this chapter, we have employed some of the principles discussed in Chapter 1 to increase the crystallizability of our target membrane protein, SPP. As evident in Chapter 1, numerous modifications must sometimes be made to the protein of interest in order to obtain quality protein crystals. Numerous crystallization conditions must be tested to determine if these strategies for SPP will lead to a high-resolution crystal structure.

CHAPTER 3: CHARACTERIZATION OF EPITOPE-SPECIFIC FAB ANTIBODY FRAGMENT FOR USE AS A GENERAL CRYSTALLIZATION CHAPERONE

3.1 INTRODUCTION

As noted in Chapter 1, crystallization of a given target protein can be a difficult step in structure determination, and is especially notorious in the case of membrane proteins. Numerous strategies have emerged to enhance the crystallizability of difficult proteins. These include protein modification such as random mutagenesis, directed evolution, or strategic mutations to improve crystallization propensity (166-169), generation or discovery of ligands to stabilize a specific protein conformation (170, 171), protein symmetrization by cross-linking or engineered metal binding sites (172, 173), limited proteolysis to remove flexible regions (76), and surface entropy reduction (174), among others. Recent advances in the crystallization environment for membrane proteins in particular include LCP and lipid-mimicking detergents to stabilize membrane proteins in a native-like environment (175-178). These methods are reviewed in Chapter 1.

For membrane proteins, the so-called crystallization chaperone approach has been particularly successful. The desirable biophysical properties of chaperones include increasing hydrophilic residues available for forming crystal contacts, thus improving the likelihood of obtaining well-ordered crystals of the chaperone-target membrane protein complex. Covalent chaperones have been utilized to crystallize several GPCRs, where ICL3 (44, 179, 180) or the N-terminus (46, 181) is replaced by T4L or BRIL and generated as a chimeric protein. Non-covalent chaperones include antibody fragments, mostly Fab and scFv generated by hybridoma or library screening (102, 182-188), V_{HH}

camlid domains (nanobodies) (104, 187, 189), and DARPs (190, 191). All of these non-covalent binding partners are target specific, namely, for each new protein of interest, a new chaperone must be sought. Previously, we proposed the use of peptide-specific antibody fragments as an alternative to protein-specific crystallization chaperones (24).

3.1.1 3D5 scFv

This project began with the observation that anti-his scFv 3D5 crystallized in a unit cell that did not use the complementarity determining regions (CDRs) in the crystal contacts, but instead the CDRs opened into a 70 Å wide channel that could possibly accommodate a membrane protein (**Figure 15**) (7). Despite the favorable crystal lattice, 3D5 is not an ideal antibody fragment for a crystallization chaperone. The low (micromolar) affinity for C-terminal histidines is not ideal for crystallization because of the pH sensitive binding to histidines. Commercial sparse matrix screens contain a wide variety of pHs which would be incompatible with the 3D5:His-tagged protein complex. 3D5 is also not very soluble in solution and after expression, only 50% of the protein was

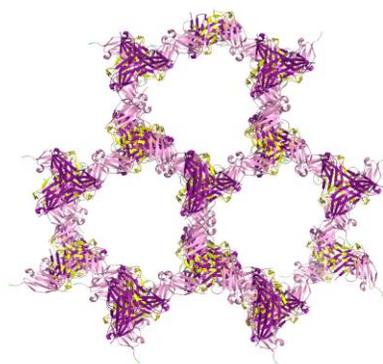


Figure 15 Anti-his scFv 3D5 crystal lattice

3D5 was crystallized in space group $P3_221$ with 2.7 Å resolution. Light chain is colored light purple, heavy chain is dark purple. CDRs are in yellow. The CDRs point in toward a wide solvent channel (7).

in monomeric form (**Table 4**).

3.1.2 Conversion of 3D5 scFv to anti-EE scFv

The Maynard (University of Texas at Austin) and the Lieberman Labs sought to convert the chaperone affinity to anti-EE (EYMPME). The EE epitope is an attractive peptide for chaperone binding because of the chemical diversity within the epitope (hydrophilic glutamates, hydrophobic methionines, restrictive proline), and the availability of commercial antibodies against the epitope for testing purposes. Upon conversion using directed evolution and phage display, the anti-EE scFv (3D5/EE_48 or scFv/EE) has 85% amino acid identity and increased from 50% monomeric protein with 3D5 to 81% monomeric protein with scFv/EE (**Table 4**).

Table 4 Biophysical characteristics of 3D5 compared to scFv variants (192)

Parameter	3D5	scFv/EE
Expression level (mg/L culture)	3.1	2.1
Solubility (mg/mL)	2.3	12.8
Melting temperature (°C)	46.5 ± 0.5	47.2 ± 0.3
% AA identity	100	85
% monomeric protein ^a	50	81
K _d (nM), 14B7-His ₆ (χ^2)	4700 (0.08)	ND ^b
K _d (nM), MBP-EE (χ^2)	ND ^b	767 (0.03)

^a % monomer measured upon initial purification

^b ND, not detected

Co-crystallization was attempted with scFv/EE and its variants, with no success (193, 194). To the best of our knowledge, the only successful example of a target independent non-covalent chaperone has involved the use of a Fab fragment that recognizes a portable small structural RNA element to crystallize a ribozyme (117, 195).

Co-crystallization with client membrane proteins has been attempted by others using Fab fragments generated from commercially-available FLAG-binding monoclonal antibodies (118), but without documented success to date. Analysis of deposited cocrystal structures of membrane proteins using crystallization chaperones in the PDB reveals that the majority rely on the Fab format (24). This observation could be due to the increased surface area available to Fab fragments to form crystal contacts, or the increased stability the constant domains provide (196), compared to scFvs.

3.1.3 Chapter overview

In this chapter, we report the generation of a Fab fragment with nanomolar affinity for the EE epitope (Fab/EE), along with detailed structural and biochemical characterization relevant to its potential use as a crystallization chaperone. This chapter is also a report of successful complex formation with soluble proteins containing EE peptides, including two EE-tagged maltose binding protein constructs and an scFv with the EE epitope inserted in the flexible linker region. The likelihood of Fab/EE being a successful crystallization chaperone is discussed.

3.1.4 Individual contributions to the work

This work is a collaborative effort between the Lieberman Lab at Georgia Tech and the Maynard Lab at the University of Texas at Austin. Kevin Entzminger and Jeongmin Hyun of the Maynard Lab cloned Fab/EE and performed thermal stability measurements. Other biophysical characterization was conducted by several members of the Lieberman group. I did all crystallography-related work, including crystal growth, data collection, refining and analysis. I largely wrote the paper, with certain sections written by Kevin Entzminger (GPCR sections) and JC Gumbart (MD methods).

3.1.5 Publications resulting from this work

This research was published in Acta Crystallographica Section D, volume 71 in 2015 (120).

3.2 METHODS

3.2.1 Molecular biology, expression and purification of Fab/EE

To convert our previously described 3D5/EE_48 scFv (scFv/EE) (193) to Fab format, the variable light and heavy chains were sequentially sub-cloned via PCR into the *NcoI-NotI* and *NheI-HindIII* restriction sites, respectively, of the pFab vector (courtesy of Dr. Georgiou, University of Texas at Austin) (197) resulting in the vector pFab-Fab/EE. The vector provides in-frame, N-terminal periplasmic leader sequences and C-terminal peptide tags, a decahistidine-tag on the light chain and a FLAG-tag on the heavy chain (**Table 7** and **Table 8**). The fidelity of the construct was confirmed by DNA sequencing (University of Texas Austin core facility) using primers 5'-GTGAGCGGATAACAATTTACACAGG (forward) and 5'-GAGAAGGAGATATACATATGAAGTCG (reverse). Fab/EE was expressed in *E. coli* BL21(DE3). 2 mL LB (Fisher) culture supplemented with 60 µg/mL ampicillin was inoculated with a single colony and incubated for ~4 hours at 37 °C with shaking at 225 RPM. The starter culture was diluted 1:100 in 500 mL Terrific Broth (TB, Fisher) in a 2 L baffled flask and grown overnight with shaking at 225 RPM and 25 °C. Cells were pelleted at 4200 x g for 10 minutes and 4 °C, then the cell pellet was resuspended in 500 mL fresh TB in 2 L flask and incubated for 1 hour at 25 °C, 225 RPM before inducing expression with 1 mM IPTG (Calbiochem) for 4.5-5 hours. Cells were pelleted and flash frozen in liquid nitrogen before placing in -80 °C freezer or cells were subjected to lysis

directly. Fab/EE was purified as reported previously for scFv/EE (198). Briefly, cell pellet was resuspended in 10 mL resuspension buffer (0.1 M Tris pH 8.0, 0.75 M sucrose) per gram of cell pellet. Osmotic shock was carried out by adding 7.5 mL 1 mM EDTA and 2.5 mg lysozyme per gram of cell and rocking or stirring for 45 minutes to 1 hour at 4 °C, then adding 1 mL 0.5 M MgCl₂ per gram of cell and stirring for an additional 45 minutes to 1 hour. After centrifuging for 20 minutes at 47,800 x g, supernatant was subjected to Ni²⁺-affinity chromatography with wash buffer 20 mM Tris pH 8, 500 mM NaCl, 20 mM imidazole and elution buffer containing either 100 mM EDTA or 500 mM imidazole. Fab/EE was further purified by preparative size exclusion chromatography using a Superdex 75 16/600 column equilibrated with 50 mM HEPES pH 7.5, 150 mM NaCl (HBS, **Figure 16**) on an ÄKTA FPLC system (GE Healthcare).

3.2.2 Biophysical characterization of Fab/EE

For all proteins, protein purity and size were assessed by standard reducing and non-reducing 12% SDS-PAGE analysis (148) using Coomassie stain for visualization of all other proteins (**Figure 16 inset**) with protein concentrations determined by micro BCA assay (Pierce) or estimated by absorbance at 280 nm combined with calculated extinction coefficients based on amino acid composition using ProtParam (149). Fractions of monomeric Fab/EE were pooled for subsequent experiments. Fab/EE solubility was determined as described previously (193) by measuring the concentration of soluble protein remaining after concentration to ~20 mg/mL and a 4-day incubation at 4 °C. Fab/EE thermal stability was measured by thermal unfolding after mixing 20 µL of 200 µM Fab/EE or HBS only control with Sypro Orange (1 µl of 1:1000 dilution; Molecular Probes) in a Real Time PCR instrument (ViiaTM7; Applied Biosystems) in increments of

0.96 °C/min from 25 °C to 90 °C and analyzed with ViiaTM7 software (Applied Biosystems). Analysis to determine the T_m , the midpoint of unfolding, was performed with ViiaTM7 software.

3.2.3 Protein crystallization, data collection, structure determination and refinement

Fab/EE (6.5 mg/mL in HBS) was crystallized at room temperature by the sitting drop vapour diffusion method. Conditions were optimized based on Wizard I/II (Emerald Biosystems) solution G4 containing 20% polyethylene glycol (PEG) 8000, 100 mM MES pH 6.0, 200 mM calcium acetate. Crystals used for structure determination were grown from a reservoir solution containing 0.1 M HEPES (pH 7.5), 100 mM calcium acetate, 20-26% (w/v) PEG 8000, and 3% 1-propanol. Crystals were harvested and cryocooled in the reservoir solution supplemented with 15% glycerol. Crystallographic data were collected at beamline 22-ID of the Southeast Regional Collaborative Access Team (SER-CAT) of the Advanced Photon Source (APS), (Darien, Illinois). Data were indexed, integrated, and scaled in HKL-3000 (199). The structure was solved by molecular replacement with Phaser (200) using a polyalanine search model prepared in CHAINSAW (201) that derived from the Fab portion of PDB ID 3SOB (202). The model was iteratively rebuilt in Coot (203) and refined in Phenix (204). Ramachandran outliers were determined using RAMPAGE (205). Crystallographic statistics are presented in **Table 5** and structure was deposited to PDB with code 4X0K.

Table 5 Fab/EE data collection and refinement statistics

Parameter	Fab/EE
Data collection	
Beamline source	APS 22-ID
X-ray wavelength (Å)	1.0
Resolution (Å)	32.97 - 2.04 (2.11 - 2.04)
Space group	P1
Unit-cell parameters	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	53.559 67.131 71.877
α , β , γ (°)	71.3 78.1 85.31
Total No. of reflections	226544
No. of unique reflections	57696 (5221)
Multiplicity	3.9 (3.9)
Completeness (%)	97.71 (93.40)
$\langle I/\sigma(I) \rangle$	14.66 (2.70)
R _{merge} (%)	8.4
Refinement statistics	
Final <i>R</i> _{cryst}	0.1641 (0.2198)
Final <i>R</i> _{free} ^a	0.2078 (0.2680)
No. of non-H atoms	
Protein	7012
Water	525
Total	7537
R.m.s. deviations	
Bonds (Å)	0.004
Angles (°)	0.9
Average B factors (Å ²)	
Protein	38.8
Water	44.3
Ramachandran plot ^b	
Most favoured (%)	97.7
Allowed (%)	2.2
Outliers (%)	0.1

^aR_{free} is calculated for a randomly chosen 5% of reflections which were not used for structure refinement

^bAs calculated by RAMPAGE (205)

3.2.4 Computational analysis of Fab/EE crystal contacts

PDBe Protein Interfaces, Surfaces and Assemblies (PISA) (206) was used to rank and analyze crystal lattice contacts by surface area and energy, as well as catalog critical amino acids in crystal contacts based on their formation of hydrogen bond or salt bridge interactions. After excluding the native heavy-light chain interface within the Fab/EE

monomer, the top three interfaces were identified as major crystal contacts and used for further analysis.

3.2.5 Molecular biology, expression and purification of soluble test proteins presenting the EE epitope

All soluble test proteins used were described previously (193). The EE-tag was appended to the C-terminus of maltose binding protein (MBP-EE) or into a surface loop of MBP replacing residues 170-175 (MBP-KEE). MBP with only a C-terminal hexahistidine-tag was used as a negative control (MBP-His₆). An scFv with the EE-tag inserted into the flexible linker region (scFv-EE₁) was used as another EE-tagged test protein for BIAcore. These proteins were expressed and purified via C-terminal hexahistidine-tag as described for Fab/EE.

3.2.6 Binding assays and complex formation

3.2.6.1 Surface plasmon resonance (SPR)

Kinetic binding assays were performed with a BIAcore 3000 (GE Healthcare) instrument using immobilized bovine serum albumin (BSA) or Fab/EE coupled to CM5 chips via NHS-EDC chemistry to a level of ~1200 response units (RU) as bait for ligand proteins. Responses due to sample refractive index changes and non-specific binding were corrected using signal from a flow cell coupled with BSA. Purified MBP-KEE, MBP-EE, scFv-EE₁ or control MBP-His₆ were injected in a duplicate dilution series from 2 to 0.125 μ M at a flow rate of 50 μ l/min to minimize mass transport effects in a HBS running buffer supplemented with 0.005% Tween-20. Surface regeneration was performed after each run with a single 30 second injection of 2 M MgCl₂.

3.2.7 Size exclusion chromatography

Fab/EE–client protein interactions were further evaluated by SEC fractionation on an ÄKTA FPLC system (GE Healthcare) at 4 °C. Equimolar amounts of Fab/EE and MBP-KEE were incubated 90 min, either together or separately, at room temperature prior to fractionation on a Superdex S200 column (GE Healthcare) equilibrated with HBS. Elution fractions for each peak were precipitated by trichloroacetic acid (TCA) (207) and analyzed with 12% reducing SDS-PAGE.

3.3 RESULTS

3.3.1 Fab/EE molecular biology, expression and purification

Fab/EE was generated by sub-cloning the variable regions into the di-cistronic plasmid pFab for bacterial expression (197). Here, each variable region is appended with a cognate human constant domain such that both polypeptide chains (V_L - C_L and V_H - C_H1) are targeted to the periplasm for correct assembly and disulfide bond formation. After osmotic shock and purification (**Figure 16**), similar levels of total and monomeric Fab/EE protein were recovered as for scFv/EE (2.4 mg/L culture versus 2.1 mg/L culture; 87% versus 81% monomeric, respectively; **Table 6**). Fab/EE thermal stability was significantly enhanced as compared to scFv/EE (**Table 6**).

3.3.2 Fab/EE structural characterization

Crystals of Fab/EE belong to space group P1 and its 2.0 Å resolution structure was solved by molecular replacement. Most residues were successfully modelled into the electron density map. Exceptions include S128 to S134 on chain H, FLAG tag on the C-terminal end of the heavy chains (chains H and A) linker residues between the C-terminal end of the light chains (chains L and B) and the last 4 of the decahistidine-tag residues

(**Table 7** and **Table 8**). The only Ramachandran outlier in Fab/EE structure is H222, part of the decahistidine-tag of chain B, and fits well into the electron density. Two nearly identical molecules are present in the asymmetric unit. Chains L and B (light chains) superimpose with a root mean squared deviation (RMSD) of 0.599 Å and chains A and H (heavy chains) with an RMSD of 0.976 Å, using secondary-structure matching (SSM) function within Coot (208) (**Figure 17a**). Fab/EE chains A and B (monomer AB) and scFv/EE chains A and B (PDB ID 3NN8) are also nearly identical, with an RMSD of 0.708 Å between two light chains and 0.581 Å between two heavy chains.

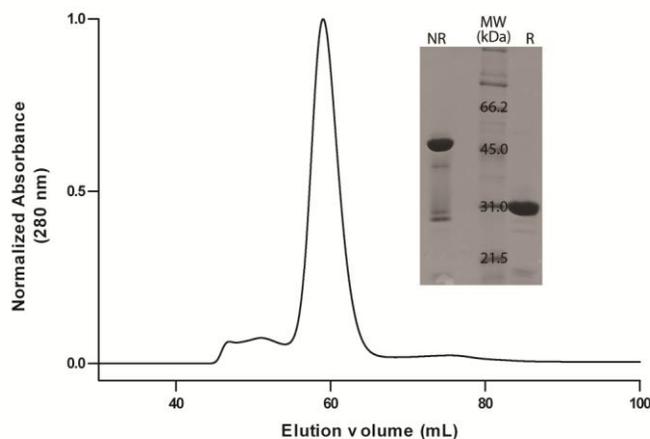


Figure 16 Fab/EE purification

Purified from the *E. coli* periplasm by Ni²⁺-affinity followed by SEC, Fab/EE elutes from an analytical gel filtration column as a single peak at the expected elution volume. Inset, non-reducing (NR) and reducing (R) SDS-PAGE gel of a fraction taken from the major peak predominantly shows a single dominant band with >95% purity ~50 kDa and 31 kDa for non-reducing and reducing samples, respectively.

Table 6 Biophysical characteristics of EE peptide-binding antibody fragments

Parameter	scFv/EE ^a	Fab/EE
Expression level (mg/L culture)	2.1	2.4
Solubility (mg/mL)	12.8	9.4
Melting temperature (°C)	47.2 ± 0.3	59.8 ± 0.1
% monomeric protein	81	87

^aAs reported in (193)

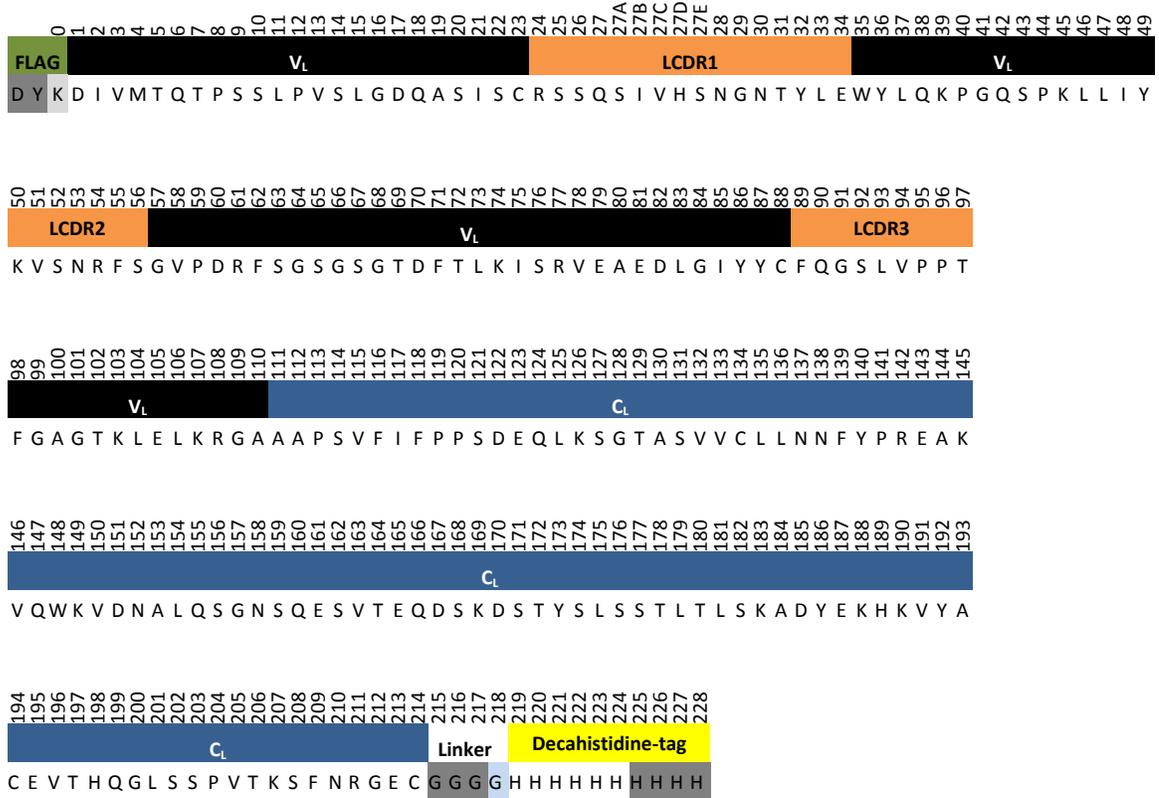
Table 7 Fab/EE heavy chains (chain H and A) Kabat numbering and modelled residues

Residues in white were modelled in both chains. Residues in light grey were not modelled in chain H. Residues in dark grey were not modelled in chains H or A. Variable and constant domains are labelled, as are complementarity determining regions.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	35A	35B	36	37	38	39	40	41	42	43	44	45	46	47	48	49
V _H											HCDR1								V _H																															
Q V Q L Q Q S G P E D V K P G A S V K I S C K A S G Y S L S T S G M G V N W V K Q S P G K G L E W L A																																																		
50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	82A	82B	82C	83	84	85	86	87	88	89	90	91	92	93	94			
HCDR2																V _H																																		
H I Y W D D D K R Y N P S L K S R A T L T V D K S S S T V Y L E L R S L T S E D S S V Y Y C A R																																																		
95	96	97	98	99	100	100A	100B	100C	100D	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	133	134	135	136	137	138	139	140			
HCDR3										V _H										C _H 1																														
R G G S S H Y Y A M D Y W G Q G T T V T V S S A S F K G P S V F P L A P S S K S T S G G T A A L																																																		
141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	162	163	164	165	166	167	168	169	171	172	173	174	175	176	177	178	179	180	182	183	184	185	186	187	188	189	190	191	192	193	194	195		
C _H 1																																																		
G C L V K D Y F P E P V T V S W N S G A L T S G V H T F P A V L Q S S G L Y S L S S V V T V P S																																																		
196	197	198	199	200	203	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	225	226	227	228	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250				
C _H 1																Hinge										Linker				FLAG-tag																				
S S L G T Q T Y I C N V N H K P S N T K V D K K V E P K S C D K T H T G G G G D Y K D D D D K																																																		

Table 8 Fab/EE light chains (chain L and B) Kabat numbering and modelled residues

Residues in white were modelled in both chains. Residues in light grey were not modelled in chain B. The residue in pale blue was not modelled in chain L. Residues in dark grey were not modelled in chains L or B. Variable and constant domains are labelled, as are complementarity determining regions.



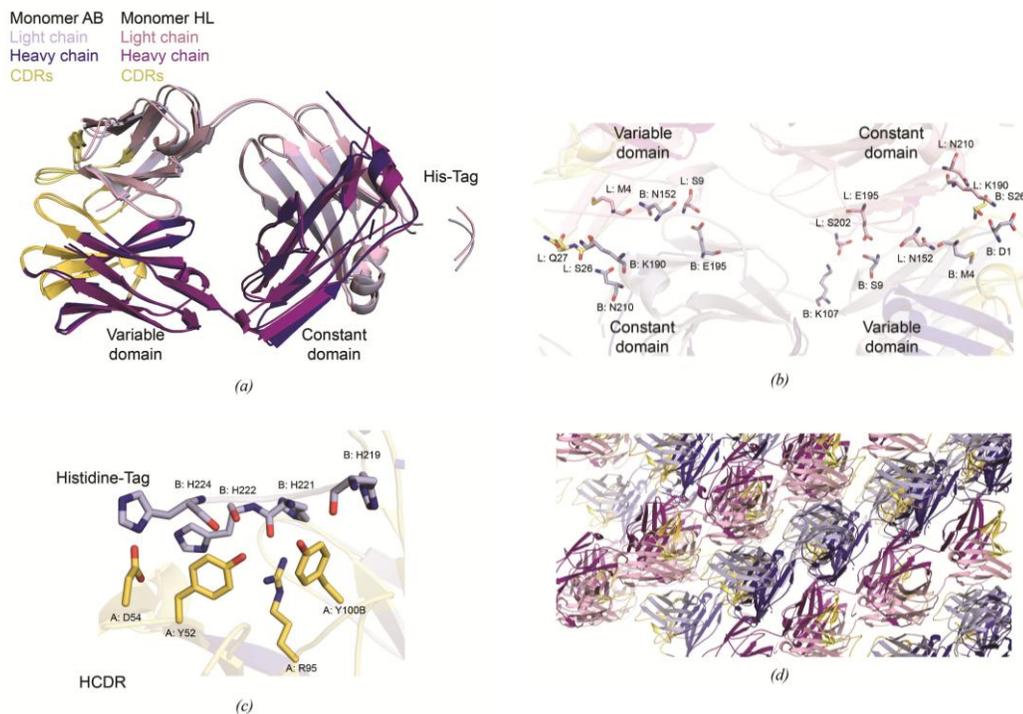


Figure 17 Structure of Fab/EE

(a) Overlay of two Fab/EE molecules in the asymmetric unit. RMSD of the two heavy chains (chain A and H) is 0.976 Å and RMSD of the two light chains (chain B and L) is 0.599 Å. (b) Fab/EE crystal contact ID2. The interface is 1109 Å² and includes 10 residues from chain L and 9 residues from chain B (both light chains, interacting residues are modelled as sticks). (c) Crystal contact ID5. Interaction between modelled histidine-tag from chain L and CDR of chain H of a different molecule. Interacting residues are modelled as sticks. (d) Fab/EE showing extended crystal contact areas and lack of channel that could accommodate a membrane protein. Colors as in a.

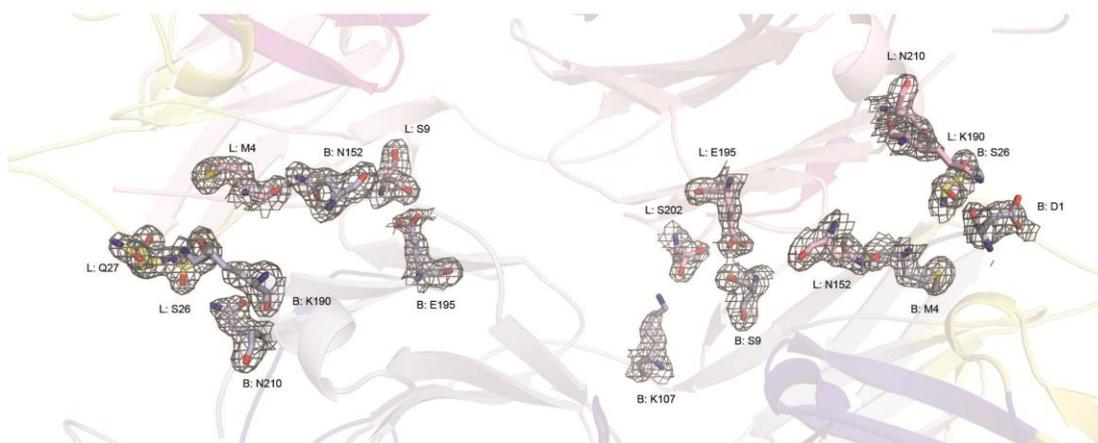
The lattice of Fab/EE demonstrates the variety of crystal contacts available to aid in the crystallization of client proteins (**Figure 17b-d**). Overall, the crystal contacts of Fab/EE had larger interface surface areas than those of scFv/EE, with the largest interface of Fab/EE calculated to be 1109 Å² (**Table 9** ID2, **Figure 17b**). The only residue that participates in crystal contacts in both Fab/EE and scFv/EE crystals is K107 (**Figure 17b**, **Table 9**), and the two interfaces are not similar. The Fab/EE interface contains the variable domain of chain B in one Fab/EE molecule interacting with the constant domain of chain L in the second Fab/EE molecule and vice versa (**Figure 17b** and **Figure 18**).

The corresponding scFv/EE interface is between the variable domains of the light chain and the heavy chain. The next largest Fab/EE contact by surface area, ID3, has interactions between heavy chains of adjacent molecules (not shown), with less than half the interface area of ID2 (**Table 9**). Unexpectedly, a portion of the decahistidine-tag on the C-terminal of the light chain of Fab/EE forms hydrogen bonding and salt bridge interactions (ID5) with the CDRs from the heavy chain. This interaction is seen in both molecules in the asymmetric unit (**Figure 17c** and **Figure 19a-b**). Finally, the P1 lattice lacks solvent channels to accommodate a client EE-tagged protein (**Figure 17d**) and the CDRs are being used in the crystal contacts.

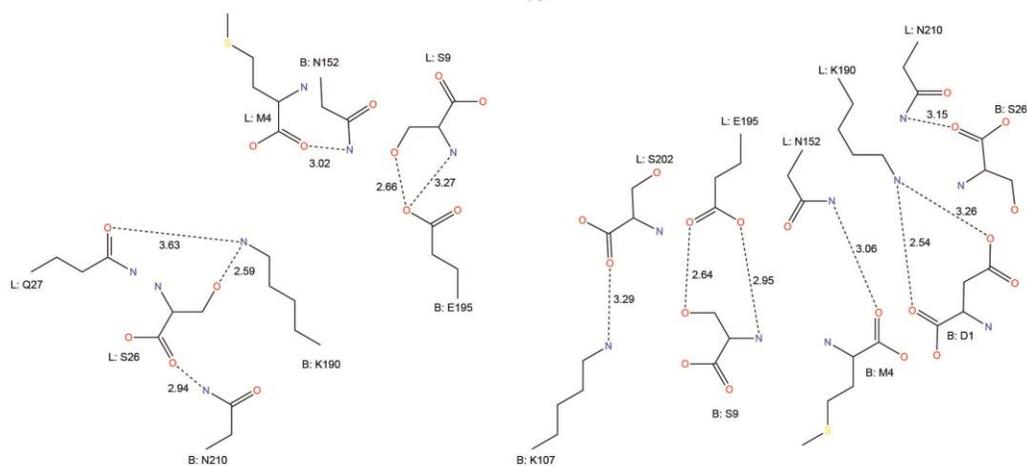
Table 9 PISA contact comparison between scFv/EE and Fab/EE

Antibody fragment	PISA interface	Interface area (Å ²)	V _H contacts that formed H-bonds/salt bridges	V _L contacts that formed H-bonds/salt bridges	C _{H1} contacts that formed H-bonds/salt bridges	C _L contacts that formed H-bonds/salt bridges
scFv/EE ^a	ID2	465	K57, R58, S65, T68	-	-	-
	ID3	456	P9, K19, S30, S32, D72, K73, S75, T77, Y79, S98	-	-	-
	ID4	393	Q6, D11, Y91, Q105, T108	L106, K107, R108, G110 (linker)	-	-
Fab/EE	ID2	1109	-	D1, M4, S9, S26, Q27, K107	-	N152, K190, E195, S202, N210
	ID3	517	Q1, S28, V71,	-	E150, K213	-
	ID4	345			S168, N216	
	ID5	325	Y52, W53, D54, D56, R95, Y100B	-	-	H219 (His-tag), H221 (His-tag), H222 (His-tag), H224 (His-tag)

^aAs reported in (119)



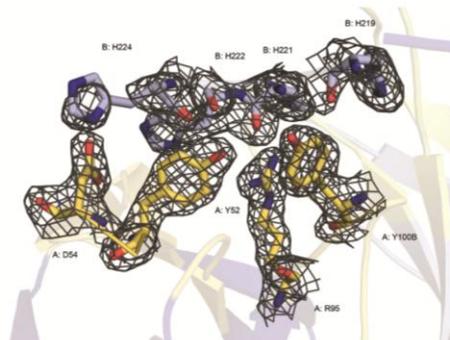
(a)



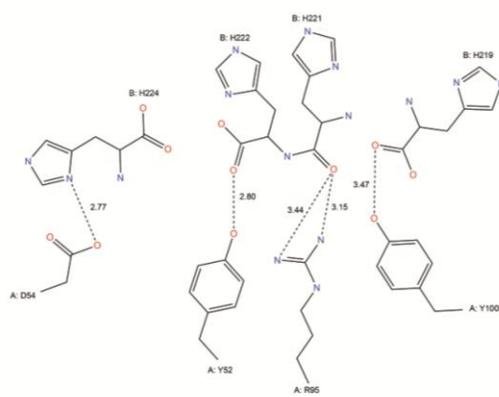
(b)

Figure 18 Interactions and final $2F_o - F_c$ electron density for crystal contact ID2

(a) Electron density for residues involved in crystal contact, contoured to 1σ . (b) Interactions between light chain (chains L and B) residues in contact ID2 depicted as dashed lines, with distances (\AA) according to PISA analysis (Table 3).



(a)



(b)

Figure 19 Interactions and final $2F_o - F_c$ electron density for crystal contact ID5

(a) Electron density for residues involved in crystal contact from chains A and B, contoured to 1σ . (b) Interactions between the histidine-tag on the C-terminus of the light chain (chain B) and the CDR of the heavy chain of a symmetry related molecule (chain A). Interactions are depicted as dashed lines, with distances (Å) according to PISA analysis.

3.3.3 Fab/EE forms stable complexes with soluble EE-tagged client proteins

To assess EE binding properties of Fab/EE, we first examined interactions with soluble proteins presenting the EE peptide by SPR. Binding affinity in the nanomolar range was measured for Fab/EE binding to MBP containing a C-terminal (MBP-EE), or an internal (MBP-KEE) EE peptide, as well as to an scFv with the EE peptide inserted into the flexible linker region (scFv-EE₁) (Table 10, Figure 20a-c). No significant affinity loss was observed upon conversion of the scFv format to Fab (767 nM for scFv/EE binding to MBP-EE (192) and 308 nM for Fab/EE binding to MBP-EE), as has

been observed previously (209-211). No binding was observed to the negative control, MBP-His₆ (Figure 20d).

Table 10 Characterization of EE-tagged protein binding kinetics of Fab/EE by SPR

	k_{on} ($M^{-1} \cdot s^{-1}$)	k_{off} (s^{-1})	K_D (nM)
MBP-EE	$3.39 \pm 1.23 \times 10^4$	$9.22 \pm 0.15 \times 10^{-3}$	308 ± 117
MBP-KEE	$1.25 \pm 0.30 \times 10^4$	$7.88 \pm 0.91 \times 10^{-3}$	612 ± 95
scFv-EE ₁	$2.95 \pm 1.47 \times 10^5$	$5.56 \pm 1.66 \times 10^{-2}$	224 ± 160

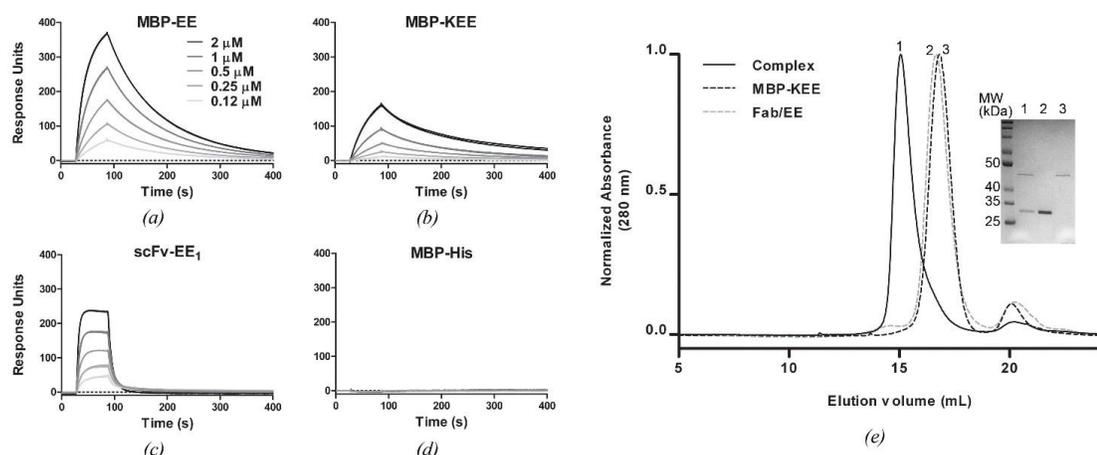


Figure 20 Complexation data for Fab/EE binding to soluble EE peptide-containing proteins

EE-peptide-containing soluble proteins were injected in duplicate for each concentration tested, and binding to immobilized Fab/EE was monitored by SPR. Both duplicate traces are shown and demonstrate Fab/EE binding to (a) MBP-EE (b) MBP-KEE, and (c) scFv-EE₁. (d) No binding was observed for MBP-His₆. (e) Equimolar amounts of MBP-KEE and Fab/EE were equilibrated at room temperature separately or together prior to separation by analytical gel filtration. The sample containing both proteins eluted as a single peak with shorter retention than either individual protein. (Inset) fractions from each major peak were analyzed by reducing SDS-PAGE. (inset lane 1) peak fraction from the putative complex peak. (inset lane 2) Peak fraction from the Fab/EE peak. (inset lane 3) peak fraction from MBP-KEE peak.

Since SPR employs surface immobilization strategies that may not accurately reflect binding in solution or in a crystallization drop, we assessed complex stability by SEC. Notably, SEC is often used to isolate protein-protein complexes prior to co-crystallization trials (118, 212). MBP-KEE was selected among the aforementioned client proteins for solution complexation with Fab/EE. A clear shift in elution volume corresponding to higher molecular mass was observed for the complex, and complexation was confirmed by SDS-PAGE (**Figure 20d**).

3.4 DISCUSSION

3.4.1 Fab/EE rationale, cloning, and biophysical properties

Previously, we reported the successful engineering of both EE- and hexahistidine peptide-specific scFvs for potential use as crystallization chaperones (119, 192). Ultimately, the anti-EE scFv proved higher affinity and preferable over its anti-hexahistidine counterpart likely due to its chemical diversity, insensitivity to pH especially near physiological conditions, and greater compatibility with a variety of peptide insert locations (terminal and internal). Here we sought to convert our EE-specific scFv to Fab format, which represents the majority of crystallization chaperone proteins in solved structures deposited to the PDB, and because Fab molecules typically exhibit better biophysical properties such as enhanced thermal stability and well defined oligomeric states compared to the scFv platform (196). We also postulated that the larger size of a Fab (~50 kDa versus ~25 kDa) would provide additional epitopes to mediate crystal contacts, especially advantageous for larger client membrane proteins. The combination of high expression yields and enhanced stability, a biophysical characteristic

correlated with crystallization success rates (213), render Fab/EE more promising than the parent scFv/EE for use in large-scale co-crystallization trials.

3.4.2 Fab/EE structural characterization

As noted in Section 3.3.2, Fab/EE was crystallized in space group P1 with CDR residues involved in crystal contacts and small solvent channels that could not accommodate a client protein, unlike the crystal lattice of scFv 3D5. Nearly all residues in scFv/EE are also in Fab/EE, with the exception of the heavy-light chain linker region in scFv/EE (194). All residues that form crystal contacts in scFv/EE are also available in Fab/EE, meaning that Fab/EE has a variety of residues and interfaces that are available for forming crystal contacts. We do not consider the Fab/EE lattice a negative aspect of Fab/EE for use as a crystallization chaperone because upon binding to target protein, Fab/EE would likely utilize other residues that are available for forming crystal contacts.

Interestingly, the hexahistidine tag forms a crystal contact with CDR regions of the heavy chain, the area of antigen binding. Negative control proteins for complexation studies by SPR, SEC, and ELISA data (not shown) has confirmed that Fab/EE does not have any detectable affinity for the hexahistidine tag, so the His-tag/HCDR interface is an artifact of crystallization.

3.4.3 Fab/EE forms stable complexes with soluble EE-tagged client proteins

As shown in Section 3.3.3, Fab/EE binds several EE-tagged soluble proteins using SPR and SEC. MBP-KEE has an internal EE-tag, MBP-EE has an EE-tag at the C-terminus, and scFv-EE₁ has the EE-tag inserted in the linker region of the scFv. The binding affinity of Fab/EE for the EE-tag and the diversity of EE-tag placement within the client protein will likely prove to be very beneficial for the use of Fab/EE as a

crystallization chaperone. Since Fab/EE does crystallize alone, it will likely be necessary to isolate the Fab/EE:EE-tagged client protein complex over SEC prior to setting up crystallization trials, though that approach does not always produce complex crystals either (118). The length of the loop containing the EE epitope may also be very important, as longer loops tend to be flexible. Disordered or flexible regions are difficult to crystallize.

In conclusion, we have developed Fab/EE for use as a crystallization chaperone for the crystallization of EE-tagged proteins. In this chapter, we characterized the Fab/EE antibody fragment using biophysical means including crystallization. For complexation studies, we used SPR and SEC to show that Fab/EE binds tightly to soluble EE-tagged proteins. In further chapters, we expand this idea to EE-tagged membrane proteins and discuss crystallization trials of Fab/EE:EE-tagged protein complexes.

CHAPTER 4: PROGRESS TOWARD CO-CRYSTALLIZATION WITH EPITOPE-SPECIFIC FAB ANTIBODY FRAGMENT AND TAGGED MEMBRANE PROTEINS

4.1 INTRODUCTION

Chapter 1 set the stage for the crystallization of membrane proteins by reviewing methods that are used to increase the likelihood that a membrane protein will crystallize, including covalent and non-covalent crystallization chaperones. In Chapter 2, our methods for the expression and purification of archaeal orthologs of SPP were discussed, including the determination of protein stability in different detergents, which is an important factor for crystallization. Chapter 3 showed the characterization of an epitope-specific crystallization chaperone, Fab/EE, that could be used to aid in the crystallization of EE-tagged membrane proteins. Examples of complexation of Fab/EE with soluble EE-tagged proteins were shown using SPR and SEC.

Here in Chapter 4, progress toward using Fab/EE as a crystallization chaperone for EE-tagged membrane proteins is shown, taking all that we learned in Chapters 2 and 3 and putting them together. The pathway taken towards crystallization of a Fab/EE:EE-tagged membrane protein will be discussed. Complexation studies were done with two test proteins, human adenosine A_{2a} G protein-coupled receptor (A_{2a}R-GFP-EE) and β -barrel membrane protein *E. coli* intimin (intimin-EE7 and intimin-EE8). The high resolution structures of A_{2a}R and intimin are published(214), and intimin is easily expressed and purified, making it ideal for structure studies.

We are interested in SPP (see Chapter 2) from a crystallization standpoint. The structure was published in 2013 (4), but the project (the co-crystallization of EE-tagged

SPP with Fab/EE) predates the structure by several years, and the structure is relatively low resolution, in an inactive confirmation, and missing loops of interest. Solution complexation studies were done with Fab/EE and EE-tagged *H. Mar* SPP (*HmSPP-EE*) and EE-tagged *M. mar* SPP (*MmSPP-EE2*).

After numerous crystallization trials, molecular dynamic simulations with wild type (WT) intimin, intimin-EE7 and intimin-EE8, and the Fab/EE:intimin-EE complexes reveal unexpected increase in flexibility when mutating native loop residues to the EE epitope, which is likely hindering crystallization of the complex. Implications of these findings for co-crystallization and future directions are discussed.

4.1.1 Individual contributions to the work

This work is highly collaborative between everyone in the SPP subgroup in the Lieberman Lab. Jeff Culver designed and made the *HmSPP-EE* construct. Sibel Kalyoncu and I worked together to generate most of the *MmSPP* constructs. Sibel made the initial *MmSPP-EE* constructs and all of the variants with the shortened TM6-TM7 loop. Swe-Htet Naing made the D to A inactive mutants of *MmSPP*. I made the remainder of the constructs. All GPCR expression, purification, and SPR was done by Kevin Entzminger (Maynard Lab, University of Texas, Austin). The intimin-EE constructs were designed by undergrads Ivan Morales and David Heaner, under the supervision of Sibel and myself. All complexation SEC data was collected by me, and the molecular dynamic simulations were done by J.C. Gumbart (Georgia Tech).

4.1.2 Publications resulting from this work

The intimin constructs were used for scFv complexation (119) and Fab/EE complexation (120). Molecular dynamic simulations were presented in the paper

detailing Fab/EE characterization (120), and some of the *MmSPP* constructs were used for an SPP enzyme assay performed by Swe (manuscript submitted).

4.2 METHODS

4.2.1 A_{2a}R expression and purification

The plasmid pITy-A_{2a}R-GFP-His₁₀ was generously provided by Dr. Anne Robinson (University of Tulane). An EE-tagged variant, pITy-A_{2a}R-GFP-EE, was generated by insertion of the EE epitope after wild-type residue Lys 209 (numbering as in GenBank AAA83270) in the ICL3, flanked by GS residues to allow for peptide accessibility (**Figure 21**). This was accomplished by SDM using QuikChange mutagenesis kit (Stratagene). Colony PCR was used to screen for modified plasmids as described previously (215). Correct EE peptide insertion of the final plasmid was confirmed by DNA sequencing. All primers used are listed in **Table 11**.

Table 11 Primers for A_{2a}R-GFP-EE SDM, colony screening, and sequencing

Primer purpose	Forward primer	Reverse primer
SDM for A _{2a} R-GFP-EE	5'-CGACGACAGCTGAAGGGTAGT GAATATATGCCAATGGAAGGTAG TCAGATGGAGAGCCAG	5'-CAGCTGTCGTCGCGCCGCCAG GAAGATCCG
pITy-A _{2a} R-GFP-EE colony screening	5'-GGTTTTGATTGTCTTGTGGC	5'-CTACCTTCCATTGGCATATAT TC
pITy-A _{2a} R-GFP-EE sequencing	5'-CACGACTTCTTCAAGTCCG	5'-GCCATCCGCGGCTTGTACAGC TGTCAT

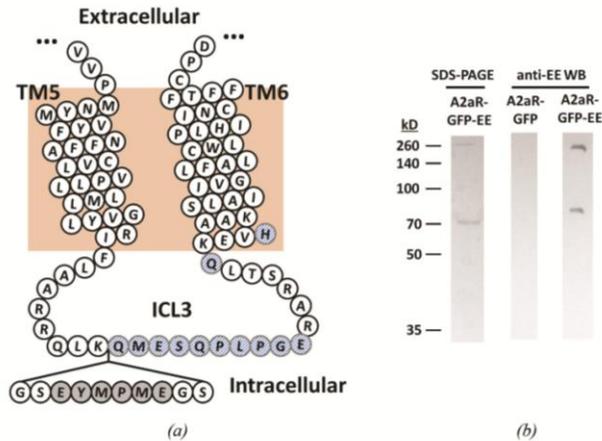


Figure 21 Design and purification of A_{2a}R-GFP-EE

(a) Collier de perles diagram of the A_{2a}R transmembrane regions 5-6 (TM5-6) depicting the insertion location for the EE peptide (shaded) with flexible linker residues. Residues homologous to those forming the Fab epitope in the Fab:GPCR cocrystal structure of β_2 adrenergic receptor (PDB ID 2R4S) are shown using hatched circles. (b) SDS-PAGE gel with silver staining shows single band at the expected monomer size for A_{2a}R-GFP-EE after purification by Ni²⁺-affinity and SEC. Western blot of the same fractions using commercial anti-EE antibody confirms the presence of the EE peptide. The lower band represents monomer and the upper band dimer at the expected sizes.

Plasmids pITy-A_{2a}R-GFP and pITy-A_{2a}R-GFP-EE were transformed into *S. cerevisiae* BJ5464 by electroporation, with individual yeast colonies screened for high expression by whole-cell GFP fluorescence and the highest-expressing clones were used for subsequent protein purification as described previously (216). Briefly, cell lysis was accomplished by vortexing and protein was purified by Ni²⁺-affinity chromatography using wash buffer 50 mM NaH₂PO₄, 300 mM NaCl, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.1% DDM, 0.1% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 0.02% cholesteryl hemisuccinate (CHS), pH 8.0 with protease inhibitors and the same buffer for elution with the addition of 500 mM imidazole. Ni²⁺-affinity chromatography was followed by SEC using a Superdex 200 16/60 (GE Healthcare) column equilibrated with GPCR buffer

(10mM NaH₂PO₄, pH 7.0 with 0.1% DDM 0.1%, CHAPS, 0.02% CHS). Purified protein was analyzed by SDS-PAGE and Western blot using the primary anti-EE peptide Glu-Glu monoclonal antibody (1:1000 dilution in blocking buffer; Covance) and incubation with horseradish peroxidase (HRP)-conjugated anti-mouse secondary antibody (ThermoFisher). Signal was developed with SuperSignal West Dura Extended Duration Substrate (Thermo Scientific), and the resulting image captured on X-ray film.

4.2.2 SPR with Fab/EE and A₂aR-GFP-EE

GPCR variants were injected similarly to MBP-EE in Section 3.2.6.1 with the exceptions of using a single dilution series from the highest concentration available upon purification and running buffer composed of HBS with 0.1% DDM. The association rate constant (k_{on}), dissociation rate constant (k_{off}), and equilibrium dissociation constant (K_d ; $K_d = k_{off}/k_{on}$) were calculated assuming a Langmuir 1:1 binding model with BIAevaluation software. Only data sets with $\chi^2 < 0.5$ were used. GraphPad Prism 5 was used for graphical representation.

4.2.3 ELISA of Fab/EE and A₂aR-GFP-EE

Fab/EE binding to A₂aR-GFP-EE was also assessed by ELISA. High-binding 96-well plates (Costar) were coated with 20 µg/mL Fab/EE overnight at 4 °C. After a 1 hour incubation with blocking buffer at room temperature, purified A₂aR-GFP-EE or A₂aR-GFP proteins were serially diluted 2-fold in blocking buffer (PBS, 5% milk). This was followed by 1 hour incubation with rabbit anti-GFP (Invitrogen), followed by washing, and 1.5 hour incubation with goat anti-rabbit HRP (Sigma). After washing, 3,3',5,5'-Tetramethylbenzidine (TMB) substrate (Vector Labs) was added, signal allowed to develop and the reaction quenched with 1 M H₂SO₄. Absorbance was read at 450 nm on a

SpectraMax M5 microplate reader (Molecular Devices). Data points represent the average of at least two measurements, including error bars equal to one standard deviation. GraphPad Prism 5 was used to fit data to a three-parameter logistic model and for graphical representation.

4.2.4 SPP molecular biology, expression and purification

4.2.4.1 *HmSPP* molecular biology, expression, and purification

Online secondary structure prediction servers (for example, Jpred 3 (217) and PSIPRED (218)) were used to predict loops available for the insertion of the EE epitope. EE-tagged *H. mar* SPP (*HmSPP*) were prepared using SDM, and plasmid fidelity was confirmed by sequencing (MWG Operon).

Protein expression and membrane isolation was carried out as stated in Section 2.2.5. Isolated membrane was solubilized by first thawing frozen membrane on ice, then adding 0.3-1 g membrane to Dounce homogenizer with enough 50 mM HEPES pH 7.5, 500 mM NaCl, and 20 mM imidazole to make a 100 mg/mL membrane solution. Membrane was resuspended using the loose plunger then the tight plunger until homogeneous. In a separate container, FC12 equal to the mass of membrane was added to 50 mM HEPES pH 7.5, 500 mM NaCl, and 20 mM imidazole to make a 100 mg/mL FC12 solution. The membrane solution was added to the detergent solution and rocked at 4 °C for at least 1 hour. Sample was then centrifuged at 181,000 x g for 45 minutes at 4 °C to remove any unsolubilized material. The supernatant was added to a superloop of appropriate size (GE Lifescience) and purified by Ni²⁺-affinity chromatography using 50 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole, and 0.1% FC12 as the wash

buffer and 50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole, and 0.1% FC12 as the elution buffer.

4.2.4.2 *MmSPP* molecular biology to insert the EE epitope and build construct library

Since the structure was not available at the time this project was planned, secondary structure prediction software was used to predict the start and end of the long loop between TM6 and TM7. The EE tag was mutated in three places in the loop between TM6 and TM7 (**Figure 22**).

After initial complexation trials and the publication of the *MmSPP* structure (4), new constructs were prepared to hopefully increase the likelihood of crystallization. The first 23 amino acids, originally thought to be a signal sequence, were added back to the truncated construct using restriction free (RF) cloning (219). 5 mutations that were outlined in Li *et al* were made to increase expression and monodispersity (D40N, E42S, A147E, V148P, A229V) (4). The long loop containing the EE epitope was also truncated by 4, 8, and 12 amino acids to limit its flexibility. As mentioned in Chapter 2, a TEV cleavage site was also added to the end of several constructs so that the flexible hexahistidine tag could be removed which increases purity by removing proteins that bind the Ni²⁺-affinity resin and decreases flexibility by removing the unstructured hexahistidine tag. Most of these mutations were made to the WT and the *MmSPP*-EE2 constructs. All *MmSPP* constructs are outlined in **Table 12** and primers used are in **Table 13**.

	TM6		TM7
<i>Mm</i> SPP-EE1	VLPVLLVLLAVY	D A I S V Y R T K H M I T L A E G V L E	E Y M P M E V V V P K R A D Y S F R K E G L N I S E G E E R G A F V M G M G D L I M P S I L V A S S H
<i>Mm</i> SPP-EE2	VLPVLLVLLAVY	D A I S V Y R T K H M I T L A E G V L E T K A P I M V V V P K R A	E Y M P M E E G L N I S E G E E R G A F V M G M G D L I M P S I L V A S S H
<i>Mm</i> SPP-EE3	VLPVLLVLLAVY	D A I S V Y R T K H M I T L A E G V L E T K A P I M V V V P K R A D Y S F R K E	E Y M P M E G E E R G A F V M G M G D L I M P S I L V A S S H

Figure 22 Alignment of *Mm*SPP-EE variants showing EE tag in three locations

The alignment shows all residues from the beginning of TM6 (V148) to the end of TM7 (V233) of *Mm*SPP-EE variants. Helices are marked by black bars at the top of the alignment and are taken from the published structure (4). The catalytic aspartate residues (D162 and D220) are highlighted yellow, and the EE epitopes are colored brown, blue, and green.

Table 12 *MmSPP* constructs

All constructs are in pET-22b(+) and were constructed by SDM of WT plasmid, with the exception of the full length (FL) which was constructed using RF cloning. The WT plasmid contains no EE tag, and EE1, EE2, and EE3 are shown in **Figure 22**. FL constructs contain the first 23 amino acids (teal), and Δ N23 constructs are missing the first 23 amino acids. The catalytic aspartate residues were mutated to alanines to make inactive mutants (orange). The long loop between TM6 and TM7 was truncated by 4, 8, and 12 residues to limit the loop flexibility (pink). The 5 missense mutations used for the crystal structure were added (dark green). Finally, a TEV cleavage site was inserted between the protein and the hexahistidine tag (purple).

WT/EE	FL	D162/D220	Loop length	Missense Mut.	TEV/NO TEV	Construct name
WT	Δ N23	DD	WT	no	no	WT <i>MmSPP</i>
WT	Δ N23	DD	WT	no	yes	
WT	Yes	DD	WT	no	no	
WT	Yes	DD	WT	no	yes	
WT	Δ N23	AA	WT	no	no	
WT	Δ N23	AD	WT	no	no	
WT	Δ N23	DA	WT	no	no	
WT	Δ N23	DD	WT	yes	no	
WT	Δ N23	DD	WT	yes	yes	<i>MmSPP</i> -EE2
WT	yes	DD	WT	yes	no	
WT	yes	DD	WT	yes	yes	
EE2	Δ N23	DD	WT	no	no	
EE2	Δ N23	DD	WT	no	yes	<i>MmSPP</i> -EE2S2
EE2	yes	DD	WT	no	no	
EE2	yes	DD	WT	no	yes	
EE2	Δ N23	AA	WT	no	no	
EE2	Δ N23	DA	WT	no	no	
EE2	Δ N23	AD	WT	no	no	
EE2	Δ N23	DD	-4	no	no	
EE2	Δ N23	DD	-4	no	yes	
EE2	Δ N23	DD	-8	no	no	
EE2	Δ N23	DD	-8	no	yes	
EE2	Δ N23	DD	-12	no	no	<i>MmSPP</i> -EE3S2
EE2	Δ N23	DD	-12	no	yes	
EE2	yes	DD	-8	no	yes	<i>MmSPP</i> -EE3S2
EE2	yes	DD	-12	no	yes	
EE2	Δ N23	DD	WT	yes	no	EE1, EE2, EE3
EE2	Δ N23	DD	WT	yes	yes	
EE2	yes	DD	WT	yes	no	
EE2	yes	DD	WT	yes	yes	
EE1	Δ N23	DD	WT	no	no	
EE3	Δ N23	DD	WT	no	no	

Table 13 Primers used to make *MmSPP* constructs outlined in Table 12

Primers for RF cloning were designed using RFloning.com (219), and primers for SDM were designed using Agilent QuikChange Primer Design Program (<https://www.genomics.agilent.com/primerDesignProgram.jsp>)

Primer Purpose	Forward Primer	Reverse Primer
<i>MmSPP</i> -EE1 SDM	GCTGGCCGAAGGCGTCCTCGAGGAGTAT ATGCCCATGGAGGTCGTGGTTCCGAAGA GAGC	GCTCTCTTCGGAACCACGACCTCCATG GGCATATACTCCTCGAGGACGCCTTCG GCCAGC
<i>MmSPP</i> -EE2 SDM	GTCGTGGTTCCGAAGAGAGCGGAGTAC ATGCCCATGGAAGAGGGGCTCAACATC AGTGAG	CTCACTGATGTTGAGCCCCCTCTCCAT GGGCATGTACTCCGCTCTCTTCGGAAC CACGAC
<i>MmSPP</i> -EE3 SDM	GACTACTCGTTCAGGAAAGAGGAGTAC ATGCCCATGGAGGGGAGGAGCGCGGC GCGTTC	GAACGCGCCGCGCTCCTCCCCCTCCAT GGGCATGTACTCCTCTTCTGAAACGA GTAGTC
PCR of whole <i>MmSPP</i> gene with TEV cleavage site for RF cloning	CGCTGCCCAGCCGGCGATGGCCATGCAG ATACGCGACTGG	GGTGGTGGTGGTGCTCGAGTGCAGAC TGGAAGTACAGATTCTCGTCGACGAA AGGAAGCCACGAAAACGAAC
PCR of first 23 amino acids for RF cloning	CGCTGCCCAGCCGGCGATGGCCATGCAG ATACGCGACTGG	GCATCGGCATGACGAGGACGATAGCG ATGATCTGGACGAAC
D162A SDM	CTCGCGGTCTACGCCGCCATATCGCTC	GACCGATATGGCGGCGTAGACCGCGA G
D220A SDM	GTCATGGGTATGGGCGCTCTCATCATGC C	GGCATGATGAGAGCGCCCATACCCAT GAC
SDM to make - 4 amino acids in TM6-TM7 loop	GCCCATGGAAGAGGGGCTCAACATCAG TCGCGGCGCGTTCGTTCATGGGTATGGGC G	CGCCCATACCCATGACGAACGCGCCG CGACTGATGTTGAGCCCCCTCTCCATG GGC
SDM to make - 8 amino acids in TM6-TM7 loop	GCGGAGTACATGCCCATGGAAGAGGGG CGCGGCGCGTTCGTTCATGGGTATGGGCG	CGCCCATACCCATGACGAACGCGCCG CGCCCCCTCTCCATGGGCATGTACTCC GC
SDM to make - 12 amino acids in TM6-TM7 loop	CATGATCACGCTGGCCGAAGGCGTCTC CCCATAATGGTTCGTGGTTCCGAAGAGAG CG	CGCTCTCTTCGGAACCACGACCATTAT GGGGAGGACGCCTTCGGCCAGCGTGA TCATG
5 missense mutations SDM	D40N/E42S TTCGCCACCGATGAGGGATTTTCAAACG CAACAAGCCCCGCC A147E/V148P GCACGAGCACCGGCAGGGGCTCGAGGG ATATCCCGAAG A229V AGACGTGCGACGACACGACAAGGATCG AG	D40N/E42S GGCGGGGCTTGTGCGTTTGAAAATCC CTCATCGGTGGCGAA A147E/V148P CTTCGGGATATCCCTCGAGCCCCTGCC GGTGCTCGTGC A229V CTCGATCCTTGTCTGTCTGTCGCACGT CT
TEV cleavage site SDM	Round 1 GTTTTCGTGGCTTCCTTTCGTGCGACGAG AATCTGTACGCACTCGAGCACCACC ACC Round 2 CGTTTTCGTGGCTTCCTTTCGTGCGACGAG AATCTGTACTTCCAGTCTGCACTCGAGC ACCACCACCACCACACTGAGATCCGGC	Round 1 GGTGGTGGTGGTGCTCGAGTGCGTAC AGATTCTCGTGCAGAAAGGAAGCCA CGAAAAC Round 2 GCCGATCTCAGTGGTGGTGGTGGT GTGCTCGAGTGCAGACTGGAAGTACA GATTCTCGTGCAGAAAGGAAGCCAC GAAAACG

4.2.4.3 *Mm*SPP-EE expression and purification

The expression and membrane isolation of *Mm*SPP-EE variants is identical to that of *Hm*SPP-EE. The purification is similar to that of *Mm*SPP-EE variants with the following exceptions. DDM is used for membrane isolation in a 1% solution with a ratio of 1:10 DDM:cell mass. For example, if 0.8 g membrane was isolated from 10 g cells, 1 g of detergent was used for a total of 100 mL resuspended membrane in 50 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole. For Ni²⁺-affinity chromatography, the wash buffer contained 50 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole, and 0.1% DDM and the elution buffer contained 50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole, and 0.1% DDM. After Ni²⁺-affinity chromatography, the protein concentrated in a 10K MWCO Amicon Ultra centrifugal filter and was further purified over size exclusion chromatography on either a Sephacryl 300 16/60 or Superose 12 10/300 column (GE Lifescience) using SPP gel filtration buffer containing 0.05% DDM as the running buffer.

4.2.5 **Fab/EE complexation with SPP-EE variants over size exclusion chromatography**

Following size exclusion chromatography, SPP variant was concentrated in a 10K MWCO Amicon Ultra centrifugal filter to <250 µL and incubated in a 1:1 molar ratio with Fab/EE (in 20 mM HEPES pH 7.5, 250 mM NaCl) at 4 °C for 2 hours prior to injection onto the size exclusion column (Superose 12 10/300, Sephacryl 300 16/60, or Sephacryl 200 10/300). SPP gel filtration buffer was used as the running buffer, containing either 0.1% FC12 for *Hm*SPP variants or 0.05% DDM for *Mm*SPP variants.

4.2.6 **Intimin molecular biology, expression, and purification**

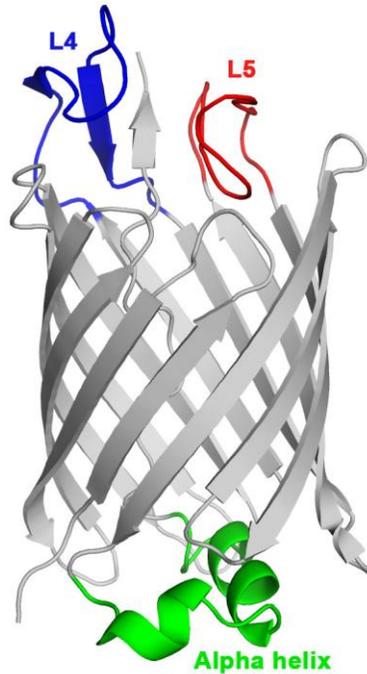


Figure 23 Intimin structure showing location of designed intimin-EE constructs

Amino acids 309-328 (L4) are colored blue, 354-370 (L5) are colored red, and 412-433 (periplasmic α -helix) are colored green.

Table 15 Primers used to make intimin-EE3, -EE4, -EE7, and -EE8 using SDM

Primer Purpose	Forward Primer	Reverse Primer
Intimin-EE3	CTTGGAGTCAGCAGATTGAACCGCAGTA TG TGAATGAATATATGCCCATGGAGGGT TCTCGCTACGATCTGGTTCAGCGTAACA ACA	TGTTGTTACGCTGAACCCAGATCGTAGC GAGAACCCTCCATGGGCATATATTCAT TCACATACTGCGGTTCAATCTGCTGAC TCCAAG
Intimin-EE4	ACAGTATTACGGTGATAATGTGGCCCTG TTAACTCTGAGTATATGCCAATGGAAC CGGGCGCAGCGACGGTGGG	CCCACCGTCGCTGCGCCCGGTTCCATT GGCATATACTCAGAGTTAAACAGGGC CACATTATCACCGTAATACTGT
Intimin-EE7	GCTCTGTTAACGGCTACTTCCGTATGAG TGGTTGGCATGAATATATGCCCATGGAG GATTACGATGAACGCCCGCAAATGGCT TTGATATTTCGTTT	AAACGAATATCAAAGCCATTTGCCGG GCGTTCATCGTAATCCTCCATGGGCAT ATATTCATGCCAACCCTCATACGGAA GTAGCCGTTAACAGAGC
Intimin-EE8	CGGCTACTTCCGTATGAGTGGTTGGCAT GAATACATGCCCATGGAAGATTACGATG AACGCCCGCAAATGGCTTTGATATTCG	CGAATATCAAAGCCATTTGCCGGGGCG TTCATCGTAATCCTCCATGGGCATGTA TTCATGCCAACCCTCATACGGAAAGTA GCCG

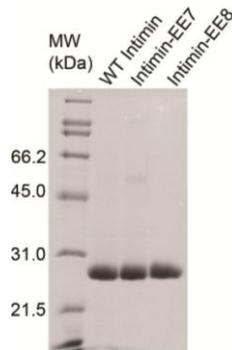


Figure 24 WT intimin, intimin-EE7 and intimin-EE8 SDS-PAGE

Three intimin variants (WT, -EE7 and -EE8) were purified by Ni²⁺-affinity and SEC. Reducing SDS-PAGE shows a single band ~31 kDa for each variant.

4.2.7 Complexation of Fab/EE with intimin-EE7, -EE8

For complexation with intimin, Fab/EE in 50 mM Tris pH 7.5, 200 mM NaCl, 0.01% sodium Azide was combined with WT intimin, intimin-EE7, or intimin-EE8 at a 1:1 molar ratio, and incubated for 2 hours on ice before injection onto a Superose 12 10/300 GL column equilibrated with 50 mM Tris pH 7.5, 200 mM NaCl, 0.01% sodium azide, 0.05% DDM. Elution fractions for each peak were concentrated by Amicon Ultra centrifugal filter (30K MWCO, Millipore) and analyzed with 12% reducing SDS-PAGE.

4.2.8 Molecular dynamic (MD) simulations

The structure of WT intimin (PDB ID 4E1S) was placed in a 140- *n*-dodecylphosphocholine (DPC) micelle using CHARMM-GUI (214, 220). DPC molecules were then mutated to DDM using a modified topology based on the CHARMM force field. Mutations for intimin-EE7 and the additional alanine insertions for intimin-EE2 were made in silico using VMD (221) and minimized using NAMD (222). An 8-residue epitope was docked to Fab/EE using ClusPro2 (223). Equilibration of this bound state over 10 ns was found to be stable and thus was used to model the

placement of the Fab/EE on both intimin-EE7 and intimin-EE8. The systems were then each solvated with water and ionized with 250 mM NaCl. The resulting systems contained approximately 172,000 atoms each. Restraints were applied for a short equilibration to ensure the epitope-containing loop had stabilized in the Fab/EE binding site. The resulting systems were then run using unrestrained MD for 50 ns for intimin-EE7 and 170 ns for intimin-EE8. All simulations were run using NAMD. The CHARMM36 (224) force field was used throughout, along with the TIP3P water model (225). A 2-fs time step was used for all bonded and short-range interactions, with long-range non-bonded electrostatics calculated every other time step using particle mesh Ewald (226). A uniform temperature of 310 K and pressure of 1 atm were maintained.

4.3 RESULTS

4.3.1 EE peptide insertion into the extracellular loop of membrane proteins

To determine candidate locations for EE peptide insertion into A₂aR, the repertoire of previously solved GPCR structures was analyzed, revealing that most previously solved GPCR structures involved modification of the flexible ICL3 (227), either by truncation or through the use of covalent (T4L or BRIL) (228) or non-covalent (Fab) (186) crystallization chaperones targeting ICL3 identified by hybridoma technology (229). Specifically, 10 of the 12 residues that form the Fab epitope in the complex structure with the β_2 adrenergic receptor (186) are contiguous in primary sequence, suggesting that the EE peptide inserted into the homologous location in A₂aR would be accessible by our Fab/EE. Thus, the EE peptide flanked by flexible linkers (sequence: GS-EYMPME-GS), was inserted into the N-terminal segment of ICL3 of an A₂aR-green fluorescence protein-construct, after K209 without mutation or truncation of

wild-type residues (**Figure 21a**) to generate A_{2a}R-GFP-EE. A_{2a}R-GFP and A_{2a}R-GFP-EE were expressed and purified from *S. cerevisiae*. Silver staining SDS-PAGE reveals predominantly monomer A_{2a}R-GFP-EE and Western blot shows EE peptide incorporation into only A_{2a}R-GFP-EE and not wild-type A_{2a}R-GFP (**Figure 21b**).

Available loops for the insertion of the EE epitope in the representative β -barrel membrane protein, *E. coli* intimin, were determined based on the high resolution crystal structure (230). Intimin was engineered with an internal EE peptide in one of three locations, the L4 region (intimin-EE7), L5 region (intimin-EE4), or the periplasmic α -helix (intimin-EE3) (**Figure 23**). Due to poor expression of intimin-EE3 and -EE4, further experiments were not conducted. An additional construct possessing an insertion of 2 alanine residues on either side of the EE peptide in intimin-EE7 was generated, analogous to A_{2a}R, which we postulated would increase the accessibility of the EE peptide for better Fab/EE binding (**Figure 23**). Both intimin-EE7 and intimin-EE8 were expressed and purified in similar protein yields to WT intimin (**Figure 24**) (214).

Though the *MmSPP* structure was published in early 2013 (4), the designed EE tagged variants of SPP predate the structure and the placement of the EE tag in *HmSPP* and *MmSPP* and was chosen based on predicted loops. The loop between TM6 and TM7 is very long, and we postulated that it would be accessible for Fab/EE binding and that the addition of Fab/EE would help stabilize the long loop. The EE tag was placed in three locations within the TM6-TM7 loop of *MmSPP* to test whether placement affected protein expression (**Figure 22**). Since expression was largely unaffected (data not shown), *MmSPP*-EE2 was chosen for complexation and crystallization experiments because it is in the center of the TM6-TM7 loop.

4.3.2 Fab/EE forms complexes with EE-tagged membrane proteins

4.3.2.1 Model α -helical membrane protein A₂aR

The accessibility and specific binding of Fab/EE to the EE peptide as presented by the A₂aR-GFP was limited to ELISA (**Figure 25a**) and SPR (**Figure 25b-c**) due to the low yields of purified protein. For both experiments Fab/EE binding was only detected for EE-tagged proteins. The calculated K_D from SPR for A₂aR-GFP-EE (**Table 16**) is higher than the K_D for the soluble proteins MBP-EE, MBP-KEE and scFv-EE₁ (**Table 10**), likely due to the presence of detergent.

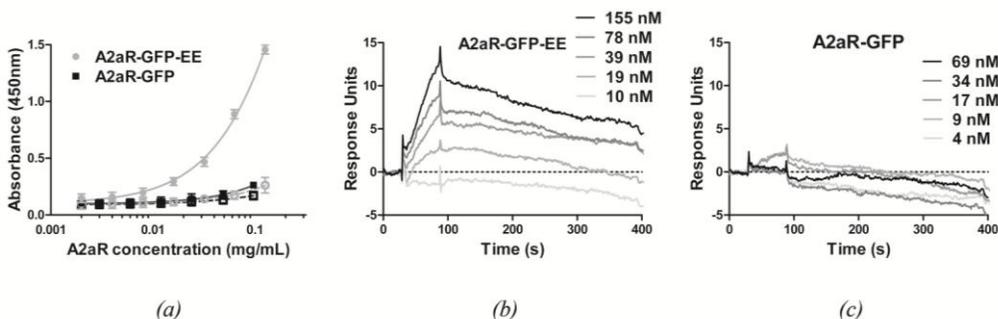


Figure 25 Fab/EE binding to EE-tagged A₂aR-GFP

(a) ELISA analysis of Fab/EE binding to A₂aR-GFP. Purified A₂aR-GFP proteins with or without the EE peptide insertion were added to ELISA wells coated with Fab/EE or blocked control wells. Captured protein was detected using an anti-GFP antibody. (b-c) SPR analysis of binding. GPCRs were injected in duplicate for each concentration tested and binding to immobilized Fab/EE was monitored by SPR. Average traces are shown and demonstrate concentration-dependent binding to A₂aR-GFP-EE. No binding was observed for A₂aR-GFP

Table 16 Characterization of EE-tagged protein binding kinetics of Fab/EE by SPR

	k_{on} ($M^{-1}\cdot s^{-1}$)	k_{off} (s^{-1})	K_D (nM)
A2aR-GFP-EE	$2.13 \pm 1.60 \times 10^5$	$3.92 \pm 1.96 \times 10^{-3}$	32 ± 24

4.3.2.2 Model β -barrel membrane protein intimin

The ability of Fab/EE to form a solution complex with an EE-tagged β -barrel membrane protein was tested with the intimin constructs by SEC (**Figure 26**). The

Fab/EE:WT intimin elution trace has two peaks, the first representing WT intimin, the second Fab/EE, as confirmed by SDS-PAGE (**Figure 26 inset, top**). By contrast, when intimin-EE7 or -EE8 is mixed with Fab/EE, the species coelute. The first peak in the SEC trace has a shorter retention time indicative of a higher molecular mass complex expected for Fab/EE:intimin-EE complexes, concomitant with a reduction in the peak corresponding to Fab/EE alone (**Figure 26 inset, middle and bottom**).

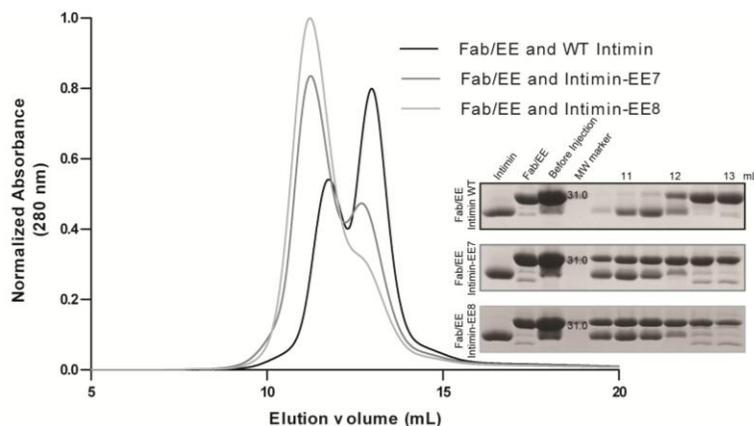


Figure 26 Elution profile of purified Fab/EE incubated with intimin WT, intimin-EE7 or intimin-EE8

Reducing SDS-PAGE analysis (*inset*) of fractions selected from 10.5 mL to 13 mL elution volumes shows coelution of Fab/EE:intimin-EE1 and Fab/EE:intimin-EE2

4.3.2.3 A-helical membrane proteins *HmSPP* and *MmSPP*

As further indication that Fab/EE complexes EE-tagged membrane proteins and as a step towards the crystallization of the Fab/EE:SPP-EE complex, the solution binding of Fab/EE to *HmSPP*-EE (**Figure 27**) and *MmSPP*-EE2 (data not shown) was also studied using SEC. Little to no coelution was detected with WT *HmSPP* and Fab/EE (orange). Fab/EE elutes in the first peak with *HmSPP*-EE (green), indicating that the two are binding over SEC. Similar results were obtained for *MmSPP*-EE2.

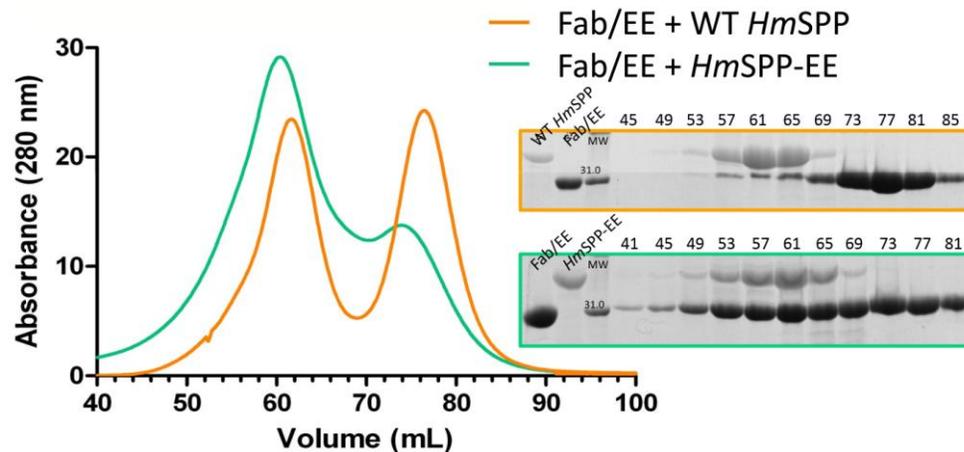


Figure 27 Fab/EE complexes with HmSPP-EE over SEC

SEC was done on a sephacryl 300 16/60 column with Fab/EE and either WT *HmSPP* (orange) or *HmSPP-EE* (green). The elution volume of each fraction is listed at the top of the SDS-PAGE. Samples of Fab/EE and *HmSPP* prior to complexation are shown to the left of the MW marker.

4.3.3 MD simulations of Fab/EE-intimin-EE complexes

We postulated that upon binding of Fab/EE to intimin-EE7, intimin-EE8, *HmSPP-EE* and *MmSPP-EE2*, loop residues would be immobilized and a stable complex for crystallization would be generated. Nearly 8000 crystallization trials of Fab/EE and EE-tagged membrane proteins have been set up (~35 x 96 conditions from commercial screens and 29 x 24 conditions from homemade screens for intimin-EE, and 36 x 96 conditions from commercial screens and 10 x 24 conditions from homemade screens for SPP-EE), with varying temperatures, concentrations, method (vapour diffusion with detergent (DDM), vapour diffusion with bicelle (3-([3-Cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO) and 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) mixture), or lipidic cubic phase (monoolein), data not shown). Both direct mixing of 1:1 Fab/EE : membrane protein-EE variant immediately prior to setting up trays and using the complex isolated from SEC as

above, were attempted. Unfortunately, none yet has resulted in diffracting crystals of Fab/EE:intimin-EE or any Fab/EE:SPP-EE complex (data not shown).

To gain insight into why co-crystallization trials have not been successful in spite of demonstrated favorable solution properties, we turned to molecular dynamics. Since the structure of *Mm*SPP is low resolution and missing the loop where the EE tag was inserted, MD studies were done on the Fab/EE:intimin-EE complexes. WT intimin, intimin-EE7, and intimin-EE8 were first modeled and allowed to equilibrate over 50 ns (**Figure 28a**). In comparison to WT intimin, intimin-EE7 and intimin-EE8 exhibit increased flexibility in residues 315-320, where the WT residues were mutated to the EE-tag. Notably, the insertion of 2 alanine residues on each side of the EE-tag in intimin-EE8 did not seem to cause a further increase in flexibility compared to intimin-EE7. After modeling the Fab/EE:intimin-EE interaction, the complex of Fab/EE:intimin-EE7 was equilibrated for 50 ns and Fab/EE:intimin-EE8 complex for 170 ns. The RMSF of residues 300-340 after 50 ns with and without Fab/EE bound (**Figure 28b**) indicate Fab/EE binding to both intimin-EE7 and intimin-EE8 slightly increases the flexibility of residues 312-325. Over the 170 ns simulation for the Fab/EE:intimin-EE8 complex, the position of Fab/EE in relation to intimin-EE8 is dynamic (**Figure 28c-d**). Such flexibility is likely due to mutations in the L4 loop in the membrane protein (**Figure 29a**). In particular, S316 of WT intimin forms hydrogen bonding interactions in both the WT intimin and the structure after 50 ns of MD simulations (**Figure 29b**). This interaction was broken when S316 was mutated to tyrosine within EYMPME of intimin-EE7, resulting in drastic changes within the loop (**Figure 29c**). The corresponding residue in intimin-EE2 is A314B (see numbering in **Table 14**) and after 50 ns simulation, the loop

resembles that of WT intimin but with E315 of intimin-EE8 in a similar position as the WT intimin residue S316 (**Figure 29d**).

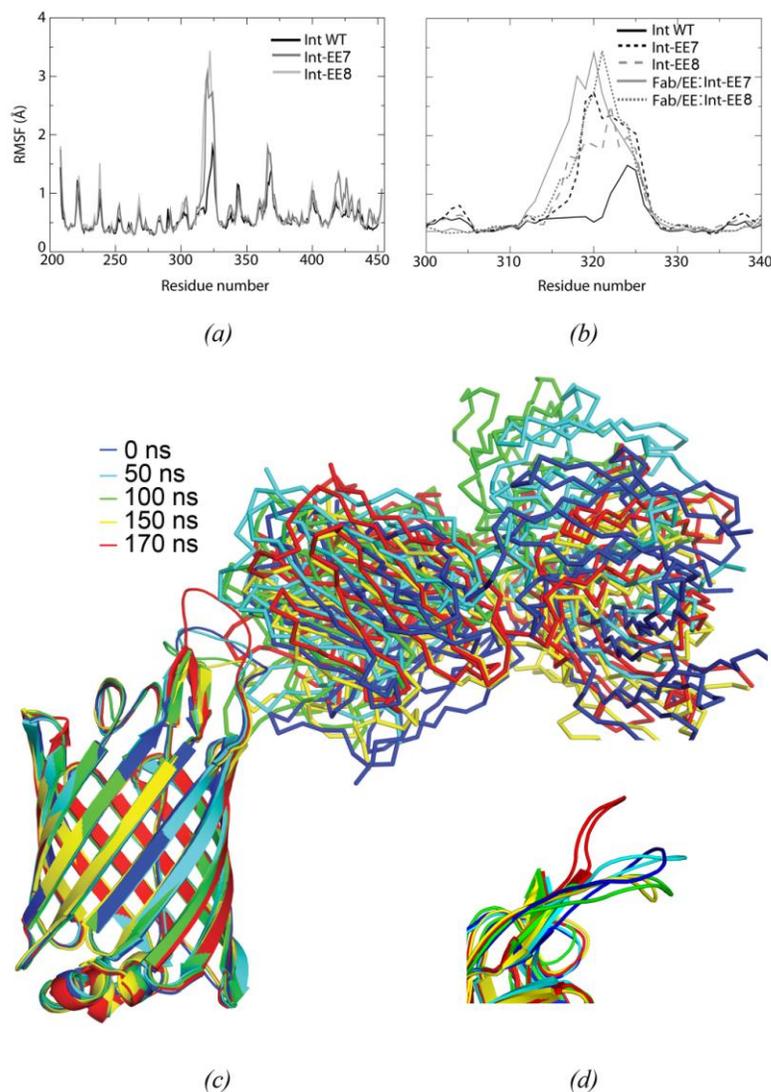


Figure 28 Molecular dynamics analysis of intimin-EE1 and intimin-EE2 with and without Fab/EE

(a) Root mean squared fluctuation (RMSF) of residues 200-450 for intimin WT, intimin-EE7, and intimin-EE8 over 50 ns simulation. (b) RMSF of residues 300-340 for intimin WT, intimin-EE7, intimin-EE8 alone and in modelled complex with Fab/EE over 50 ns simulation. (c) Fab/EE:intimin-EE8 proposed complex structure at different time points during the simulation, colored from blue to red. Intimin-EE8 is shown as a cartoon and Fab/EE as ribbon. (d) A zoomed view of the intimin-EE8 loop containing the EE peptide during the same timepoints as in *c* with Fab/EE omitted for clarity.

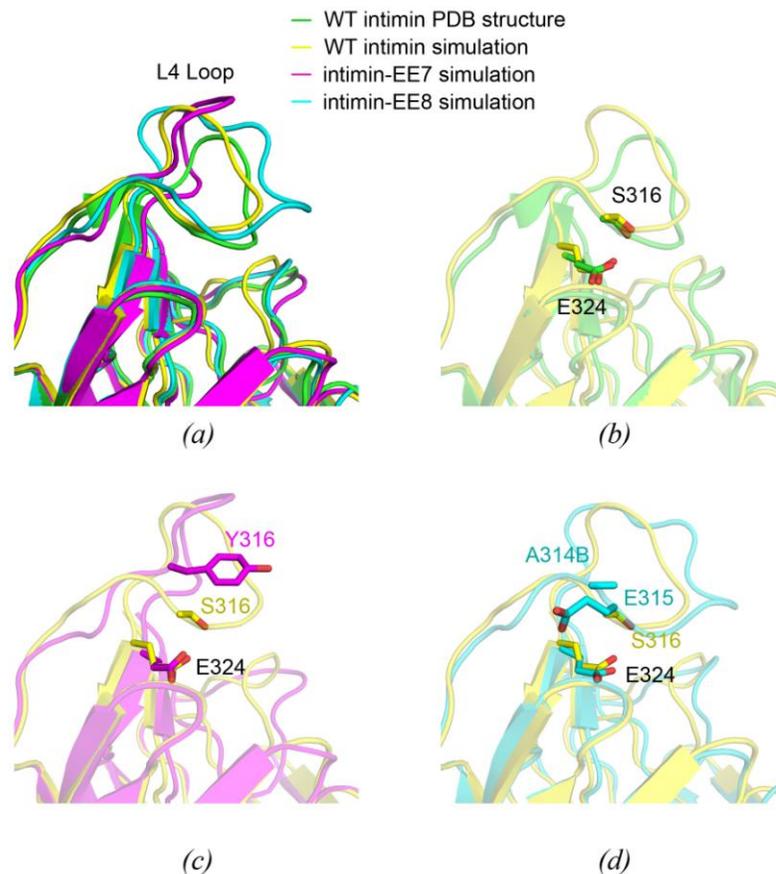


Figure 29 Analysis of intimin L4 interactions

(a) L4 loops of 4 intimin structures: WT intimin (PDB ID 4E1S), and WT intimin, intimin-EE7 and intimin-EE8 after 50 ns simulation. (b) L4 loop of WT intimin from the PDB structure and WT intimin after 50 ns simulation. S316 is hydrogen bonding to E324 in 4E1S, and interaction is unchanged after 50 ns simulation. (c) WT intimin and intimin-EE7 after 50 ns equilibration showing L4 loop residue 316 and E324 for each structure. (d) WT intimin and intimin-EE8 after 50 ns equilibration showing L4 loop residues S316 and E324 of WT intimin, A314A and E315 of intimin-EE2.

4.3.4 Fab/EE forms a solution complex with SPP constructs with truncated TM6-TM7 loop

Because the MD studies indicated that the insertion of the EE tag into a long soluble loop is likely detrimental to crystallization of the complex due to loop flexibility, *MmSPP* constructs with a shortened TM6-TM7 loop were made using SDM by deleting 4 residues at a time. Because the deleted residues are near the EE tag, we were uncertain if

the complexation over SEC would be successful. The SEC of Fab/EE and WT *MmSPP* resulted in 4 peaks (**Figure 30**, blue). The first small peak (labeled 1) was too low in concentration to be visible on SDS-PAGE. The second peak (labeled 3) contained only WT *MmSPP*. The third and fourth peaks, (labeled 4 and 5) contained 2 different species of Fab/EE. Both of the SECs with the shortened EE tagged variants contained 3 peaks (**Figure 30**, red and yellow). In each case, the first small peak contained some form of aggregated *MmSPP*-EE2S (2 or 3). The second peak (labeled 2) contained the Fab/EE:*MmSPP*-EE2S (2 or 3) complex. The third peak (labeled 4) contained uncomplexed Fab/EE.

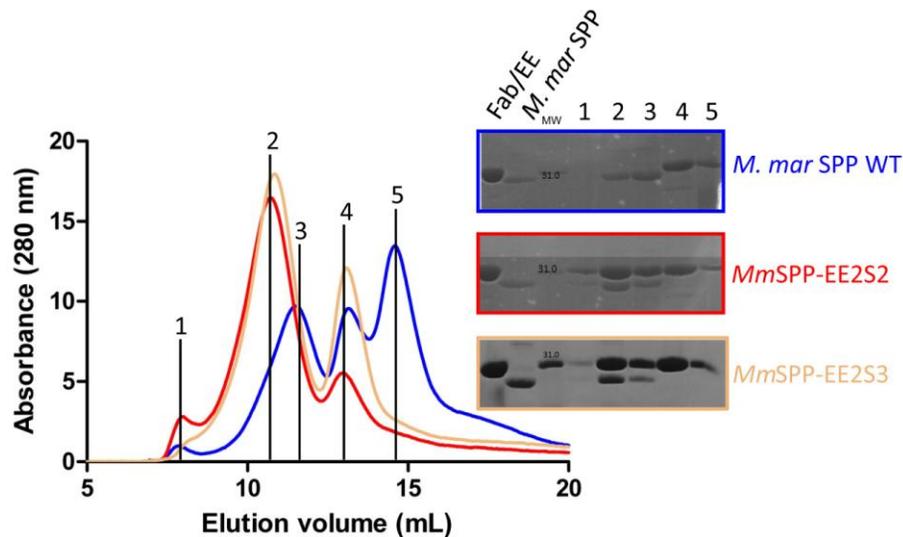


Figure 30 Fab/EE complexes with *MmSPP*-EE2S constructs

The complexation of Fab/EE with WT *MmSPP* (blue), *MmSPP*-EE2S2 (red), and *MmSPP*-EE2S3 (yellow) was done using SEC on a superose 12 10/300 column. The same fractions from each SEC were run on SDS-PAGE and the fraction locations are marked with black lines. Fab/EE and *MmSPP* samples prior to complexation were also run on SDS-PAGE on the left side of the MW marker.

Crystallization trials of the Fab/EE:*Mm*SPP-EE2S3 complex as isolated over SEC have been done (8 x 96 conditions from commercial screens). Only a single well contained 2 rod shaped crystals (~25 x 5 µm in size, grown in 0.2 M ammonium sulphate, 14 % PEG 4000, 0.1 M CHES pH 10, 5.94 mg/mL complex, 20 °C for 23 days), but the crystals dissolved after being under the microscope for a brief period of time, and were not reproducible.

4.4 DISCUSSION

4.4.1 Conclusions: implications for chaperone-mediated crystallization of MPs

Our engineered Fab/EE exhibits numerous favorable characteristics for use as a crystallization chaperone for difficult targets. Introduction of the EE epitope within membrane proteins is straightforward by site directed mutagenesis, leading to readily detected high affinity solution complexes. However, even when placement of the EE epitope does not interfere with protein expression and purification, an unintended consequence may be the removal of native contacts and thus an increase in conformational heterogeneity that is detrimental to crystallographic efforts. Such conformational changes can be challenging to predict, even for test proteins whose structures are known, and would be especially enigmatic in the case of target proteins of unknown structure. If a less flexible region of the target protein is not known, a reduction in loop flexibility is likely achievable by shortening the EE epitope-containing loop to a minimum number of residues that can still be complexed with Fab/EE. As shown in **Figure 30**, this is an ongoing project in the lab and will likely take time to determine the optimal loop length for the complexation and crystallization of Fab/EE with the shortened

constructs. This approach may be successful in combination with other methods that were discussed in Chapter 1.

CHAPTER 5: MONOMERIC STREPTAVIDIN AS A CRYSTALLIZATION CHAPERONE FOR A BIOTINYLATED SOLUBLE TEST PROTEIN

5.1 INTRODUCTION

Streptavidin is a ~15 kDa protein from the bacterium *Streptomyces avidinii* (231) that forms a homotetramer to bind up to 4 molecules of D-biotin with unusually high affinity ($K_D=10^{-14}$ M) (232). The crystallographic structure of streptavidin was solved in 1989 in the apo and biotin-bound forms (**Figure 31**) (233).

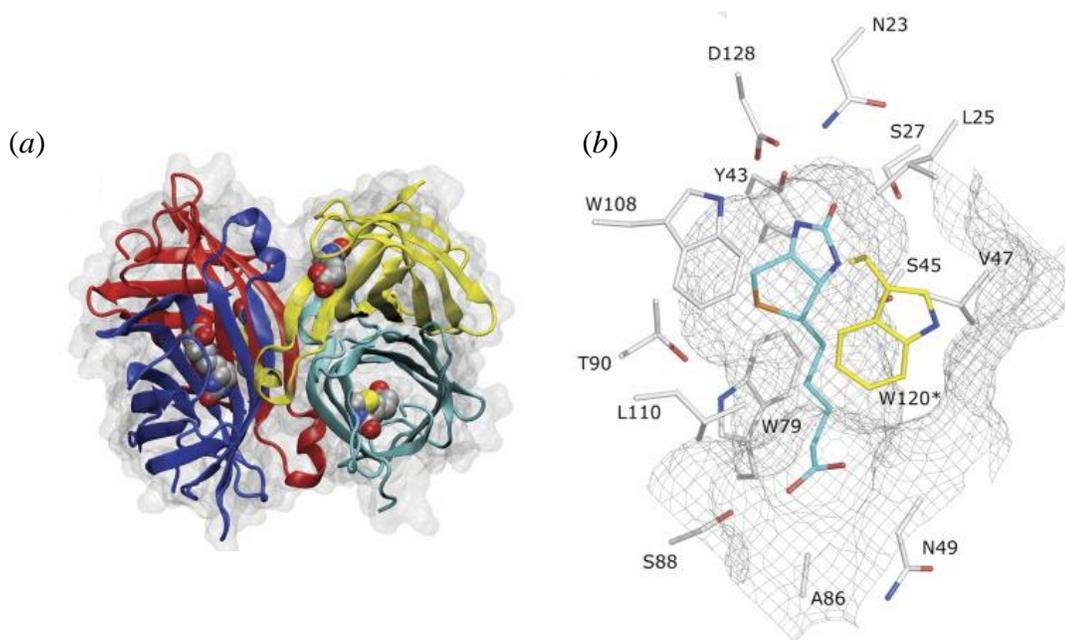


Figure 31 Structure of WT streptavidin

(a) The tetrameric structure of streptavidin is represented as four monomers in different colors with secondary structure shown as cartoon (PDB ID 1MK5 (8)). The bound biotin is shown as van der Waal's spheres and surface representation is transparent grey. (b) The ligand binding cavity is shown as grey mesh and sidechains of residues within 4 Å are shown as sticks. Carbon atoms of the bound biotin are shown in cyan. W120 is colored in yellow, and is part of the hydrophobic lid from a neighboring subunit (9). Figure was reprinted with permission.

The streptavidin-biotin complex has been used in a range of biological applications including ELISA, immunohistochemistry, cell-surface labeling, affinity purification, and FACS (9) because the complex is strong with rapid binding and resistance to extremes in pH, temperature, and denaturing agents (234). Tetravalent streptavidin, however, has drawbacks. The binding of multiple biotinylated ligands can cause aggregation of the target protein, which could alter the biological function and cause artifacts. For example, when used for labeling cell surface receptors, streptavidin can cause artificial crosslinking, which alters cellular signaling and induces receptor internalization (6). Germane to this thesis, as a crystallization chaperone, tetravalent streptavidin could induce protein aggregation or sample heterogeneity.

To address the issue of multivalency of streptavidin, three approaches have been taken. The first is to reduce the number of biotin binding sites within the tetramer. A heterotetramer containing two biotin-binding and two non-binding W120A mutant subunits was prepared by mixing mutant and WT streptavidin before in vitro refolding (235). Similarly, a monovalent streptavidin has been made with three non-binding mutant and one WT subunit (236). The monovalent streptavidin has femtomolar affinity for biotin the single biotin binding site, but the preparation process is cumbersome and the tetrameric protein is large which can prevent binding to partially buried biotin sites. The second approach to engineer a non-tetravalent streptavidin is by constructing a single-chain version of the tetramer so that each biotin binding site can be engineered separately on one plasmid using molecular biology. The topology of the subunits was modified so that the C-terminus of one chain was near the N-terminus of another chain and the chains

were connected by a short linker region (9, 237). Again, the large size of the tetrameric protein could be detrimental to some labeling studies.

An third approach to obtain monovalent streptavidin is to engineer a stable monomer. Most of the reported monomers are unstable, prone to aggregation, and exhibit low affinity for biotin (238-240). Some have concluded that the loss of biotin affinity was due to the absence of the hydrophobic lid from adjacent subunits (**Figure 31**) (241, 242), but a dimeric version engineered by the Park lab at State University of New York (SUNY) Buffalo indicates that something else is responsible for the high biotin affinity (122). The dimer consists of two covalently linked subunits and contains the same binding site residues as the tetramer, but the affinity for biotin is 17 nM. This suggests that tetramerization is not the only parameter that causes the high affinity for biotin. To engineer a monomeric streptavidin, the Park lab introduced biotin binding-site residues of rhizavidin, a biological streptavidin dimer (243), into a streptavidin monomer. Other stabilizing mutations were introduced based on rational design (**Figure 32**). The resultant monomer (mSA) has 55% sequence identity with streptavidin and 57% identity with rhizavidin. The mutations improved the affinity of the monomer for biotin by 13-fold and led to the highest reported thermal stability of any engineered monomeric streptavidin (6).

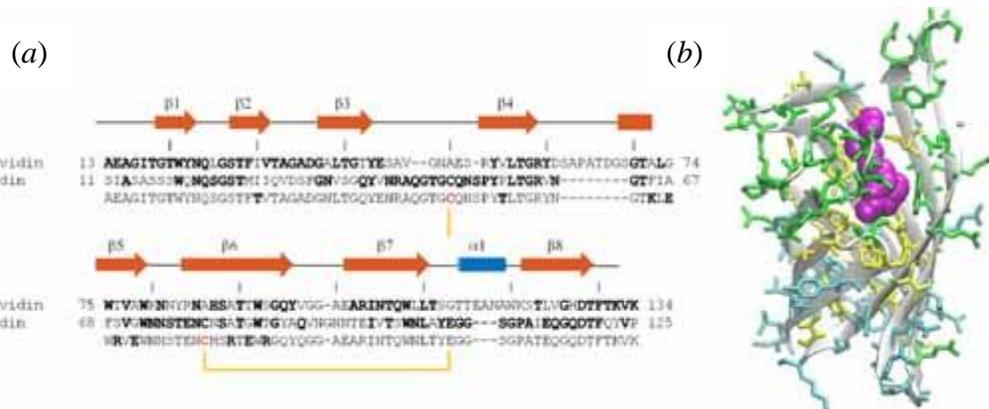


Figure 32 Design of monomeric streptavidin based on rhizavidin dimer

(a) Alignment of streptavidin and rhizavidin sequences based on structure. Residues in the final hybrid mSA are in bold and the engineered disulfide bond is shown in yellow. (b) The modelled structure of mSA. Residues of streptavidin origin are in cyan, rhizavidin origin are in green, and those common to both proteins are in yellow. The residues that differ from both streptavidin and rhizavidin are not shown. Biotin is in purple van der Waals spheres (6). Figure reproduced with permission.

One unanticipated feature of mSA is rapid dissociation of bound biotin, which limits the usefulness of the streptavidin monomer for the aforementioned applications. To overcome this issue, the Park lab introduced an additional mutation, L25H, which slowed the dissociation rate (k_{off}) of the new mSA2 to $4.2 \times 10^{-4} \text{ s}^{-1}$, and yielded a two-fold improvement in affinity ($K_d=0.52 \text{ nM}$) (6). Besides its potential suitability for a variety of research applications, a major practical improvement for mSA2 was the addition of an MBP fusion tag to the N-terminus (MBP-mSA2) to allow for the expression of soluble protein, instead of inclusion bodies.

The MBP-mSA2 plasmid was generously gifted to us by Dr. Sheldon Park to explore its use as a crystallization chaperone. Since monomeric streptavidin can bind biotin with high affinity, is very stable, and easily crystallized, we hypothesized mSA2 could potentially be a general crystallization chaperone for biotinylated proteins. As described in this chapter, our system uses biotin ligase (BirA) to enzymatically

biotinylate a specific lysine residue within a 15 amino acid tag (GLNDIFEAQKIEWHE, Avi-tag), which is inserted into a loop within the test protein of interest, in this case, MBP (MBPavi). Complexation of mSA2 bound to MBPavi and crystallization trials are discussed.

5.1.1 Individual contributions to the work

The use of monomeric streptavidin as a crystallization chaperone has been an idea in the lab since its inception. Dr. Derrick Watkins and Lindsey Porter experimented with converting the tetrameric streptavidin to monomeric and refolding from inclusion bodies. Dr. Park then shared the plasmid for an early, less stable version of streptavidin monomer (Tr-mStrav). By using Lindsey's protocols, I was successful with the refolding of Tr-mStrav and binding to a chemically biotinylated membrane protein (data not shown) with the help of 2013 REU summer student Rachel Wills. The project in its current form has been optimized by me and I planned and conducted all the experiments presented in this chapter.

5.2 METHODS

5.2.1 Protein expression and purification

5.2.1.1 BirA BL21 expression and amylose-affinity chromatography

BirA plasmid (pET-28a/BirA) was generously provided by Dr. Sheldon Park and encodes MBP-BirA fusion, with an N-terminal hexahistidine-tag and a TEV protease site between MBP and BirA. Expression and purification was carried out as previously reported (244), with some exceptions. Briefly, after transformation of the plasmid into *E. coli* BL21(DE3), a single colony was placed in 100 mL LB containing 50 µg/mL kanamycin. Culture was incubated overnight at 37 °C with shaking at 225 RPM. The

starter culture was diluted 1:100 in each 1 L fresh LB culture supplemented with 50 $\mu\text{g}/\text{mL}$ kanamycin, and cells were grown at 28 °C with shaking at 225 RPM until the OD_{600} reached 0.6. The temperature was then dropped to 18 °C, and expression was induced by the addition of IPTG to a final concentration of 1 mM. Cells were cultivated for 22 hours then harvested by centrifugation at 4225 x g for 10 minutes. Cell pellet was flash frozen in liquid nitrogen and stored at -80 °C.

Cells mass corresponding to ~3 L culture were thawed and resuspended in BirA cell lysis buffer (25 mM Tris pH 8, 200 mM NaCl) and lysed by French press using 18,000 PSI cell pressure, passing through twice. Lysate was rocked at 4 °C for 30 minutes then centrifuged at 47,800 x g for 40 minutes. Lysate was then purified using either amylose- or Ni^{2+} -affinity chromatography. For amylose-affinity chromatography, BirA cell lysis buffer was used for the wash, and cell lysis buffer supplemented with 20 mM maltose was used for elution. For Ni^{2+} -affinity chromatography, cell lysis buffer was supplemented with 20 mM imidazole for the wash, and 400 mM imidazole for the elution (**Figure 34a**). Since BirA activity is inhibited by NaCl, the purified protein is then exchanged into 25 mM Tris pH 8 via gel filtration on Superdex 75 16/60 column or by using an 30K MWCO Amicon Ultra centrifugal filter. MBP-BirA fusion protein does not need to be cleaved for biotin ligase activity. Henceforth, BirA refers to the fusion MBP-BirA . Purified BirA can be stored at 4 °C for at least 2 weeks. For longer term storage, BirA in 50% glycerol can be stored at -20 °C.

5.2.1.2 Molecular biology, BL21 expression and purification of MBPavi

The pAK400 expression vector with the MBP-KEE gene (see Section 3.2.5) was modified using two rounds of site directed mutagenesis to mutate EYMPME to the 15

amino acid Avi-tag GLNDIFEAQKIEWHE using primers listed in **Table 17**; see **Figure 33** for alignment of relevant residues of WT MBP with each round of SDM for MBPavi prepared using Clustal Omega (150). Plasmid fidelity is confirmed by DNA sequencing (MWG Operon, www.operon.com).

Table 17 Primers used for both rounds of SDM to make MBPavi

	Forward Primer	Reverse Primer
SDM Round 1	MBP-Avi_Rd1_For	MBP-Avi_Rd1_Rev
	GCT GAT TGC TGC TGA CGG	CCC ACG TCT TTA ATG TCC
	GGG TTA TGC GTT CAA GGG	TCC ATG GGC ATG TAT TCC
	CCT GAA TGA TAT TTT TGA	TGC GCT TCA AAA ATA TCA
	AGC GCA GGA ATA CAT GCC	TTC AGG CCC TTG AAC GCA
	CAT GGA GGA CAT TAA AGA CGT GGG	TAA CCC CCG TCA GCA GCA ATC AGC
SDM Round 2	MBP-Avi_Rd2_For	MBP-Avi_Rd2_Rev
	GCG CCA GCG TTA TCC ACG	CGT TCA AGG GCC TGA ATG
	CCC ACG TCT TTA ATG TCT	ATA TTT TTG AAG CGC AGA
	TCA TGC CAC TCG ATT TTC	AAA TCG AGT GGC ATG
	TGC GCT TCA AAA ATA TCA	AAG ACA TTA AAG ACG
	TTC AGG CCC TTG AAC G	TGG GCG TGG ATA ACG CTG GCG C

```

(a) MBP          MFNLQEPYFTWPLIAADGGYAFKYENGK-----YDIKDVGVNDAGAKAGLTFLLVDL
MBP-AviRound1  MFNLQEPYFTWPLIAADGGYAFKGLNDIFEAOEYMPMEIDIKDVGVNDAGAKAGLTFLLVDL
                ***** * *****

MBP          IKNKHMNADTDYSIAEAAFNKGETAMTIN
MBP-AviRound1 IKNKHMNADTDYSIAEAAFNKGETAMTIN
                *****

(b) MBP          MFNLQEPYFTWPLIAADGGYAFKYENGK-----YDIKDVGVNDAGAKAGLTFLLVDL
MBP-AviRound2  MFNLQEPYFTWPLIAADGGYAFKGLNDIFEAQKIEWHEIDIKDVGVNDAGAKAGLTFLLVDL
                ***** * *****

MBP          IKNKHMNADTDYSIAEAAFNKGETAMTIN
MBP-AviRound2 IKNKHMNADTDYSIAEAAFNKGETAMTIN
                *****

```

Figure 33 Alignment of WT MBP and MBPavi after each round of SDM

Clustal Omega alignment of WT MBP with MBPavi after SDM round 1 (a) and round 2 (b). EYMPME from MBP-KEE is underlined, and the Avi-tag is inside the red box.

MBPavi was expressed in the same manner as Fab/EE (see Section 3.2.1, (120)).

Briefly, MBPavi was expressed in *E. coli* BL21(DE3). 2 mL LB (Fisher) culture

supplemented with 34 µg/mL chloramphenicol was inoculated with a single colony and incubated for ~4 hours at 37 °C with shaking at 225 RPM. The starter culture was diluted 1:100 in 500 mL Terrific Broth (TB, Fisher) in a 2 L baffled flask and grown overnight with shaking at 225 RPM and 25 °C. Cells were pelleted at 4200 x g for 10 minutes and 4 °C, then the cell pellet was resuspended in 500 mL fresh TB in 2 L flask and incubated for 1 hour at 25 °C, 225 RPM before inducing expression with 1 mM IPTG for 4.5-5 hours. Cells were pelleted and flash frozen in liquid nitrogen before placing in -80 °C freezer. Cells were then thawed and resuspended in 25 mM Tris pH 8, 250 mM NaCl and lysed by passing through French press at least twice at 18,000 PSI cell pressure. Cellular debris was pelleted by ultracentrifugation at 181,000 x g and lysate was purified by amylose-affinity chromatography using the same buffers as for BirA (**Figure 34b**).

5.2.1.3 MBP-mSA2 BL21 expression, purification, and cleavage

The pET-28a vector containing the gene for MBP-mSA2 was generously provided by Dr. Sheldon Park. pET-28a/MBP-mSA2 plasmid was transformed into BL21(DE3) *E. coli* cells and a single colony was used to inoculate 100 mL LB containing 50 mM KH₂PO₄, 50 mM Na₂HPO₄, 25 mM (NH₄)₂SO₄, 1 mM MgSO₄, 0.5% glycerol, 1% glucose, and 50 µg/mL kanamycin. After overnight growth at 37 °C and 225 RPM, the starter culture was diluted 1:100 in 1 L LB containing 50 mM KH₂PO₄, 50 mM Na₂HPO₄, 25 mM (NH₄)₂SO₄, 1 mM MgSO₄, 0.5% glycerol, 0.5% glucose, and 50 µg/mL kanamycin. The cultures were incubated at 37 °C and 225 RPM until the OD₆₀₀ reached 0.8-1.0 at which point expression was induced by the addition of 50 µM IPTG. Protein expression was carried out for 16 hours at 20 °C. Cells were harvested by

centrifugation at 4300 x g for 10 minutes and flash frozen in liquid nitrogen before storing at -80 °C.

Cells (~5 g) harboring MBP-mSA2 were thawed on ice and resuspended in 15-25 mL MBP-mSA2 cell lysis buffer (25 mM Tris pH 8, 150 mM NaCl, 1 mM EDTA, 10% glycerol, 0.2% Triton X-100, 0.2 mg/mL lysozyme, 0.9 mM biotin, and Roche protease inhibitor) and incubated on ice for 30 minutes prior to lysis by French press at 18,000 PSI. Cellular debris was pelleted by centrifugation at 18,000 x g for 20 minutes at 4 °C. Cell lysate was purified by amylose-affinity chromatography using 50 mM Tris pH 8, 150 mM NaCl, 1 mM EDTA, 0.9 mM biotin as the wash buffer and elution buffer containing 20 mM maltose (**Figure 34c**).

5.2.1.4 MBP-mSA2 TEV cleavage

Purified MBP-mSA2 was exchanged into 50 mM Tris pH 8, 150 mM NaCl, 0.9 mM biotin, 0.5 mM EDTA and 1.0 mM DTT and concentrated to ~10 mg/mL using a 30K MWCO Amicon Ultra centrifugal filter. TEV protease (in PBS, see Chapter 2 for TEV protease expression and purification) was added in a 10:1 MBP-mSA2:TEV protease mass ratio. The sample was incubated at room temperature for 1 hour, then heated in a 75 °C water bath for 10 minutes. After cooling sample on ice for about 2 minutes, the precipitate was pelleted by centrifugation at 10,000 x g for 10 minutes. Supernatant was further purified on a Superdex 75 16/60 gel filtration column, and fractions containing cleaved mSA were concentrated in a Amicon Ultra centrifugal filter with MWCO of 3K. (**Figure 35**)

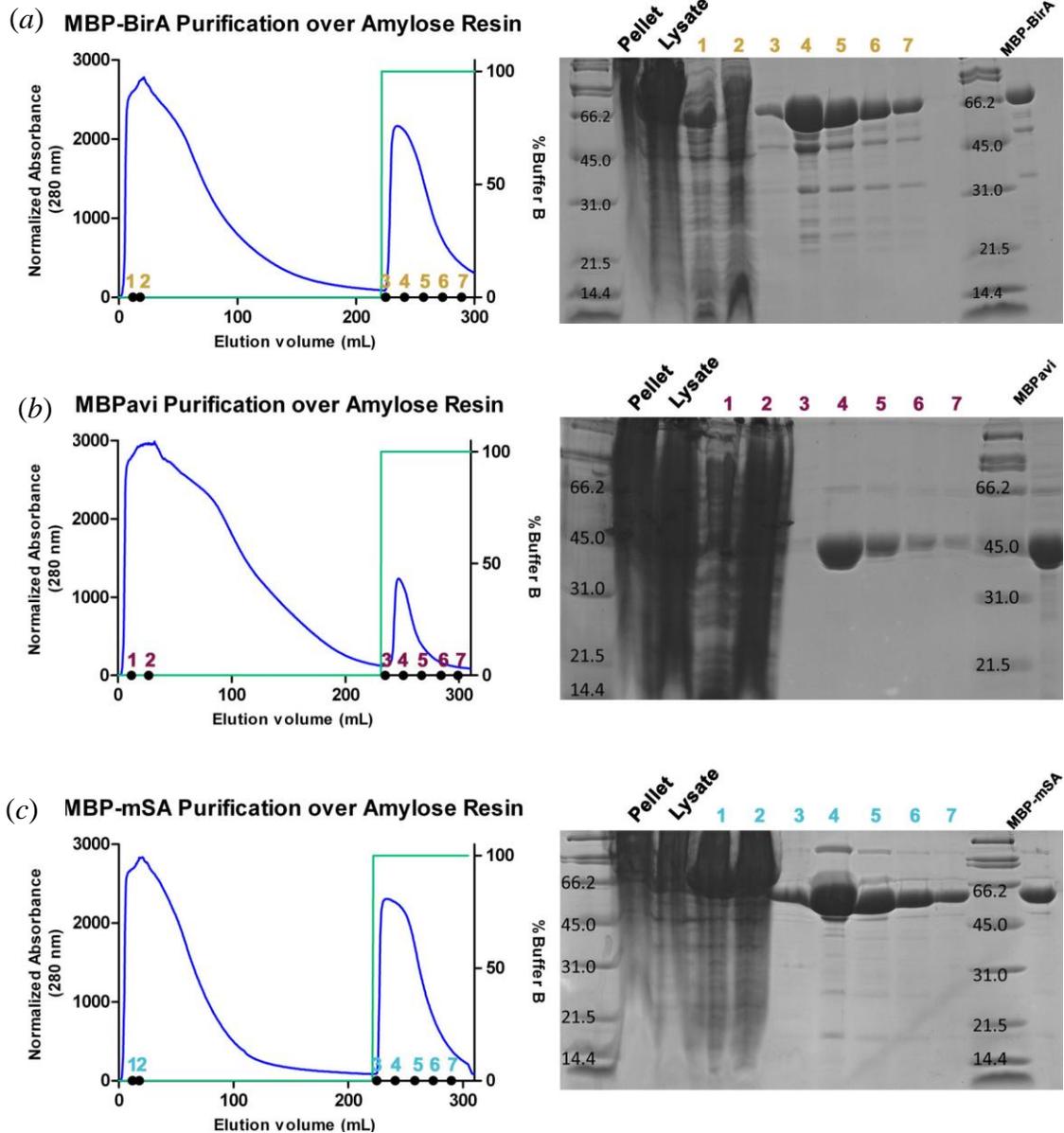


Figure 34 Purification of BirA, MBPavi, and MBP-mSA using amylose resin

(a) BirA, (b) MBPavi, and (c) MBP-mSA were purified using amylose resin. Each gel shows the cell debris (pellet) after lysis and centrifugation, the cell lysate before purification, two flowthrough fractions, and elution fractions from the amylose purification. Fraction locations are marked with black dots along the x-axis of the chromatogram. Elution fractions were pooled and concentrated, and the purified sample is shown in the last lane.

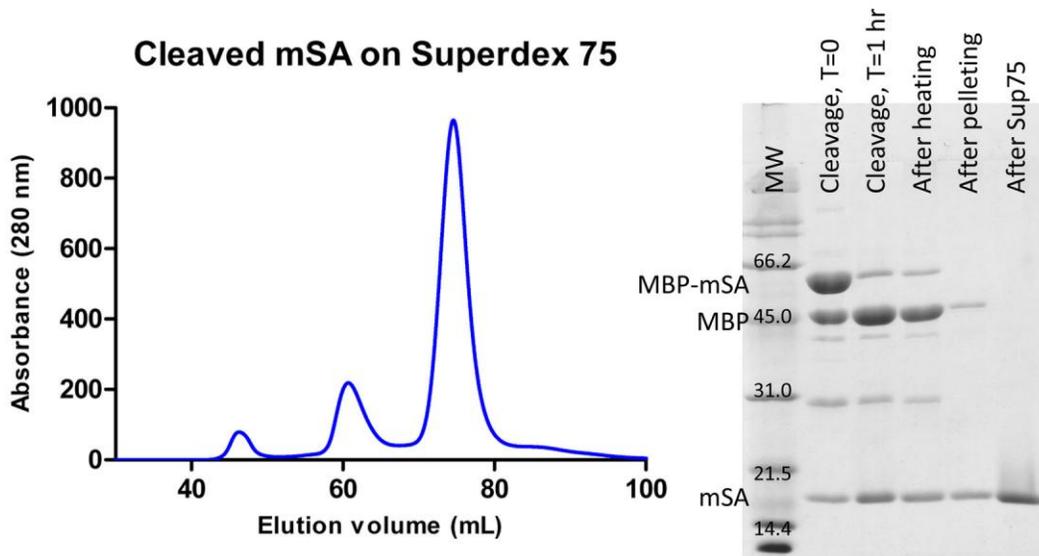


Figure 35 Cleavage of mSA and gel filtration over Superdex 75 column

mSA was cleaved, heated, and sample was run over Superdex 75 column to remove impurities. SDS-PAGE samples are as follows: sample immediately following addition of TEV protease (Cleavage, T=0), sample 1 hour after adding TEV protease (Cleavage, T=1 hr), sample after heating for 10 minutes at 75°C (After heating), sample after centrifuging at 10K x g for 10 minutes (After pelleting), and mSA purified over Superdex 75 column (After Sup75).

5.2.2 Enzymatic biotinylation of MBPavi using BirA and purification of biotinylated MBPavi

Since BirA activity is inhibited by NaCl, MBPavi is exchanged into 25 mM Tris pH 8 prior to biotinylation. In advance, 10X Biomix A (100 mM Tris pH 8), 10X Biomix B (100 mM adenosine triphosphate (ATP), 100 mM magnesium acetate, 500 μ M D-biotin) is prepared and stored at -80 °C. The biotinylation reaction contains 1X Biomix A, 1X Biomix B, 40 μ M MBPavi, and 500 μ g/mL MBP-BirA, and is carried out for 12-16 hours at 30 °C.

To remove excess biotin prior to purification of bMBPavi, the BirA/MBPavi mixture was buffer exchanged into avidin wash buffer (50 mM Tris pH 8, 150 mM NaCl) then loaded onto a column packed with 15 mL Monomeric Avidin UltraLink Resin

(Pierce, #53146) on an AKTA Pure FPLC instrument. The column was then washed with avidin wash buffer until the A_{280} was near baseline then washed with 5 column volumes (CV) avidin elution buffer (50 mM Tris pH 8, 150 mM NaCl, 2 mM biotin), followed by 5 CV avidin regeneration buffer (50 mM glycine pH 2.5). Peak elution fractions were pooled and concentrated in 10K Amicon Ultra centrifugal filter.

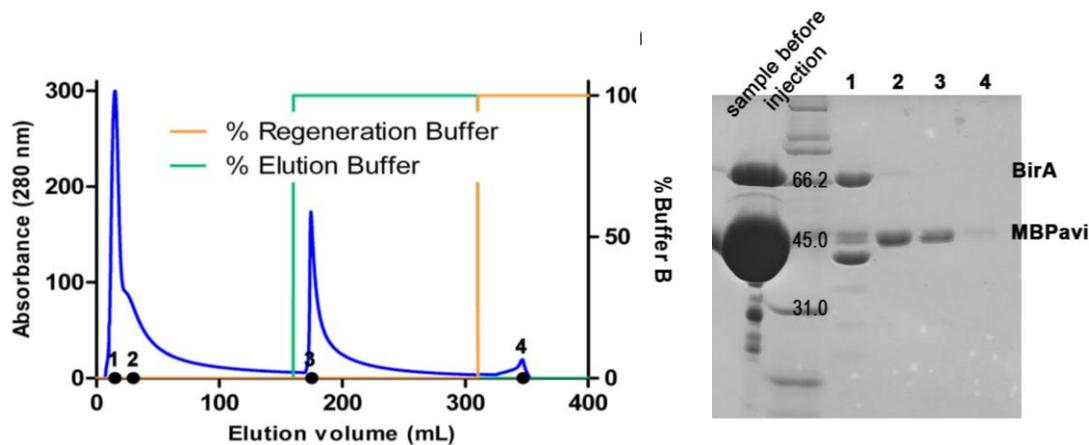


Figure 36 Purification of biotinylated MBPavi using avidin resin

MBPavi is biotinylated using BirA and the BirA/MBPavi mixture is purified over monomeric avidin resin. Fraction locations are marked with black dots along the x-axis of the purification trace. Elution buffer contains biotin to elute the biotinylated protein, and the regeneration buffer contains low pH for resin regeneration.

5.2.3 Complexation of bMBPavi with mSA

bMBPavi was exchanged into either 25 mM Tris pH 8 or 25 mM Tris pH 8, 150 mM NaCl TBS to remove excess biotin. Protein concentration was assessed using Beer's law where absorbance was measured at 280 nm using a Nanodrop 2000 UV/Vis spectrophotometer, and molecular weight, and extinction coefficients calculated by ExPASy ProtParam (149) based on protein sequences. Protein purity was assessed by 12% SDS-PAGE. bMBPavi and mSA were incubated on ice in a 4:1 mSA2:bMBPavi molar ratio for 1 hour prior to injection onto a Superdex 75 16/60 gel filtration column

using either 25 mM Tris pH 8 or TBS as the mobile phase. Samples were evaluated on SDS-PAGE and fractions containing both bMBP and mSA were concentrated in a 3K Amicon Ultra centrifugal filter for crystallization trials.

5.2.4 Crystallization trials of bMBP:mSA2 complex

After complex purification over gel filtration using either TBS or Tris pH 8 as the running buffer, fractions containing the complex were pooled and concentrated in a 3K MWCO Amicon Ultra centrifugal filter. Crystal trays were prepared using commercial 96 well screens (Hampton, Emerald Biosystems) and an Art Robbins Gryphon drop setter. Starting protein concentration ranged from 5 mg/mL to 21.9 mg/mL, and concentration was measured using Nanodrop 2000 with the sum of molecular weights and extinction coefficients as calculated by ExPASy ProtParam (149). The crystallization drop contained 0.3 μ L protein sample and 0.3 μ L crystallization solution from a commercial 96-well screen (Hampton or Emerald Biosystems). All trays were stored in a 20 °C incubator. Drops were imaged using a Rigaku Minstrel DT and visualized with both visible and UV light settings.

5.3 RESULTS

5.3.1 Insertion of the Avi-tag into MBP

MBP was chosen as the test protein because it is easily crystallized (a BLASTp (144) search of the *E. coli* MBP sequence returns >100 structures, see (245-247) for recent examples), easily expressed in *E. coli* as evidenced by the use of MBP as a fusion protein to ease expression of difficult proteins in *E. coli* (248), and already available in several plasmids in the lab. The insertion location for the Avi-tag was chosen based on available loops within MBP, which are limited (see **Figure 33** for alignment of MBP and

MBPavi, and **Figure 37** for Avi-tag placement in MBP). After insertion, MBPavi was still easily expressed, and the biotinylation was successful, so other Avi-tag insertion locations were not explored.

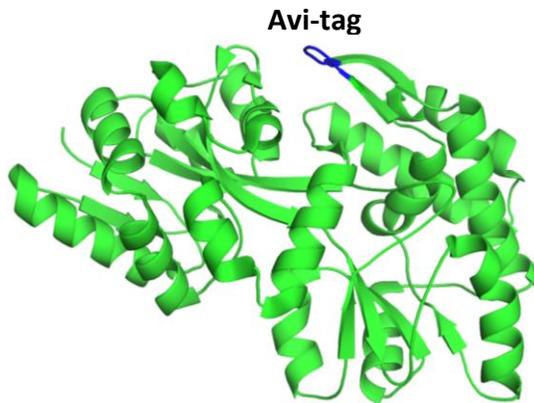


Figure 37 MBP structure showing placement of Avi-tag

Structure of MBP (green) was rendered in Pymol (1) from the MBP portion of PDB ID 3MQ9 (3). 6 residues were replaced by the Avi-tag and are colored blue. The other 9 residues were inserted in that location, making the loop 9 residues longer.

5.3.2 Purification of components

All components critical to the success of the project, namely, MBPavi, MBP-mSA2, and BirA were successfully expressed and purified in high yield and purity (**Figure 35**). BirA was used to enzymatically biotinylate MBPavi (**Figure 36**), and MBP-mSA2 was cleaved to free mSA2 (**Figure 35**). bMBPavi and mSA2 were complexed over gel filtration (**Figure 38**), and the isolated complex was used in crystallization trials.

Though each protein was obtained with high purity, some facets of the protocol could be optimized. First, the MBP-mSA2 cleavage reaction does not go to completion (see **Figure 35**, lane “cleavage, T=1 hr”), even after 24 hour room temperature incubation (data not shown). The cause of this is unknown, and the amount of protein lost is considered negligible. The sample is then heated to precipitate TEV protease, uncleaved MBP-mSA, and cleaved MBP. Heating the sample also causes protein loss (compare

mSA bands of **Figure 35**, lane “cleavage, T=1 hr” and lane “After pelleting”), but the sample is difficult to purify without heating due to the overlapping peaks of MBP and mSA during gel filtration.

The avidin purification of bMBPavi could also be optimized. The avidin column flowthrough should contain a uniform peak of BirA and unbiotinylated MBPavi, but **Figure 36**, lane 1 shows mostly BirA and the shoulder of the flowthrough peak, **Figure 36**, lane 2 contains MBPavi. The likely cause of this is that the capacity of the avidin resin has been reached so the rest of the bMBPavi cannot bind the column. This issue could be rectified easily by increasing the volume of monomeric avidin resin or separating the sample into two purification runs.

5.3.3 Complexation of biotinylated MBPavi with monomeric streptavidin

Because the lack of NaCl in a protein crystallization buffer allows the salt in the crystallization condition to have more impact, the bMBP:mSA2 gel filtration was done in both Tris pH 8 (without salt) and TBS. In both cases, the first small peak contained a small amount of aggregated mSA, the second peak contained bMBP:mSA2 complex, the shoulder of the second peak contained uncomplexed bMBP, and the last peak contained mSA (**Figure 38**). The elution volumes of the complex and the individual proteins were significantly different when Tris pH 8 was used instead of TBS pH 8. This is likely due to non-specific interactions of the protein with the gel filtration matrix in the condition without NaCl.

In the size exclusion chromatography of the complex, the sample contains a 4:1 ratio of mSA:bMBP. An excess of mSA2 was originally chosen so that no bMBP peak would be present, eliminating the overlap of the complex with free bMBP. Some portion

of the mSA2, however, is not able to bind free bMBP so some percentage of the mSA still has bound biotin (**Figure 38**). Biotin can be dissociated from streptavidin by low pH or elevated temperature, so a buffer exchange of mSA into 50 mM glycine pH 2.5 at 4 °C or 25 mM Tris pH 8 at 40 °C should eliminate much of the bound biotin and increase the yield of bMBP:mSA2 complex.

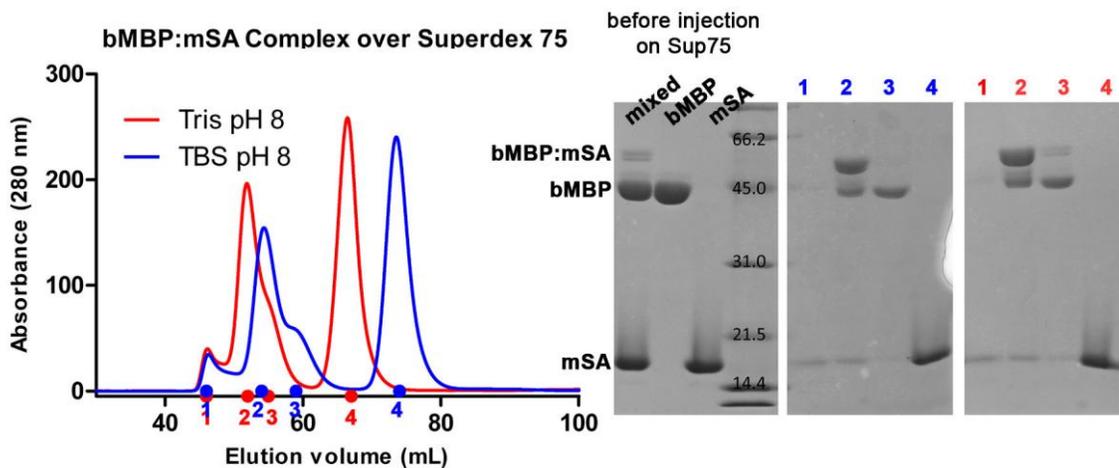


Figure 38 Gel filtration and SDS-PAGE of the bMBP:mSA complex

(a) purified mSA and bMBP were mixed and incubated on ice for 1 hour prior to injection on a Superdex 75 column. The mobile phase was either 25 mM Tris pH 8 (red) or 25 mM Tris pH 8, 200 mM NaCl (blue). (b) SDS-PAGE of samples prior to injection on the GF column (left) and fractions from each GF run (middle and right). Fraction locations are marked with either red or blue dots on the x-axis of the GF trace. The middle gel represents samples from the GF run in TBS and the right gel represents samples from the GF run in Tris pH 8.

5.3.4 Progress in the crystallization of bMBP:mSA2 complex

Crystallization trays of bMBP:mSA complex were prepared using both Tris pH 8 and TBS as the protein crystallization buffer (12 x 96 conditions from commercial screens and 3 x 24 conditions from homemade screens, ~1200 conditions total). As a general rule of thumb, about 50% of wells in a sparse matrix tray should contain protein precipitate immediately after the tray is prepared. The optimal concentration of protein in

a crystallization buffer without NaCl was much lower (6.7 mg/mL) than that of protein in TBS crystallization buffer (>22 mg/mL).

Crystallization trays of MBP_{avi} alone were also prepared to test if the Avi-tag hinders crystallization. One well did contain crystals, though they diffracted poorly and the structure was not determined (data not shown).

5.4 DISCUSSION

Though separately MBP and streptavidin are soluble, easily crystallized proteins, crystallization trials of the bMBP:mSA2 complex were unsuccessful. As noted with the insertion of the EE tag in Chapter 4, the insertion of the Avi-tag likely increases the flexibility of the loop which negatively impacts the crystallization. After several crystallization trials using different commercial screens and a range of protein concentrations, no complex crystals were identified.

Though the crystallization of bMBP_{avi}:mSA2 was unsuccessful here, as indicated in Chapter 4, we are still optimistic that this approach could be successful. A BLASTp (144) search of the 15 amino acid Avi-tag reveals 3 structures with the Avi-tag modeled into the published structure. In PDB ID 2BCK, amino acids GLNDIF were modeled into the structure (249). Similarly, in PDB ID 2YF6, GL was modeled (250). In PDB ID 4L79, all 15 amino acids of the Avi-tag were modeled, the first 9 of which were in an alpha helix and the last 6 were in a loop region (251). In all 3 structures, the modeled Avi-tag forms crystal contacts with neighboring molecules, perhaps locking the disordered structure into a specific confirmation (**Figure 39**). Without stabilizing crystal contacts, the tag will likely be flexible and therefore difficult to crystallize.

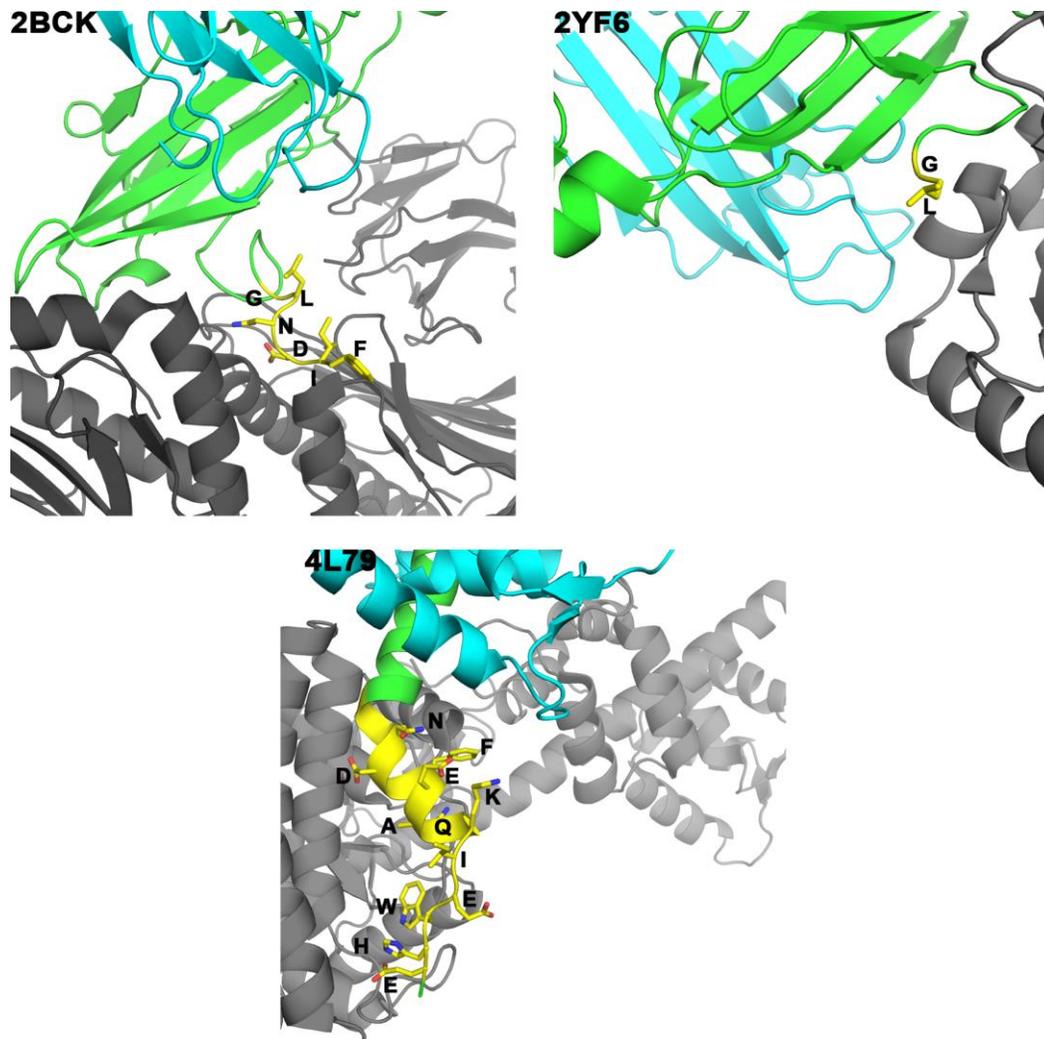


Figure 39 Structures of 3 partial or full Avi-tags from the PDB

PDB ID of the structure is shown in the upper left corner of each image. The asymmetric unit is colored by chain (blue and green) with the Avi-tag in yellow sticks with the amino acid labelled in black. Symmetry related molecules forming crystal contacts with the Avi-tag are shown in gray.

5.4.1 Concluding remarks and future directions

With the available PDB structures indicating that the Avi-tag is flexible, another approach may be needed in order to use monomeric streptavidin as a crystallization chaperone without the negative effects of the flexibility of the 15 amino acid Avi-tag inserted into the protein of interest. A 13 amino acid miniprotein (RCCHPQCGAVEECCR) was designed by the lab of Dr. Robert O. Fox of the University

of Texas Medical Branch as a ligand for tetrameric streptavidin in place of biotin. Though the work was never published, the crystal structure in the PDB (PDB ID 1HQQ (2)) reveals that the miniprotein has a rigid structure with all the cysteines forming disulfide bonds, and the N- and C-termini are within 10 Å of one another. The miniprotein streptavidin ligand could theoretically be inserted into a loop of a protein of interest and the rigid structure of the miniprotein tag could allow for crystallization of the miniprotein-tagged:mSA2 complex. As shown here, our Avi-tag approach may have added conformational heterogeneity in the crystallization drop, but perhaps a similar, more rigid, approach could prove effective.

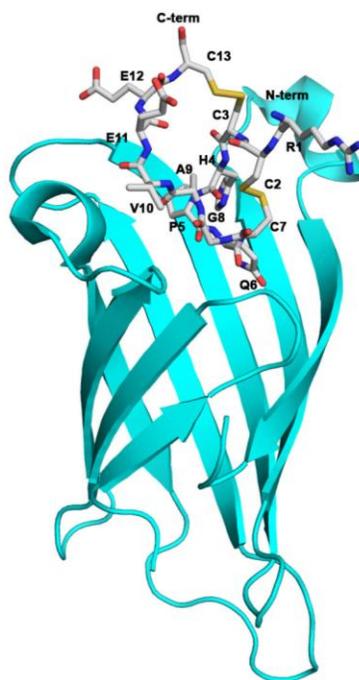


Figure 40 Streptavidin in complex with 13 amino acid miniprotein ligand

A monomer of tetrameric streptavidin is shown in cyan cartoon, with the other three subunits omitted for clarity. The 13 amino acid miniprotein ligand is shown as white sticks, with each side chain labelled. N- and C-term are also labelled, and are 9.8 Å away from one another. Image was rendered in Pymol from PDB ID 1HQQ (2)

CHAPTER 6: MOLECULAR BASIS OF NOVEL 1- DEOXYGALACTONOJIRIMYCIN ARYLTHIOUREA PHARMACOLOGICAL CHAPERONES BINDING TO ALPHA GALACTOSIDASE A FOR THE TREATMENT OF FABRY DISEASE

6.1 INTRODUCTION

Fabry disease (FD) is an X-linked lysosomal storage disorder (LSD) resulting from mutations in the *GLA* gene which encodes glycoside hydrolase α -galactosidase A (α -Gal A) (252). Mutations in α -Gal A result in the aggressive accumulation of glycosphingolipids in the lysosome, mainly globotriaosylceramide (Gb3). Gb3 accumulation could result in minor to severe symptoms, including neuropathic pain, gastrointestinal issues, hearing loss, cardiomyopathy, skin angiokeratomas, and renal problems. FD incidence is estimated to be around 1:100,000 live births, though some believe that is an underestimation (252), and incidence in Italian and Taiwan male newborns is much higher (1:3100 and 1:1250, respectively) (253, 254).

The three-dimensional crystal structure of α -Gal A shows two domains in each monomer. Residues 32-328 make up the N-terminal domain, and residues 329-421 fold into a C-terminal domain containing eight antiparallel β -strands packed into a β sandwich. The active site is contained within the N-terminal domain, and 15 residues come together to form the pocket (Trp 47, Asp 92, Asp 93, Tyr 134, Cys 142, Lys 168, Asp 170, Cys 172, Glu 203, Leu 206, Tyr 207, Arg 227, Asp 231, Asp 266, and Met 267) (255). 13 of those residues made direct hydrogen bonding interactions with galactose (256, 257). The catalytic nucleophile is Asp 170 and the catalytic acid/base is Asp 231 (257). More than 585 mutations in α -Gal A have been found to cause some form of FD

(Human Gene Mutation Database, www.hgmd.cf.ac.uk), including all 13 residues interacting with galactose. About 60% of all known mutations result in unstable α -Gal which is targeted for degradation in the ER.

In 2001, enzyme replacement therapy (ERT) was introduced for the treatment of FD where purified enzyme is injected into the patient biweekly (252, 258, 259). Drug efficacy has been confirmed in FD patients, but problems include stability of the protein in blood, and patient immunological response (260). Alternatively, pharmacological chaperones (PCs, the use of “chaperone” here is unrelated to its use in previous chapters) are being tested for FD. PCs are small molecules which bind to the destabilized enzyme in the ER, allowing proper folding and trafficking to the lysosome. Once inside the lysosome, the PC dissociates from α -Gal A, and its glycoside hydrolase activity is restored.

Competitive inhibitors act as PCs for α -Gal A. A mimic of the terminal galactose unit of Gb3, deoxygalactonojirimycin (DGJ), has been shown to enhance the residual activity of α -Gal A by stabilizing the mutant enzyme and allowing trafficking to the lysosome (261, 262). DGJ is a low molecular weight molecule, can be administered orally, is free from immunological reactions, and is currently in clinical trials for the treatment of FD (252, 263). However, DGJ is largely protonated at physiological pH, and is highly hydrophilic, making it difficult to pass through biological membranes. Previously, C1- and N-substitutions to increase amphiphilicity has been done, with limited success due to a significant decrease in the affinity of α -Gal A for the chaperone (264). A look at the DGJ-bound α -Gal A crystal structure reveals a hydrogen bonding interaction between Asp 170 and the protonated amino group of DGJ (257, 265), which is

likely broken with *N*-substitutions (**Figure 41a**). In the case of galactose, the interaction is stabilized by Asp 231, which is not possible for a C1 carbon substituent (**Figure 41b**). We hypothesize that the Asp 231 interaction could be maximized with a DGJ-arylthiourea, with the N'H proton being the hydrogen bond donor (**Figure 41c**).

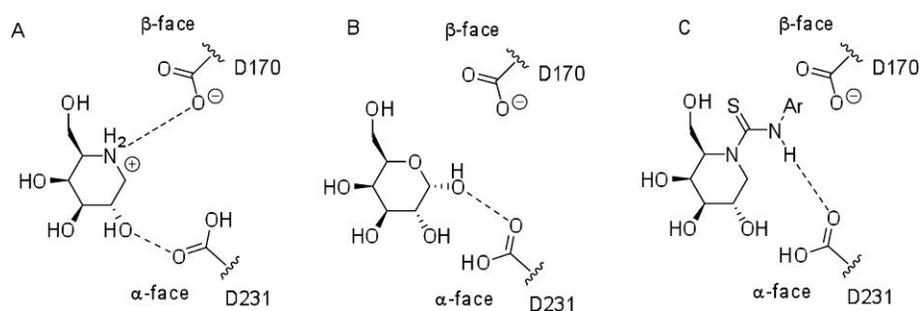


Figure 41 Design criteria for the novel FD PCs, DGJ-ArTs

Key hydrogen bonding interactions between the catalytic residues D170 and D231 with DJG (a), galactose (b), and the novel DGJ-ArTs (c).

In this chapter, the synthesis of 4 DGJ-ArTs is discussed, as well as the molecular basis for DGJ-ArT binding based on an X-ray crystal structure of DGJ-ArT-bound α -Gal A. Inhibition and activity assays are done to evaluate the potential of DGJ-ArTs for the treatment of FD.

6.1.1 Individual contributions to the work

This work is highly collaborative between the laboratories of Katsumi Higaki at Tottori University in Yonago, Japan, José García Fernández at Universidad de Sevilla in Sevilla, Spain, and Raquel Lieberman at Georgia Tech. Yi Yu and Katsumi Higaki performed inhibition and stabilization assays in vitro. Yu Yi, Katsumi Higaki and Takahiro Tsukimura performed chaperone test and cell toxicity test on cultured fibroblasts. Yi Yu, Katsumi Higaki, and Naoe Nakasone performed transfection/chaperone test on COS7 cells, immunofluorescence analysis of anti-Gb3 and

western blot analyses. Katsumi Higaki, Eiji Nanba and Kousaku Ohno designed the overall biological study. Hitoshi Sakuraba and Yoshiyuki Suzuki supervised the work. Dr. Jason Drury crystallized the protein, performed crystal soaking experiments, and collected the x-ray data. I did all the data processing and model building, including modeling the inhibitor into the active site and analysis of hydrogen bonding interactions. Members of the García Fernández Lab performed remaining experiments.

6.1.2 Publication resulting from this work

This research was published in the journal ACS Chemical Biology issue 9 in 2014 (266). This chapter was adapted with permission from Yu *et al.*, ACS Chemical Biology, 2014. **9**(7): p. 1460-1469. <http://pubs.acs.org/doi/abs/10.1021/cb500143h> Copyright (2014) American Chemical Society.

6.2 METHODS

6.2.1 Synthesis of DGJ-arylthioureas and other DGJ Derivatives

1-Deoxygalactonojirimycin (DGJ) was prepared following a protocol previously reported (267). The 4 novel DGJ-ArT pharmacological chaperones (*N'*-(1-naphthyl) (DGJ-NphT), *N'*-(*p*-methoxyphenyl) (DGJ-*p*MeOPhT), *N'*-(*p*-methylthiophenyl) (DGJ-*p*MeSPhT), and *N'*-(*p*-fluorophenyl) (DGJ-*p*FPhT), **Figure 42**) were synthesized by reacting the substituent component (1-naphthyl, *p*-methoxyphenyl, *p*-methulthiophenyl, or *p*-fluorophenyl isothiocyanate, respectively) with DGJ (**Figure 43**). Two additional DGJ derivatives were synthesized for comparative purposes (**Figure 42**). The bicyclic isothiourea-type derivitave 5*N*,6*S*-(*p*-fluorophenylimino methylidene)-6-thio-1-deoxynojirimycin (*p*FPhIM-DGJ) was prepared by HCl-promoted cyclization of

monocyclic thiourea precursor DGJ-*p*FPhT (**Figure 43**). DGJ *N'*-benzylthiourea (DGJ-BnT) was prepared by reaction of DGJ and benzyl isothiocyanate.

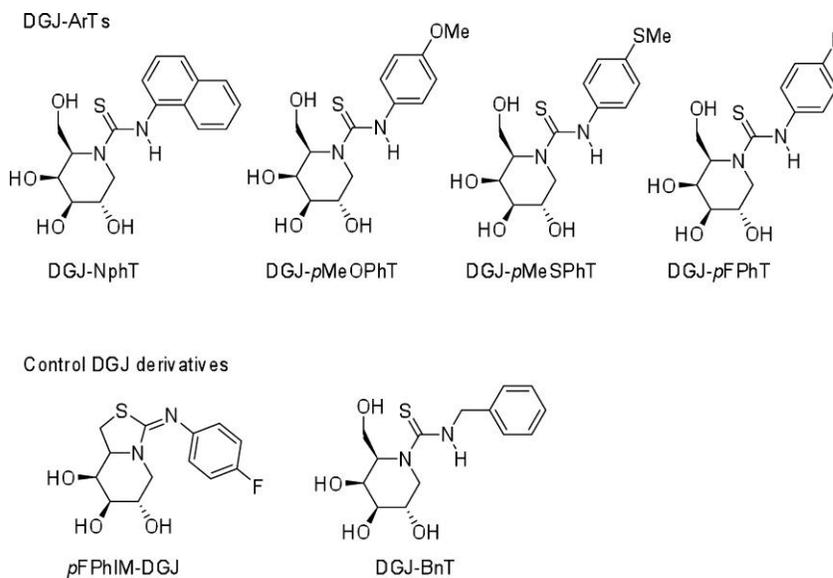


Figure 42 Chemical structures of DGJ-ArTs and DGJ derivatives

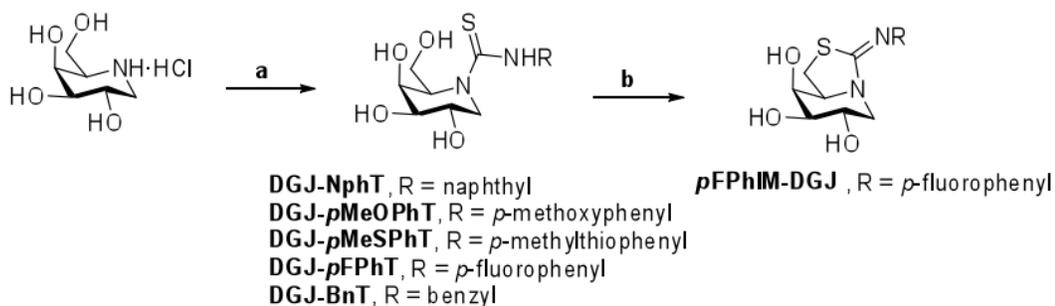


Figure 43 Synthesis of DGJ-ArT pharmacological chaperones and bicyclic DGJ derivative

Reaction conditions: a) RNCS, pyridine, diethylamine, room temperature for 18 hours, 47-94% yield; b) HCl, methanol, room temperature for 12 hours, 45-99% yield.

6.2.2 Cell culture, transfection, and pharmacological chaperone test

Human skin fibroblasts and COS7 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) with 10% FBS. Immortalized human fibroblast lines were

established by transfecting SV40 large T cDNA expression vector, pET321-T (268). To test the pharmacological chaperones, human fibroblasts were cultured in DMEM with or without the chaperone for 96 hours as described previously (269, 270). The chaperones were exposed to the immortalized human fibroblast cells for 48 hours. Lipofectamine 2000 was used for the transfection of COS7 cells with wild-type and mutant α -Gal A cDNA (271). After 5 hours, the medium was replaced with fresh medium containing or lacking the pharmacological chaperones and cells were incubated for 48 hours. Cytotoxicity of the compounds in the human fibroblast cells was evaluated using the lactate dehydrogenase (LDH) cytotoxicity assay.

6.2.3 Measurement of enzyme activities in the lysosome

4-MU-conjugated substrates were used to measure lysosomal enzyme activities (4-MU-conjugated α -D-galactopyranoside for α -Gal A, β -D-galactoside for β galactosidase (β -Gal), *N*-acetyl- β -D-glucosaminide for total hexosaminidase (Hex) and α -*N*-acetyl-D-galactosaminide for α -*N*-acetylgalactosaminidase (α -NAGA)). For the α -Gal assay, 10 μ L cell lysate (in 0.1% Triton X-100 in dH₂O) was mixed with 20 μ L 4-MU substrate (5 mM 4-MU α -D-galactopyranoside and 0.1 M *N*-acetyl-D-galactosamine in 0.1 M citrate buffer, pH 4.5). The mixture was incubated at 37 °C for 1 hour and terminated with the addition of 0.2 M glycine-NaOH, pH 10.7. The free 4-MU was measured with a Tecan Infinite F500 fluorescence plate reader using excitation wavelength 340 nm and emission wavelength 460 nm. Protein concentration was measured using Wako Protein Assay Rapid Kit, and enzyme activity was normalized by protein concentration.

6.2.4 Inhibition and stabilization of α -Gal A *in vitro*

0.1% Triton X-100 extracts from normal human skin fibroblasts were mixed with 4-MU substrates in absence or presence of DGJ derivatives (269) for the inhibition assay. For heat-induced degradation, extracts were incubated at 48 °C for the time indicated in 0.1 M citrate buffer, pH 7. The reaction was terminated with the addition of 0.1 M citrate buffer, pH 4.5 and the enzyme activity was measured as in section 6.2.3.

6.2.5 Crystallization and X-ray data collection of α -Gal A in complex with DGJ-*p*FPhT

Apo α -Gal A (10 mg/mL) was crystallized at pH 4.5 as described previously (256). 0.15 μ L of 1 mM DGJ-*p*FPhT encapsulated in β -cyclodextran (β CD) (272) was added directly to the crystallization drop and incubated for 4.5 hours at room temperature. Crystals were cryocooled in mother liquor supplemented with 30% ethylene glycol. Diffraction data were collected at SER-CAT beamline 22-ID at the APS in Argonne, Illinois. Data were indexed, integrated, and scaled using HKL-2000 (273) in space group $P3_221$. Molecular replacement using a single chain of apo α -Gal A (PDB ID 3GXN (256)) was performed in Phaser (200) and the model was refined using Phenix (204). eLBOW (274) within Phenix was used to generate the restraints of the DGJ-*p*FPhT molecule using coordinates generated in PRODRG (275). The final occupancy of DGJ-*p*FPhT within the α -Gal A structure is 0.6. The crystallographic statistics can be found in **Table 18** and the structure has been deposited in the PDB (PDB ID 4NXS).

Table 18 Data collection and refinement statistics for α -Gal A in complex with DGJ-*p*FPhT pharmacological chaperone

Parameter	A-Gal A in complex with DGJ- <i>p</i> FPhT
Data collection	
Beamline source	APS 22-ID
X-ray wavelength (Å)	1.0
Resolution (Å)	38.68 - 2.549 (2.64 - 2.549)
Space group	P 32 2 1
Unit-cell parameters	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	90.783 90.783 217.382
α , β , γ (°)	90 90 120
Total No. of reflections	180519
No. of unique reflections	33780 (2979)
Multiplicity	5.3 (3.6)
Completeness (%)	97.25 (87.34)
$\langle I/\sigma(I) \rangle$	15.14 (2.56)
R_{sym} (%)	0.105 (0.471)
Refinement statistics	
Final R_{cryst}	0.1871 (0.2628)
Final R_{free} ^a	0.2317 (0.3118)
No. of non-H atoms	
Protein	6261
ligands	179
water	96
R.m.s. deviations	
Bonds (Å)	0.007
Angles (°)	0.93
Average B factors (Å ²)	
Protein	51.5
Ligand	76.6
Water	45.3
Ramachandran plot ^b	
Most favoured (%)	97
Outliers (%)	0

^a R_{free} is calculated for a randomly chosen 5% of reflections which were not used for structure refinement

^bAs calculated by RAMPAGE (205)

6.2.6 Immunofluorescence staining

The cells were fixed to coverslips with 4% paraformaldehyde in PBS, then permeabilized with 0.1% Triton X-100 in PBS, and incubated with the primary antibodies for 1 hour. Bound primary antibodies were detected using secondary antibodies conjugated with Alexa-Fluor. Fluorescence images were collected using a Leica TSC SP-2 confocal laser microscope. All of the procedures were carried out at room temperature. Gb3 intensity was measured using Leica confocal software and intensity was normalized using 4',6-diamidino-2-phenylindole (DAPI) intensity. For each experiment, over 100 cells in 10 randomly collected images were evaluated.

6.2.7 Immunoblotting

Cultured cells were lysed by sonication in 10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, and a Roche protease inhibitor cocktail. Detergent-resistant membrane microdomains were obtained by centrifugation at 100,000 x g for 30 minutes in 1% Triton X-100. Procedure was carried out at 4 °C. Immunoblotting was performed as described previously (276) and signal from HRP-conjugated secondary antibodies were visualized using enhanced chemiluminescence (ECL) detection kit (GE Lifescience). Images were obtained using Fujifilm LAS-4000 lumino image analyzer.

6.3 RESULTS

6.3.1 Comparison of α -Gal A crystal structures in complex with DGJ and DGJ-*p*FPhT

The 2.55 Å resolution crystal structure of α -Gal A in complex with DGJ-*p*FPhT shows the glycomimetic in the active site as expected. The 2Fo-Fc density is clear for the iminosugar and the fluorobenzyl ring, but no 2Fo-Fc density is present for the thiourea

linker (**Figure 44a**). The orientation of the iminosugar and side chains involved in hydrogen bonding are similar to that of DGJ in PDB ID 3GXT (**Figure 44b**), and the polypeptide chain is largely unchanged from other ligand-bound α -Gal A structures, as seen previously (256). Hydroxyl groups OH-2, OH-3, OH-4, and OH-6 of DGJ-*p*FPhT are involved in extensive hydrogen bonding network with α -Gal A residues Asp 92, Asp 93, Lys 168, Arg 227, and the catalytic nucleophile Asp 170, and the catalytic acid/base Asp 231 (**Figure 44c**). Similar hydrogen bonding pattern is seen with DGJ, with the exception of Asp 170 and Asp 231, as predicted (**Figure 44d**). Hydrogen bonding distances for DGJ-*p*FPhT are slightly longer than hydrogen bonds for DGJ and the fluorobenzyl ring is nestled among hydrophobic residues Leu 206, Ala 230, and Tyr 207. The cocrystal structure confirms that the desired orientation of the pharmacological chaperone within the α -Gal A active site was achieved, with the predicted hydrogen bonding interaction between Asp 231 and the N'H proton.

6.3.2 Effects of DGJ arylthiourea compounds on α -Gal A *in vitro*

The inhibitory activity of the DGJ-ArTs was tested in human skin fibroblast cells and compared to the activity of DGJ and the two DGJ derivatives, bicyclic isothiourea *p*FPhIM-DGJ and alkylthiourea DGJ-BnT (**Figure 42**, **Figure 45a**). IC₅₀ values were calculated from the inhibition assay and range from 0.0083 to 1.6 μ M for the 4 DGJ-ArTs at pH 5 (**Table 19**). The inhibitory effect of the two control DGJ derivatives, *p*FPhIM-DGJ and DGJ-BnT, are 3- to 4-fold less than the inhibitory potential of the 4 designed DGJ-ArTs. The potency of the DGJ-ArT inhibitors increases from pH 7 to pH 5 by 4.3 to 7.9 fold. Inhibition was also tested for β -Gal, α -NAGA, and Hex to test for selectivity.

No inhibition was observed for β -Gal and Hex, and moderate inhibition was detected for α -NAGA (**Figure 45b**).

The ability of the DGJ-ArTs to protect the heat-induced inactivation of α -Gal A was measured in normal fibroblasts. After incubating cell extracts for 40 minutes at 48 °C, α -Gal A retained about 20% of its initial activity. The 4 DGJ-ArTs were able to protect against the heat inactivation, with the most effective compounds being DGJ-*p*MeOPhT and DGJ-*p*MeSPhT, with 95 and 85% protection after 1 hour incubation and 30 μ M concentration. The designed DGJ-ArTs were significantly more effective than reference pPhIM-DGJ, and slightly more effective than DGJ-BnT (**Figure 46**).

6.3.3 The activity of α -Gal A is increased in normal and FD fibroblast cells after treatment with DGJ-ArTs

Normal and FD (Q279E mutant) fibroblasts were cultured in the presence and absence of each DGJ-ArT and reference DGJ derivatives at varying concentrations. After 96 hours, α -Gal A activity was determined (**Figure 47a-b**). 3 of the 4 designed DGJ-ArTs showed enhanced α -Gal A activity in the normal fibroblasts. Only DGJ-*p*MeSPhT did not give enhanced activity, with the other three designed DGJ-ArTs giving 1.2 to 1.5 fold increase (**Figure 47a**). In FD cells expressing mutant (Q279E) α -Gal A, all 4 designed DGJ-ArTs enhanced the α -Gal A activity. The effect of DGJ-*p*MeOPhT was the greatest, with a 3-fold activity increase over cell treated with DGJ and more than a 7-fold increase as compared to untreated cells (**Figure 47b**).

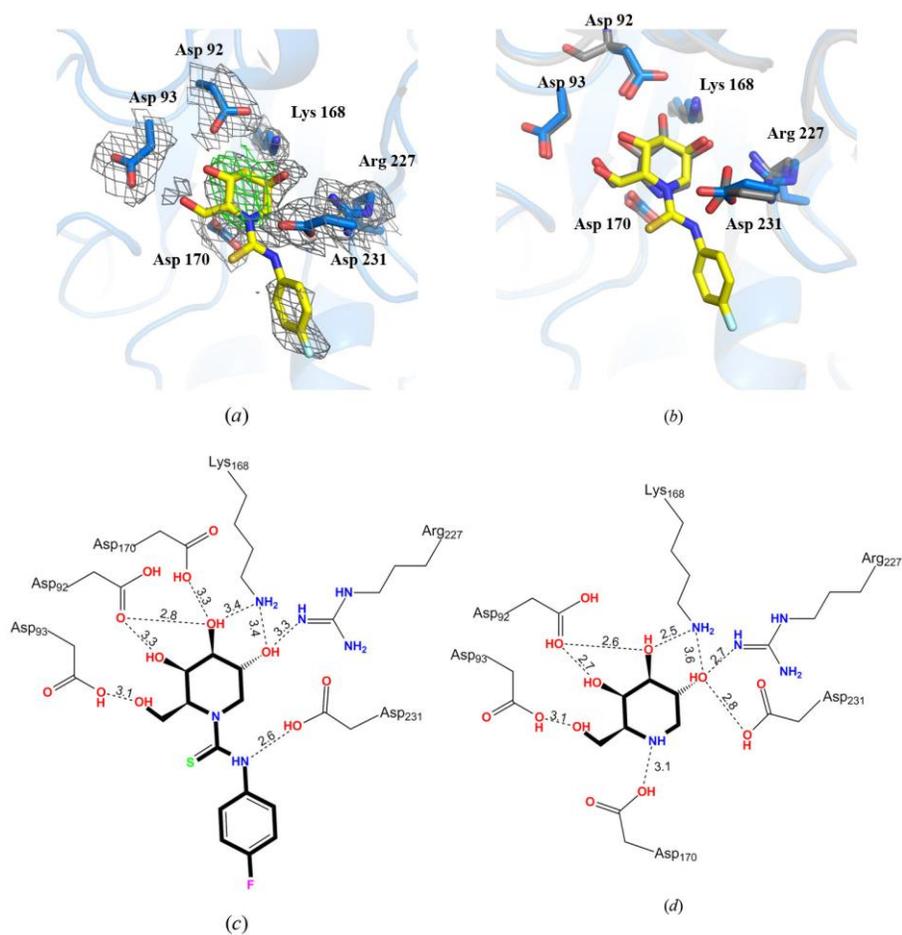


Figure 44 Crystal structure of α -Gal A in complex with DGJ-pFPhT

(a) 2.55 Å resolution structure showing final 2Fo-Fc density at 1 σ in gray mesh for the DGJ-pFPhT (yellow) and interacting residues of α -Gal A (blue). Green mesh is Fo-Fc density contoured at 3s (σ) after initial molecular replacement with apo α -Gal A from PDB ID 3GXN and first round of refinement, prior to modelling the chaperone. (b) Overlay of DGJ-bound structure (gray) and DGJ-pFPhT-bound structure (blue and yellow). (c) Hydrogen bonding network of α -Gal A and DGJ-pFPhT. Distances shown are average distances from two molecules in the asymmetric unit. (d) Hydrogen bonding network involving DGJ in the active site of α -Gal A from PDB ID 3GXT. Distances are in angstroms.

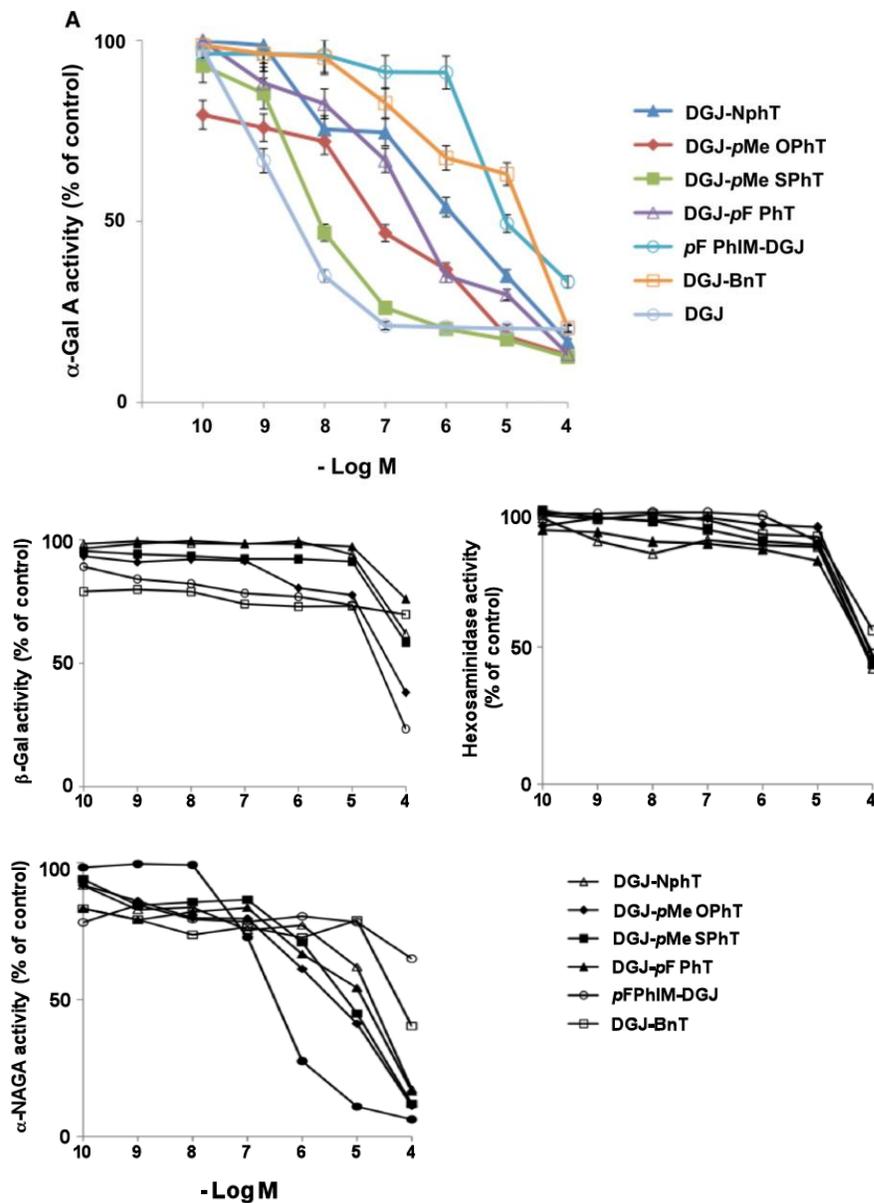


Figure 45 Effects of DGJ-ArTs and derivatives on human normal fibroblast cells in vitro

(a) Inhibition assay of DGJ-ArTs on α -Gal A *in vitro*. Enzyme activity was measured in cell lysate in the presence or absence of increasing concentrations of DGJ or DGJ derivative compounds. Each data point represents the mean \pm the standard error of the mean (SEM) of three determinations, each done in triplicate. (b) Effects of 4 DGJ-ArTs and DGJ derivatives on β -Gal, Hex, and α -NAGA. Assay was done as noted in a.

Table 19 IC₅₀ values of DGJ derivative compounds against Gal A in vitro

Activity was determined as described in **Figure 45**.

	pH 5 (μM)	pH 7 (μM)	Fold increase from pH 7 to pH 5
DGJ-NphT	1.60 ± 0.092	0.37 ± 0.045	4.3
DGJ- <i>p</i> MeOPhT	0.074 ± 0.0064	0.016 ± 0.0020	4.6
DGJ- <i>p</i> MeSPhT	0.0083 ± 0.00051	0.0014 ± 0.00031	5.9
DGJ- <i>p</i> FPhT	0.34 ± 0.047	0.043 ± 0.0062	7.9
<i>p</i> FPhIM-DGJ	9.67 ± 0.88	2.81 ± 0.39	3.44
DGJ-BnT	20.26 ± 2.43	4.41 ± 0.65	4.59
DGJ	0.0030 ± 0.00048	0.00078 ± 0.00013	6.25

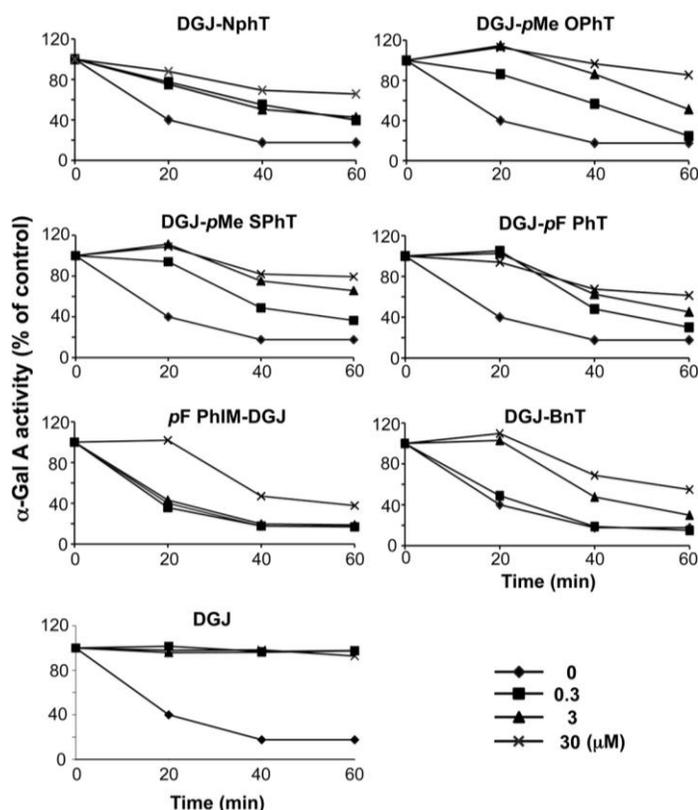


Figure 46 Stabilization activity of DGJ and DGJ derivatives on α-Gal A in skin fibroblast cells.

Cell lysate was incubated at pH 7.0 and 48°C for the indicated time and the α-Gal A activity was measured. Each point represents the mean of triplicates obtained from three separate experiments. Values are expressed as a percentage of the α-Gal A activity in absence of compound.

Cell toxicity of the pharmacological chaperones was measured by LDH assay and only DGJ-*p*MeSPhT was found to be cytotoxic (**Figure 48a**). DGJ-*p*MeOPhT was selected for immunofluorescence study, using 30 μ M concentrations. Treatment of Q279E FD fibroblasts with DGJ-*p*MeOPhT significantly reduced Gb3 accumulation, meaning the mutant enzyme is actively processing the substrate in the lysosome.

6.3.4 The effect of DGJ-*p*MeOPhT on mutants of α -Gal A

α -Gal A (WT and 17 missense mutants) were transiently transfected into COS7 cells. DGJ-*p*MeOPhT significantly enhanced the activity of 15 out of 17 of the mutants, with the exception of E66Q and G373D (**Figure 48b**). The results for the other 3 DGJ-*p*MeOPhTs were similar (data not shown).

6.3.5 Autophagic proteins are upregulated in FD cells and restored by DGJ-*p*MeOPhT and DGJ-FPhT

SV-40 mediated transformed cell lines from normal and Q279E FD fibroblasts (FD-SV) were established for these experiments. DGJ-*p*MeOPhT and DGJ-*p*FPhT, increased α -Gal A activity in FD-SV, and had significant effect on activity over DGJ at 30 μ M concentration (**Figure 49a**).

Autophagy-related proteins within were then analyzed using FD-SV cells. LC3-II, a specific autophagosome marker, was significantly increased in FD-SV cells over control cells, and a small increase in p62 was seen in FD-SV cells (**Figure 49b**). No difference in levels of beclin-1 and Bip was detected. FD-SV cells treated with chaperones DGJ-*p*MeOPhT and DGJ-*p*FPhT showed marked decrease in the levels of LC3-II and p62 (**Figure 49c**).

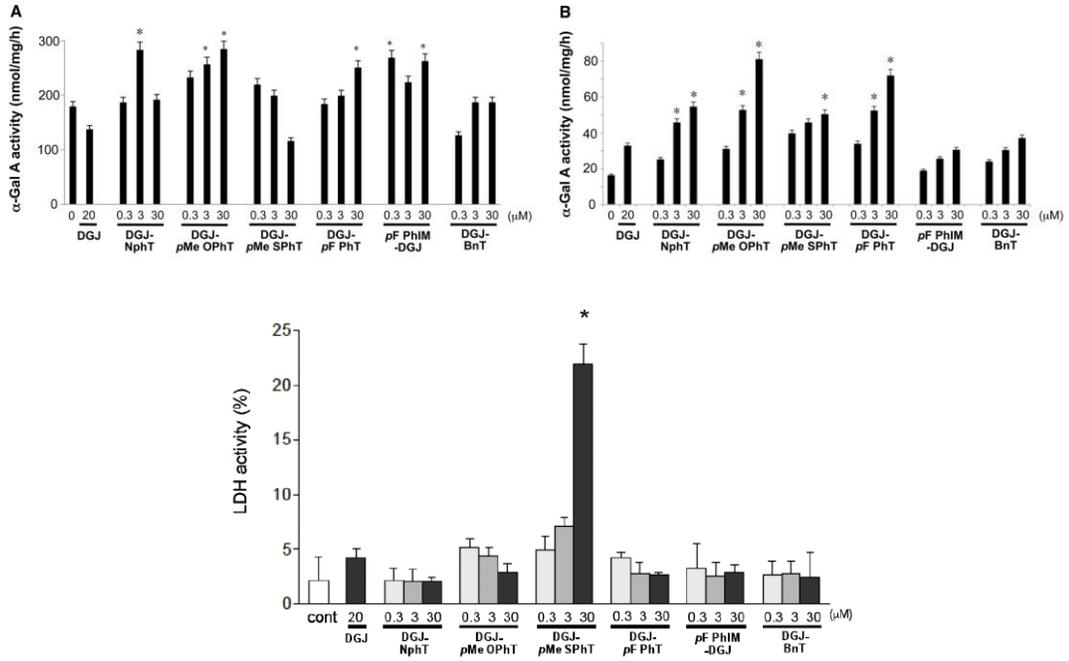


Figure 47 Gal-A activity is enhanced by DGJ-ArTs in normal and FD fibroblasts

(a) Activity of 4 DGJ-ArTs, 2 DGJ derivatives, and DGJ on normal fibroblasts. (b) Activity of DGJ and derivatives on FD fibroblasts, measured in nmol/mg/hour. For a, b, and d, each bar represents the mean \pm SEM of three experiments, each done in triplicate. (c) Cell toxicity assay measuring LDH levels on normal skin fibroblasts for each of the DGJ derivatives.

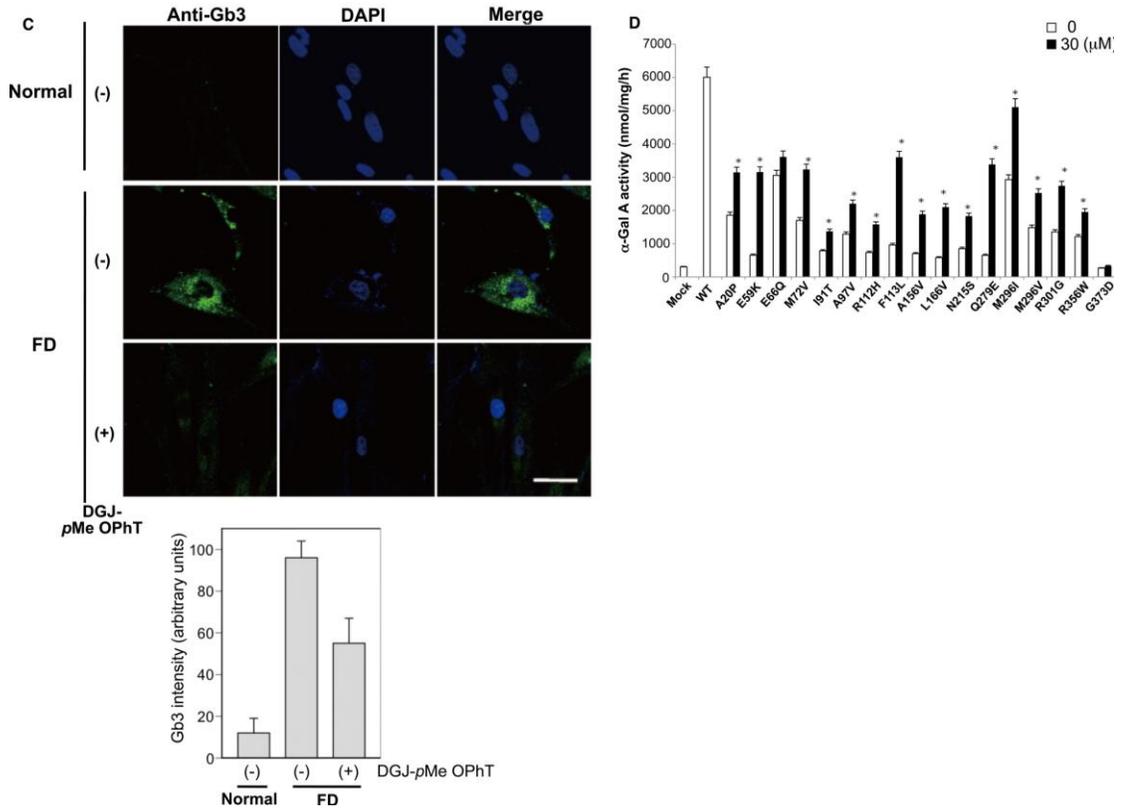


Figure 48 Activity of DGJ-pMeOPhT on normal and FD fibroblasts, and COS7 cells expressing WT and mutant α -Gal A.

(a) immunofluorescence of anti-Gb3 on normal and FD fibroblasts in the presence and absence of DGJ-pMeOPhT with quantification of Gb3 intensity. (b) Activity of DGJ-pMeOPhT on COS7 cells with WT and mutant α -Gal expression. Each bar represents the mean \pm SEM of three experiments, each done in triplicate.

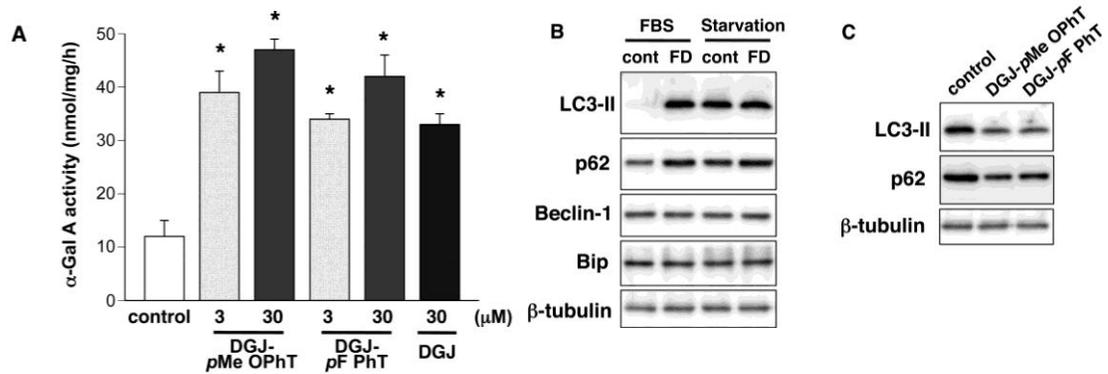


Figure 49 Effects of DGJ-pMeOPhT and DGJ-pFPhT in the impairment of autophagy of FD fibroblasts

(a) α -Gal A activity of transformed FD (SV-FD) fibroblasts in the presence and absence of DGJ and two DGJ-ArTs. Each bar represents the mean \pm SEM of three experiments, each done in triplicate. * $p < 0.05$, statistically different from the activity of untreated samples. (b) Western blot analysis of proteins related to autophagy from normal and FD fibroblasts in FBS and in serum starvation medium. (c) Effect of DGJ-pMeOPhT and DGJ-pFPhT on levels of LC3-II and p62 in FD fibroblast cell lysate. B-tubulin was used for control.

6.3.6 Effects of proteostasis regulators on α -Gal A activity are additive

Two proteostasis regulators, 4-PBA and celastrol, were added with DGJ-pMeOPhT to test the added effect of mutant α -Gal A activity. Without the addition of DGJ-pMeOPhT, 4-PBA had a statistically significant effect on mutant α -Gal A activity. Notably, 0.1 mM 4-PBA in the presence of 20 μ M DGJ-pMeOPhT had a very significant, additive effect on the activity. Celastrol had no effect on α -Gal A activity (**Figure 50**).

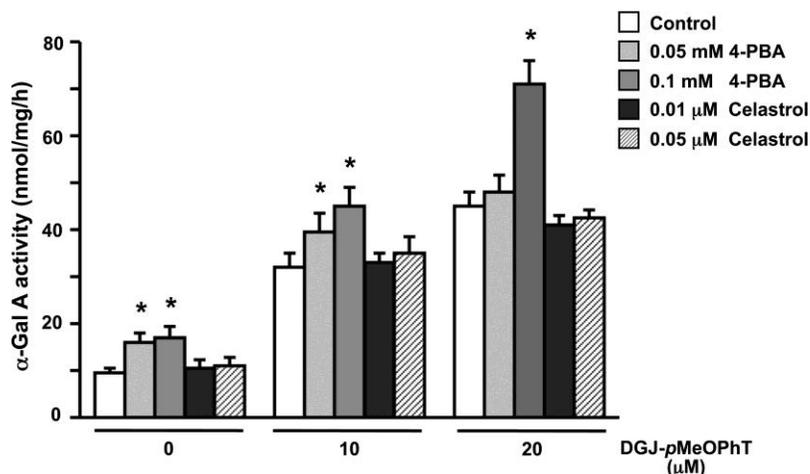


Figure 50 Effects of DGJ-pMeOPhT with proteostasis regulators on α -Gal A activity

FD-SV cells were treated with 4-PBA and celastrol in the presence or absence of DGJ-pMeOPhT. Each bar represents the mean \pm SEM of three experiments, each done in triplicate. * $p < 0.05$, statistically different from the activity of α -Gal A in the absence of 4-PBA and celastrol controls.

6.4 DISCUSSION

Currently, DGJ is in clinical trials for the treatment of Fabry disease, but DGJ has some drawbacks, mainly its hydrophilicity which limits diffusion through cell membranes. Because of these issues, a new family of pharmacological chaperones for Fabry disease, DGJ-ArTs, were designed to have strong hydrogen bonding capabilities between Asp 231 and the N^oH proton. In this study, the 4 designed DGJ-ArTs were synthesized and tested *in vitro*.

The molecular basis for DGJ-ArT binding to α -Gal A was elucidated by X-ray crystallography structure, and confirmed the hypothesis that the N^oH proton of the arylthiourea would make a strong hydrogen bond donor for the catalytic residue, Asp 231. Asp 170, while interacting with the endocyclic nitrogen in DGJ, interacts with OH4 in DGJ-pFPhT. Other hydrogen bonding between α -Gal A and bound ligand DGJ-pFPhT, is very similar to that of DGJ (**Figure 44**), with the notable exception that hydrogen

bonding distances for the DGJ-*p*FPhT-bound structure are on average slightly longer than the distances for the DGJ-bound structure. The longer distances in the DGJ-*p*FPhT-bound structure could be because the fluoro substituent forms hydrogen bonding interactions with a residue from a neighboring molecule, specifically Arg 193. The DGJ-*p*FPhT-bound structure was also solved at pH 4.5, likely changing the protonation state of some of the interacting residues and increasing bond length. Based on the crystallographic binding data, further substituents on the aryl ring can be explored to optimize the potency of the chaperone.

An inhibition assay was done to determine IC₅₀ values for the 4 DGJ-ArTs and the reference DGJ derivatives. DGJ-*p*MeSPhT had an IC₅₀ value on the same order of magnitude as DGJ and all 4 DGJ-ArTs had significantly higher inhibitory potential than the reference DGJ derivatives. The difference in IC₅₀ values confirms the superiority of DGJ-ArTs over the reference DGJ derivatives as active-site directed α -Gal A ligands. Notably, the designed DGJ-ArTs have a significant increase in inhibitory potency from pH 5.0 to pH 7.0. This is important because the pharmacological chaperone must bind tightly to the enzyme in the endoplasmic reticulum (neutral) to stabilize the enzyme, but dissociate in the lysosome (acidic) to resume normal function (**Table 19**).

The activity assay with normal fibroblast cells revealed that little difference was seen between the designed DGJ-ArTs and the reference DGJ derivatives (**Figure 47a**), but a marked difference was observed in the FD fibroblast cells, which agrees with the starting hypothesis (**Figure 47b**). It is interesting that the optimal concentration for enhanced enzyme activity (30 μ M) is about 3 times the IC₅₀ value (**Table 19** and **Figure 47b**). This is observed for DGJ and other imino-sugar pharmacological chaperones. It

must be noted that the K_m of α -Gal A for the natural substrate Gb3, and at sufficiently high substrate concentrations the competition for binding will favour the natural substrate over the pharmacological chaperone which will drive the reaction toward Gb3 processing (277). It is unknown if the enzyme can balance hydrolysis at substrate concentrations below K_m , but the neutral nature of the DGJ-ArTs will likely be an advantage over the protonated inhibitor DGJ.

The activity assay with 17 mutants of α -Gal A and DGJ-*p*MeOPhT (**Figure 48b**) is consistent with other studies of pharmacological chaperones where the enzyme being studied retains full or partial catalytic activity upon successful trafficking from the endoplasmic reticulum to the lysosome (269, 276). DGJ-*p*MeOPhT was successful in enhancing the activity of α -Gal A in 15 of the 17 mutants tested, showing its potential for treatment of Fabry disease.

Since impairment of autophagy is a hallmark of cellular pathology in lysosomal storage disorders (278, 279), the levels of autophagy-related proteins, specifically LC3-II, p62, were tested (**Figure 49**). Levels of Bip, a protein related to ER stress, and beclin-1, a regulator of autophagy, were unchanged in FD-SV cells when exposed to starvation. This indicates that autophagy is impaired in the FD-SV cells, as expected. When FD-SV cells were treated with 2 DGJ-ArTs, levels of LC3-II and p62 both decreased, showing the great potential of the pharmacological chaperones for the treatment of Fabry disease.

In this chapter, we discuss 4 novel DGJ-ArT pharmacological chaperones for the treatment of Fabry disease. The molecular basis for their binding has been studied by X-ray crystallography and the compounds have been tested in normal and FD fibroblast cells. The neutral and amphiphilic nature of the designed DGJ-ArTs will likely make

them more suitable than the parent compound, DGJ. Based on our studies, DGJ-ArTs should be further explored for use as a treatment for Fabry disease, and design principles used here should be expanded to other lysosomal storage disorders.

CHAPTER 7: PERSPECTIVES AND FUTURE DIRECTIONS ON THE USE OF EPITOPE-SPECIFIC CRYSTALLIZATION CHAPERONES

7.1 PERSPECTIVES ON MEMBRANE PROTEIN CRYSTALLOGRAPHY

Membrane protein crystallography requires a lot of trial and error, and there are no shortcuts. Nevertheless, as evidenced throughout this thesis, many lessons can be learned in the process leading to eventual success in membrane protein crystallization. The high-risk/high-payoff goal of my thesis project was to solve a high resolution co-crystal structure of EE-tagged SPP in complex with the anti-EE Fab fragment, a goal that has yet to be attained. Indeed, the paper that details the Nobel Prize-winning structure of β 2AR in complex with its G protein ((105), see Chapter 1), has 20 authors (including key player in membrane protein crystallography in LCP, Martin Caffery), one of which (Brian Kobilka himself) was involved in the cloning of the first GPCR 25 years prior (280). Even after several GPCR structures had been solved and many of the techniques that were used for the β 2AR-G protein structure had already been used for previous structures, the researchers still had to test ~50 ligands, test ~50 detergents, generate new antibody fragments, optimize the placement of the lysozyme, and screen hundreds of crystals before the structure was obtained.

One important lesson from this thesis work is that it is not productive to set up crystallization trials of the same protein without making a major change. The changes I always made were in the crystallization trials: the sparse matrix screen used, temperature, protein concentration, method of protein complexation (adding the Fab/EE to the SPP-EE in the crystallization drop, or isolating the complex over gel filtration). Those parameters

are very important, but when no crystals are obtained after several sparse matrix trays, larger changes are needed. Protein screening and engineering seem to be keys to solving membrane protein structures, which we did not realize in the early stages of this project. We knew that the TM6-TM7 loop of *H. mar* SPP was very long (and likely flexible). We knew that we could not successfully solubilise *H. mar* SPP in DDM. We knew that FC12 was only once used successfully for every step of membrane protein isolation, purification, and crystallization, namely, the *E. coli* porin OmpF in 2010, a structure that had been previously solved in different detergents and was crystallized in FC12 as an accidental by-product (281). Perhaps we should have switched the target protein from *H. mar* SPP, which expressed extremely well, to *M. mar* SPP earlier, but this genome sequence was not available at the start of the project.

At the Frontiers of Structural Biology Keystone Conference in 2014, I discussed membrane protein crystallography with many people. Dr. Bergeron sent me his protocol, which consisted of making a small clone library (4-8 variants, starting and ending at different points based on alignments and secondary structure predictions) of the protein of interest all in a vector that included a cleavable hexahistidine tag. Small scale expression trials were then done by growing and inducing 10 mL cultures and running samples on SDS-PAGE. This is a very simple way to test protein expression of several variants in parallel, similar to the expression trials we conducted with intimin, and likely how SPP screening started out before I joined the project. The next step I took (with all the *M. mar* SPP constructs) was to scale up the expression to 3 L of culture, isolate membrane, and purify, which was too slow. The method of fusing GFP to the target protein and testing whole cell or in gel fluorescence has worked quite well for some

investigators in assessing and optimizing protein expression (See Section 2.4), and a GFP fusion can be used with a fluorescence detector in line with size exclusion chromatography to assess the shape, intensity, and elution volume of the protein peak (see Section 2.2.7 and 2.3.4) without the need to purify the protein. This requires constructing GFP-fusions and acquiring an in-line fluorescence detector, though.

Once expression and purification are optimized, a high (or medium) throughput way to determine protein stability in different conditions should be performed. The thermostability assay developed by the lab of Dr. Robert Stroud (159) is a useful tool if buried cysteine residues are present in the target protein. If no buried cysteines exist, SE-HPLC and SDS-PAGE are alternatives (see Section 2.4). Our method of thermal unfolding using CD was effective in determining stabilizing detergents (see **Figure 12**), but time consuming; an alternative is needed.

7.2 PERSPECTIVES ON THE USE OF EPITOPE-SPECIFIC CRYSTALLIZATION

CHAPERONES FOR MEMBRANE PROTEINS

From my thesis work, we now know that the insertion of the 6 amino acid EE epitope can cause a detrimental increase in flexibility of the loop where insertion occurred (see **Figure 28**). The flexibility of the linker region is likely the culprit for the difficulty in co-crystallization, as protein engineering efforts to decrease the loop flexibility with a shortened TM6-TM7 loop in *M. mar* SPP led to a promising co-crystal lead with complexed Fab/EE (see Section 4.3.4). Our method of using an epitope-specific antibody fragment as a chaperone is similar to the covalent addition of T4L to GPCRs (see Section 1.2.5). The addition of the antibody fragment, of course, is not covalent, but

both methods involve a short linker region (native loop residues) flanking a well folded protein. The covalent addition of T4L has been successful for numerous structures, likely because researchers tried many constructs prior to successful crystallization. In our attempt to develop a very general method that could be applied to any membrane protein, engineering of the protein to determine a suitable epitope insertion site was not attractive. Our method was supposed to be easier, less time-consuming, and less expensive than the generation of new crystallization chaperones for each target protein. It likely still will be, but several constructs of the membrane protein are probably going to need to be screened, paying careful attention to the loops within each protein target. Methods should be devised to scale up the process of evaluating chaperone binding to each potential membrane protein construct. In addition to optimization of Fab/EE:EE-tagged protein complexation, other epitopes (FLAG, for example) and fragment types (nanobody, for example) should be explored so that a toolbox of epitope-specific crystallization chaperones is generated, giving the researcher a choice for the optimal success.

7.3 CONCLUSIONS

In this thesis, I have presented the road taken to develop a stable, soluble, easily expressed and purified, epitope-specific antibody fragment as a general crystallization chaperone for membrane proteins. Though the long-term goal in obtaining a co-crystal structure of Fab/EE and SPP-EE was not realized, lessons that were learned on membrane proteins, detergents, antibody fragments, complexation, and complex crystallization are very important to the Lieberman lab and lay the foundation for the future success of epitope-specific crystallization chaperones.

REFERENCES

1. Robert, X. and P. Gouet, *Deciphering key features in protein structures with the new endscrip server*. Nucleic Acids Res, 2014. **42**(W1): p. W320-W324.
2. Yang, H.W., D.Q. Liu, X. Fan, M.A. White, and R.O. Fox, *Conformational ensemble analysis of ligand binding in streptavidin mini-protein complexes*. To be published.
3. Yang, H., J. Wang, X. Jia, M.W. McNatt, T. Zang, B. Pan, W. Meng, H.W. Wang, P.D. Bieniasz, and Y. Xiong, *Structural insight into the mechanisms of enveloped virus tethering by tetherin*. Proc Natl Acad Sci U S A, 2010. **107**(43): p. 18428-18432.
4. Li, X., S. Dang, C. Yan, X. Gong, J. Wang, and Y. Shi, *Structure of a presenilin family intramembrane aspartate protease*. Nature, 2013. **493**(7430): p. 56-61.
5. Voss, M., B. Schröder, and R. Fluhrer, *Mechanism, specificity, and physiology of signal peptide peptidase (spp) and spp-like proteases*. BBA - Biomembranes, 2013. **1828**(12): p. 2828-2839.
6. Lim, K.H., H. Huang, A. Pralle, and S. Park, *Stable, high-affinity streptavidin monomer for protein labeling and monovalent biotin detection*. Biotechnol Bioeng, 2013. **110**(1): p. 57-67.
7. Kaufmann, M., P. Lindner, A. Honegger, K. Blank, M. Tschopp, G. Capitani, A. Plückthun, and M.G. Grütter, *Crystal structure of the anti-his tag antibody 3d5 single-chain fragment complexed to its antigen*. JMB, 2002. **318**(1): p. 135-147.
8. Hyre, D.E., I. Le Trong, E.A. Merritt, J.F. Eccleston, N.M. Green, R.E. Stenkamp, and P.S. Stayton, *Cooperative hydrogen bond interactions in the streptavidin-biotin system*. Protein sci, 2006. **15**(3): p. 459-467.
9. Laitinen, O.H., H.R. Nordlund, V.P. Hytönen, and M.S. Kulomaa, *Brave new (strept)avidins in biotechnology*. Trends Biotechnol, 2007. **25**(6): p. 269-277.
10. Wallin, E. and G. von Heijne, *Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms*. Protein Sci, 1998. **7**(4): p. 1029-1038.
11. Almén, M.S., K.J.V. Nordström, R. Fredriksson, and H.B. Schiöth, *Mapping the human membrane proteome: A majority of the human membrane proteins can be classified according to function and evolutionary origin*. BMC Biology, 2009. **7**: p. 50-50.

12. Overington, J.P., B. Al-Lazikani, and A.L. Hopkins, *How many drug targets are there?* Nat Rev Drug Discov, 2006. **5**(12): p. 993-996.
13. Sherrington, R., E.I. Rogaev, Y. Liang, E.A. Rogaeva, G. Levesque, M. Ikeda, H. Chi, C. Lin, G. Li, K. Holman, T. Tsuda, L. Mar, J.F. Foncin, A.C. Bruni, M.P. Montesi, S. Sorbi, I. Rainero, L. Pinessi, L. Nee, I. Chumakov, D. Pollen, A. Brookes, P. Sanseau, R.J. Polinsky, W. Wasco, H.A. Da Silva, J.L. Haines, M.A. Pericak-Vance, R.E. Tanzi, A.D. Roses, P.E. Fraser, J.M. Rommens, and P.H. St George-Hyslop, *Cloning of a gene bearing missense mutations in early-onset familial alzheimer's disease.* Nature, 1995. **375**(6534): p. 754-760.
14. Cutting, G.R., *Cystic fibrosis genetics: From molecular understanding to clinical application.* Nat Rev Genet, 2015. **16**(1): p. 45-56.
15. Hizawa, N., *Beta-2 adrenergic receptor genetic polymorphisms and asthma.* J Clin Pharm Ther, 2009. **34**(6): p. 631-643.
16. Drenth, J., *Principles of protein x-ray crystallography.* 2010: Springer.
17. Sherwood, D. and J. Cooper, *Crystals, x-rays and proteins: Comprehensive protein crystallography.* 2010.
18. Surade, S. and Tom L. Blundell, *Structural biology and drug discovery of difficult targets: The limits of ligandability.* Chem Biol, 2012. **19**(1): p. 42-50.
19. Scapin, G., *Structural biology and drug discovery.* Curr Pharm Design. **12**(17): p. 2087-2097.
20. Fernandez, J.M. and J.P. Hoeffler, *Introduction*, in *Gene expression systems*, J.M. Fernandez and J.P. Hoeffler, Editors. 1999, Academic Press: San Diego. p. 1-5.
21. Luo, P., X. Yu, W. Wang, S. Fan, X. Li, and J. Wang, *Crystal structure of a phosphorylation-coupled vitamin c transporter.* Nat Struct Mol Biol, 2015. **22**(3): p. 238-241.
22. Akyuz, N., E.R. Georgieva, Z. Zhou, S. Stolzenberg, M.A. Cuendet, G. Khelashvili, R.B. Altman, D.S. Terry, J.H. Freed, H. Weinstein, O. Boudker, and S.C. Blanchard, *Transport domain unlocking sets the uptake rate of an aspartate transporter.* Nature, 2015. **518**(7537): p. 68-73.
23. Li, F., J. Liu, Y. Zheng, R.M. Garavito, and S. Ferguson-Miller, *Crystal structures of translocator protein (tspo) and mutant mimic of a human polymorphism.* Science, 2015. **347**(6221): p. 555-558.
24. Lieberman, R.L., J.A. Culver, K.C. Entzminger, J.C. Pai, and J.A. Maynard, *Crystallization chaperone strategies for membrane proteins.* Methods, 2011. **55**(4): p. 293-302.

25. Tanabe, H., Y. Fujii, M. Okada-Iwabu, M. Iwabu, Y. Nakamura, T. Hosaka, K. Motoyama, M. Ikeda, M. Wakiyama, T. Terada, N. Ohsawa, M. Hato, S. Ogasawara, T. Hino, T. Murata, S. Iwata, K. Hirata, Y. Kawano, M. Yamamoto, T. Kimura-Someya, M. Shirouzu, T. Yamauchi, T. Kadowaki, and S. Yokoyama, *Crystal structures of the human adiponectin receptors*. *Nature*, 2015. **520**(7547): p. 312-316.
26. Tanabe, H., K. Motoyama, M. Ikeda, M. Wakiyama, T. Terada, N. Ohsawa, T. Hosaka, M. Hato, Y. Fujii, Y. Nakamura, S. Ogasawara, T. Hino, T. Murata, S. Iwata, M. Okada-Iwabu, M. Iwabu, K. Hirata, Y. Kawano, M. Yamamoto, T. Kimura-Someya, M. Shirouzu, T. Yamauchi, T. Kadowaki, and S. Yokoyama, *Expression, purification, crystallization, and preliminary x-ray crystallographic studies of the human adiponectin receptors, adipor1 and adipor2*. *J Struct Funct Genomics*, 2015. **16**(1): p. 11-23.
27. Dong, Y.Y., A.C.W. Pike, A. Mackenzie, C. McClenaghan, P. Aryal, L. Dong, A. Quigley, M. Grieben, S. Goubin, S. Mukhopadhyay, G.F. Ruda, M.V. Clausen, L. Cao, P.E. Brennan, N.A. Burgess-Brown, M.S.P. Sansom, S.J. Tucker, and E.P. Carpenter, *K2p channel gating mechanisms revealed by structures of trek-2 and a complex with prozac*. *Science*, 2015. **347**(6227): p. 1256-1259.
28. Cooper, D.R., T. Boczek, K. Grelewska, M. Pinkowska, M. Sikorska, M. Zawadzki, and Z. Derewenda, *Protein crystallization by surface entropy reduction: Optimization of the ser strategy*. *Acta Crystallogr D*, 2007. **63**(5): p. 636-645.
29. Longenecker, K.L., M.E. Lewis, H. Chikumi, J.S. Gutkind, and Z.S. Derewenda, *Structure of the rgs-like domain from pdz-rhogef: Linking heterotrimeric g protein-coupled signaling to rho gtpases*. *Structure*, 2001. **9**(7): p. 559-569.
30. Bielnicki, J., Y. Devedjiev, U. Derewenda, Z. Dauter, A. Joachimiak, and Z.S. Derewenda, *B. Subtilis ykud protein at 2.0 a resolution: Insights into the structure and function of a novel, ubiquitous family of bacterial enzymes*. *Proteins*, 2006. **62**(1): p. 144-151.
31. Moon, A.F., G.A. Mueller, X. Zhong, and L.C. Pedersen, *A synergistic approach to protein crystallization: Combination of a fixed-arm carrier with surface entropy reduction*. *Protein sci*, 2010. **19**(5): p. 901-913.
32. Banatao, D.R., D. Cascio, C.S. Crowley, M.R. Fleissner, H.L. Tienson, and T.O. Yeates, *An approach to crystallizing proteins by synthetic symmetrization*. *PROC NATL ACAD SCI U S A*, 2006. **103**(44): p. 16230-16235.
33. Forse, G.J., N. Ram, D.R. Banatao, D. Cascio, M.R. Sawaya, H.E. Klock, S.A. Lesley, and T.O. Yeates, *Synthetic symmetrization in the crystallization and structure determination of celsr from thermotoga maritima*. *Protein Sci*, 2011. **20**(1): p. 168-178.

34. Laganowsky, A., M. Zhao, A.B. Soriaga, M.R. Sawaya, D. Cascio, and T.O. Yeates, *An approach to crystallizing proteins by metal-mediated synthetic symmetrization*. Protein Sci, 2011. **20**(11): p. 1876-1890.
35. Huard, D.J.E., K.M. Kane, and F.A. Tezcan, *Re-engineering protein interfaces yields copper-inducible ferritin cage assembly*. Nat Chem Biol, 2013. **9**(3): p. 169-176.
36. Saul, F.A., B. Vulliez-le Normand, F. Lema, and G.A. Bentley, *Crystal structure of a recombinant form of the maltodextrin-binding protein carrying an inserted sequence of a b-cell epitope from the pres2 region of hepatitis b virus*. Proteins, 1997. **27**(1): p. 1-8.
37. Wiltzius, J.J., S.A. Sievers, M.R. Sawaya, and D. Eisenberg, *Atomic structures of iapp (amylin) fusions suggest a mechanism for fibrillation and the role of insulin in the process*. Protein sci, 2009. **18**(7): p. 1521-1530.
38. Mendillo, M.L., C.D. Putnam, and R.D. Kolodner, *Escherichia coli muts tetramerization domain structure reveals that stable dimers but not tetramers are essential for DNA mismatch repair in vivo*. JBC, 2007. **282**(22): p. 16345-16354.
39. Nauli, S., S. Farr, Y.J. Lee, H.Y. Kim, S. Faham, and J.U. Bowie, *Polymer-driven crystallization*. Protein Sci, 2007. **16**(11): p. 2542-2551.
40. Corsini, L., M. Hothorn, K. Scheffzek, M. Sattler, and G. Stier, *Thioredoxin as a fusion tag for carrier-driven crystallization*. Protein Sci, 2008. **17**(12): p. 2070-2079.
41. Hanson, M.A., V. Cherezov, M.T. Griffith, C.B. Roth, V.P. Jaakola, E.Y. Chien, J. Velasquez, P. Kuhn, and R.C. Stevens, *A specific cholesterol binding site is established by the 2.8 a structure of the human β 2-adrenergic receptor*. Structure, 2008. **16**(6): p. 897-905.
42. Cherezov, V., D.M. Rosenbaum, M.A. Hanson, S.G. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka, and R.C. Stevens, *High-resolution crystal structure of an engineered human β 2-adrenergic g protein-coupled receptor*. Science, 2007. **318**(5854): p. 1258-1265.
43. Jaakola, V.P., M.T. Griffith, M.A. Hanson, V. Cherezov, E.Y. Chien, J.R. Lane, A.P. Ijzerman, and R.C. Stevens, *The 2.6 angstrom crystal structure of a human a2a adenosine receptor bound to an antagonist*. Science, 2008. **322**(5905): p. 1211-1217.
44. Wu, B., E.Y. Chien, C.D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F.C. Bi, D.J. Hamel, P. Kuhn, T.M. Handel, V. Cherezov, and R.C. Stevens, *Structures of the cxcr4 chemokine gpcr with small-molecule and cyclic peptide antagonists*. Science, 2010. **330**(6007): p. 1066-1071.

45. Chun, E., Aaron A. Thompson, W. Liu, Christopher B. Roth, Mark T. Griffith, V. Katritch, J. Kunken, F. Xu, V. Cherezov, Michael A. Hanson, and Raymond C. Stevens, *Fusion partner toolchest for the stabilization and crystallization of g protein-coupled receptors*. Structure, 2012. **20**(6): p. 967-976.
46. Thompson, A.A., W. Liu, E. Chun, V. Katritch, H. Wu, E. Vardy, X.-P. Huang, C. Trapella, R. Guerrini, G. Calo, B.L. Roth, V. Cherezov, and R.C. Stevens, *Structure of the nociceptin/orphanin fq receptor in complex with a peptide mimetic*. Nature, 2012. **485**(7398): p. 395-399.
47. Kapust, R.B. and D.S. Waugh, *Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused*. Protein sci, 1999. **8**(8): p. 1668-1674.
48. Sachdev, D. and J.M. Chirgwin, *Fusions to maltose-binding protein: Control of folding and solubility in protein purification*. Methods Enzymol, 2000. **326**: p. 312-321.
49. Braun, P., Y. Hu, B. Shen, A. Halleck, M. Koundinya, E. Harlow, and J. LaBaer, *Proteome-scale purification of human proteins from bacteria*. Proc Natl Acad Sci U S A, 2002. **99**(5): p. 2654-2659.
50. Shih, Y.P., W.M. Kung, J.C. Chen, C.H. Yeh, A.H. Wang, and T.F. Wang, *High-throughput screening of soluble recombinant proteins*. Protein sci, 2002. **11**(7): p. 1714-1719.
51. Edwards, A.M., C.H. Arrowsmith, D. Christendat, A. Dharamsi, J.D. Friesen, J.F. Greenblatt, and M. Vedadi, *Protein production: Feeding the crystallographers and nmr spectroscopists*. Nat struct bio, 2000. **7 Suppl**: p. 970-972.
52. Stevens, R.C., *Design of high-throughput methods of protein production for structural biology*. Structure, 2000. **8**(9): p. R177-R185.
53. Smyth, D.R., M.K. Mrozkiewicz, W.J. McGrath, P. Listwan, and B. Kobe, *Crystal structures of fusion proteins with large-affinity tags*. Protein sci, 2003. **12**(7): p. 1313-1322.
54. Miller-Gallacher, J.L., R. Nehme, T. Warne, P.C. Edwards, G.F. Schertler, A.G. Leslie, and C.G. Tate, *The 2.1 a resolution structure of cyanopindolol-bound beta1-adrenoceptor identifies an intramembrane na+ ion that stabilises the ligand-free receptor*. PLoS One, 2014. **9**(3): p. e92727.
55. Doré, Andrew S., N. Robertson, James C. Errey, I. Ng, K. Hollenstein, B. Tehan, E. Hurrell, K. Bennett, M. Congreve, F. Magnani, Christopher G. Tate, M. Weir, and Fiona H. Marshall, *Structure of the adenosine a2a receptor in complex with zm241385 and the xanthines xac and caffeine*. Structure. **19**(9): p. 1283-1293.

56. Warne, T., M.J. Serrano-Vega, C.G. Tate, and G.F.X. Schertler, *Development and crystallization of a minimal thermostabilised g protein-coupled receptor*. Protein Expres Purif, 2009. **65**(2): p. 204-213.
57. Pautsch, A., J. Vogt, K. Model, C. Siebold, and G.E. Schulz, *Strategy for membrane protein crystallization exemplified with ompa and ompx*. Proteins, 1999. **34**(2): p. 167-172.
58. Avoigt, C., S. Kauffman, and Z.-G. Wang, *Rational evolutionary design: The theory of in vitro protein evolution*, in *Advances in protein chemistry*, H.A. Frances, Editor. 2001, Academic Press. p. 79-160.
59. Schlinkmann, K.M. and A. Pluckthun, *Directed evolution of g-protein-coupled receptors for high functional expression and detergent stability*. Methods Enzymol, 2013. **520**: p. 67-97.
60. Sarkar, C.A., I. Dodevski, M. Kenig, S. Dudli, A. Mohr, E. Hermans, and A. Plückthun, *Directed evolution of a g protein-coupled receptor for expression, stability, and binding selectivity*. Proc Natl Acad Sci U S A, 2008. **105**(39): p. 14808-14813.
61. Bernaudat, F., A. Frelet-Barrand, N. Pochon, S. Dementin, P. Hivin, S. Boutigny, J.B. Rioux, D. Salvi, D. Seigneurin-Berny, P. Richaud, J. Joyard, D. Pignol, M. Sabaty, T. Desnos, E. Pebay-Peyroula, E. Darrouzet, T. Vernet, and N. Rolland, *Heterologous expression of membrane proteins: Choosing the appropriate host*. PLoS One, 2011. **6**(12): p. e29191.
62. Sahdev, S., S.K. Khattar, and K.S. Saini, *Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies*. Mol Cell Biochem, 2008. **307**(1-2): p. 249-264.
63. Hou, X., L. Pedi, M.M. Diver, and S.B. Long, *Crystal structure of the calcium release-activated calcium channel orai*. Science, 2012. **338**(6112): p. 1308-1313.
64. Brohawn, S.G., J. del Marmol, and R. MacKinnon, *Crystal structure of the human k2p traak, a lipid- and mechano-sensitive k⁺ ion channel*. Science, 2012. **335**(6067): p. 436-441.
65. Parker, J.L. and S. Newstead, *Molecular basis of nitrate uptake by the plant nitrate transporter nrt1.1*. Nature, 2014. **507**(7490): p. 68-72.
66. Pryor, E.E., Jr., P.S. Horanyi, K.M. Clark, N. Fedoriw, S.M. Connelly, M. Koszelak-Rosenblum, G. Zhu, M.G. Malkowski, M.C. Wiener, and M.E. Dumont, *Structure of the integral membrane protein caax protease ste24p*. Science, 2013. **339**(6127): p. 1600-1604.

67. Penmatsa, A., K.H. Wang, and E. Gouaux, *X-ray structure of dopamine transporter elucidates antidepressant mechanism*. Nature, 2013. **503**(7474): p. 85-90.
68. Gruswitz, F., S. Chaudhary, J.D. Ho, A. Schlessinger, B. Pezeshki, C.M. Ho, A. Sali, C.M. Westhoff, and R.M. Stroud, *Function of human rh based on structure of rhcg at 2.1 a*. Proc Natl Acad Sci U S A, 2010. **107**(21): p. 9638-9643.
69. Standfuss, J., G. Xie, P.C. Edwards, M. Burghammer, D.D. Oprian, and G.F. Schertler, *Crystal structure of a thermally stable rhodopsin mutant*. J Mol Biol, 2007. **372**(5): p. 1179-1188.
70. Qin, L., I. Kufareva, L.G. Holden, C. Wang, Y. Zheng, C. Zhao, G. Fenalti, H. Wu, G.W. Han, V. Cherezov, R. Abagyan, R.C. Stevens, and T.M. Handel, *Crystal structure of the chemokine receptor cxcr4 in complex with a viral chemokine*. Science, 2015. **347**(6226): p. 1117-1122.
71. Wacker, D., C. Wang, V. Katritch, G.W. Han, X.P. Huang, E. Vardy, J.D. McCorvy, Y. Jiang, M. Chu, F.Y. Siu, W. Liu, H.E. Xu, V. Cherezov, B.L. Roth, and R.C. Stevens, *Structural features for functional selectivity at serotonin receptors*. Science, 2013. **340**(6132): p. 615-619.
72. He, Y., K. Wang, and N. Yan, *The recombinant expression systems for structure determination of eukaryotic membrane proteins*. Protein & Cell, 2014. **5**(9): p. 658-672.
73. Privé, G.G., *Detergents for the stabilization and crystallization of membrane proteins*. Methods, 2007. **41**(4): p. 388-397.
74. le Maire, M., P. Champeil, and J.V. Moller, *Interaction of membrane proteins and lipids with solubilizing detergents*. Biochim Biophys Acta, 2000. **1508**(1-2): p. 86-111.
75. Tribet, C., R. Audebert, and J.-L. Popot, *Amphipols: Polymers that keep membrane proteins soluble in aqueous solutions*. Proc Natl Acad Sci U S A, 1996. **93**(26): p. 15047-15050.
76. Wernimont, A. and A. Edwards, *in situ proteolysis to generate crystals for structure determination: An update*. PLoS ONE, 2009. **4**(4): p. e5094.
77. Deniaud, A., L. Liguori, I. Blesneac, J.L. Lenormand, and E. Pebay-Peyroula, *Crystallization of the membrane protein hvdac1 produced in cell-free system*. BBA - Biomembranes, 2010. **1798**(8): p. 1540-1546.
78. Dong, A., X. Xu, A.M. Edwards, C. Chang, M. Chruszcz, M. Cuff, M. Cymborowski, R. Di Leo, O. Egorova, E. Evdokimova, E. Filippova, J. Gu, J. Guthrie, A. Ignatchenko, A. Joachimiak, N. Klostermann, Y. Kim, Y. Korniyenko, W. Minor, Q. Que, A. Savchenko, T. Skarina, K. Tan, A. Yakunin,

- A. Yee, V. Yim, R. Zhang, H. Zheng, M. Akutsu, C. Arrowsmith, G.V. Avvakumov, A. Bochkarev, L.G. Dahlgren, S. Dhe-Paganon, S. Dimov, L. Dombrowski, P. Finerty, Jr., S. Flodin, A. Flores, S. Graslund, M. Hammerstrom, M.D. Herman, B.S. Hong, R. Hui, I. Johansson, Y. Liu, M. Nilsson, L. Nedyalkova, P. Nordlund, T. Nyman, J. Min, H. Ouyang, H.W. Park, C. Qi, W. Rabeh, L. Shen, Y. Shen, D. Sukumard, W. Tempel, Y. Tong, L. Tresagues, M. Vedadi, J.R. Walker, J. Weigelt, M. Welin, H. Wu, T. Xiao, H. Zeng, and H. Zhu, *In situ proteolysis for protein crystallization and structure determination*. Nat Methods, 2007. **4**(12): p. 1019-1021.
79. Chiu, M.L., P. Nollert, M.C. Loewen, H. Belrhali, E. Pebay-Peyroula, J.P. Rosenbusch, and E.M. Landau, *Crystallization in cubo: General applicability to membrane proteins*. Acta Crystallogr D Biol Crystallogr, 2000. **56**(Pt 6): p. 781-784.
80. Burg, J.S., J.R. Ingram, A.J. Venkatakrishnan, K.M. Jude, A. Dukkupati, E.N. Feinberg, A. Angelini, D. Waghray, R.O. Dror, H.L. Ploegh, and K.C. Garcia, *Structural biology. Structural basis for chemokine recognition and activation of a viral g protein-coupled receptor*. Science, 2015. **347**(6226): p. 1113-1117.
81. Lee, Y., T. Nishizawa, K. Yamashita, R. Ishitani, and O. Nureki, *Structural basis for the facilitative diffusion mechanism by semisweet transporter*. Nature Comm, 2015. **6**: p. 6112.
82. Fenalti, G., N.A. Zatsepin, C. Betti, P. Giguere, G.W. Han, A. Ishchenko, W. Liu, K. Guillemyn, H. Zhang, D. James, D. Wang, U. Weierstall, J.C. Spence, S. Boutet, M. Messerschmidt, G.J. Williams, C. Gati, O.M. Yefanov, T.A. White, D. Oberthuer, M. Metz, C.H. Yoon, A. Barty, H.N. Chapman, S. Basu, and J. Coe, *Structural basis for bifunctional peptide recognition at human delta-opioid receptor*. Nature, 2015. **22**(3): p. 265-268.
83. Weinert, T., V. Olieric, S. Waltersperger, E. Panepucci, L. Chen, H. Zhang, D. Zhou, J. Rose, A. Ebihara, S. Kuramitsu, D. Li, N. Howe, G. Schnapp, A. Pautsch, K. Bargsten, A.E. Protá, P. Surana, J. Kottur, D.T. Nair, and F. Basilico, *Fast native-sad phasing for routine macromolecular structure determination*. Nat Methods, 2015. **12**(2): p. 131-133.
84. Caffrey, M., *Membrane protein crystallization*. J Struct Biol, 2003. **142**(1): p. 108-132.
85. Caffrey, M., *Crystallizing membrane proteins for structure-function studies using lipidic mesophases*. Biochem Soc Trans, 2011. **39**(3): p. 725-732.
86. Caffrey, M. and C. Porter, *Crystallizing membrane proteins for structure determination using lipidic mesophases*. JoVE, 2010(45): p. e1712.

87. Xu, F., H. Wu, V. Katritch, G.W. Han, K.A. Jacobson, Z.G. Gao, V. Cherezov, and R.C. Stevens, *Structure of an agonist-bound human a2a adenosine receptor*. Science, 2011. **332**(6027): p. 322-327.
88. Rosenbaum, D.M., C. Zhang, J.A. Lyons, R. Holl, D. Aragao, D.H. Arlow, S.G. Rasmussen, H.J. Choi, B.T. Devree, R.K. Sunahara, P.S. Chae, S.H. Gellman, R.O. Dror, D.E. Shaw, W.I. Weis, M. Caffrey, P. Gmeiner, and B.K. Kobilka, *Structure and function of an irreversible agonist-beta(2) adrenoceptor complex*. Nature, 2011. **469**(7329): p. 236-240.
89. Singh, S.K., A. Yamashita, and E. Gouaux, *Antidepressant binding site in a bacterial homologue of neurotransmitter transporters*. Nature, 2007. **448**(7156): p. 952-956.
90. Hibbs, R.E. and E. Gouaux, *Principles of activation and permeation in an anion-selective cys-loop receptor*. Nature, 2011. **474**(7349): p. 54-60.
91. Abramson, J., I. Smirnova, V. Kasho, G. Verner, H.R. Kaback, and S. Iwata, *Structure and mechanism of the lactose permease of escherichia coli*. Science, 2003. **301**(5633): p. 610-615.
92. Hunte, C. and H. Michel, *Crystallisation of membrane proteins mediated by antibody fragments*. Curr Opin Struct Biol, 2002. **12**(4): p. 503-508.
93. Milstein, C., *The hybridoma revolution: An offshoot of basic research*. BioEssays, 1999. **21**(11): p. 966-973.
94. Tomita, M. and K. Tsumoto, *Hybridoma technologies for antibody production*. Immunotherapy, 2011. **3**(3): p. 371-380.
95. Röthlisberger, D., K.M. Pos, and A. Plückthun, *An antibody library for stabilizing and crystallizing membrane proteins – selecting binders to the citrate carrier cits*. FEBS Letters, 2004. **564**(3): p. 340-348.
96. Tanha, J., G. Dubuc, T. Hiramata, S.A. Narang, and C.R. MacKenzie, *Selection by phage display of llama conventional vh fragments with heavy chain antibody vhh properties*. J Immunol Methods, 2002. **263**(1–2): p. 97-109.
97. Fang, Y., H. Jayaram, T. Shane, L. Kolmakova-Partensky, F. Wu, C. Williams, Y. Xiong, and C. Miller, *Structure of a prokaryotic virtual proton pump at 3.2 a resolution*. Nature, 2009. **460**(7258): p. 1040-1043.
98. Shaffer, P.L., A. Goehring, A. Shankaranarayanan, and E. Gouaux, *Structure and mechanism of a na+-independent amino acid transporter*. Science, 2009. **325**(5943): p. 1010-1014.
99. Rasmussen, S.G.F., H.-J. Choi, D.M. Rosenbaum, T.S. Kobilka, F.S. Thian, P.C. Edwards, M. Burghammer, V.R.P. Ratnala, R. Sanishvili, R.F. Fischetti, G.F.X.

- Schertler, W.I. Weis, and B.K. Kobilka, *Crystal structure of the human β 2 adrenergic g-protein-coupled receptor*. *Nature*, 2007. **450**(7168): p. 383-387.
100. Inaba, K., S. Murakami, A. Nakagawa, H. Iida, M. Kinjo, K. Ito, and M. Suzuki, *Dynamic nature of disulphide bond formation catalysts revealed by crystal structures of dsbb*. *EMBO J*, 2009. **28**(6): p. 779-791.
 101. Lee, S.Y., A. Lee, J. Chen, and R. MacKinnon, *Structure of the kvap voltage-dependent k^+ channel and its dependence on the lipid membrane*. *Proc Natl Acad Sci U S A*, 2005. **102**(43): p. 15441-15446.
 102. Hunte, C., J. Koepke, C. Lange, T. Rossmann, and H. Michel, *Structure at 2.3 Å resolution of the cytochrome bc(1) complex from the yeast *saccharomyces cerevisiae* co-crystallized with an antibody fv fragment*. *Structure*, 2000. **8**(6): p. 669-684.
 103. Ostermeier, C., A. Harrenga, U. Ermler, and H. Michel, *Structure at 2.7 Å resolution of the *paracoccus denitrificans* two-subunit cytochrome c oxidase complexed with an antibody fv fragment*. *Proc Natl Acad Sci U S A*, 1997. **94**(20): p. 10547-10553.
 104. Rasmussen, S.G., H.J. Choi, J.J. Fung, E. Pardon, P. Casarosa, P.S. Chae, B.T. Devree, D.M. Rosenbaum, F.S. Thian, T.S. Kobilka, A. Schnapp, I. Konetzki, R.K. Sunahara, S.H. Gellman, A. Pautsch, J. Steyaert, W.I. Weis, and B.K. Kobilka, *Structure of a nanobody-stabilized active state of the β 2 adrenoceptor*. *Nature*, 2011. **469**(7329): p. 175-180.
 105. Rasmussen, S.G., B.T. DeVree, Y. Zou, A.C. Kruse, K.Y. Chung, T.S. Kobilka, F.S. Thian, P.S. Chae, E. Pardon, D. Calinski, J.M. Mathiesen, S.T. Shah, J.A. Lyons, M. Caffrey, S.H. Gellman, J. Steyaert, G. Skinotis, W.I. Weis, R.K. Sunahara, and B.K. Kobilka, *Crystal structure of the β 2 adrenergic receptor-gs protein complex*. *Nature*, 2011. **477**(7366): p. 549-555.
 106. Low, C., Y.H. Yau, E. Pardon, C. Jegerschold, L. Wahlin, E.M. Quistgaard, P. Moberg, S. Geifman-Shochat, J. Steyaert, and P. Nordlund, *Nanobody mediated crystallization of an archeal mechanosensitive channel*. *PLoS One*, 2013. **8**(10): p. e77984.
 107. Shea, C., L. Bloedorn, and M.A. Sullivan, *Rapid isolation of single-chain antibodies for structural genomics*. *J Struct Funct Genomics*, 2005. **6**(2-3): p. 171-175.
 108. Venturi, M. and C. Hunte, *Monoclonal antibodies for the structural analysis of the na^+/h^+ antiporter *nhaa* from *escherichia coli**. *Biochim Biophys Acta*, 2003. **1610**(1): p. 46-50.
 109. Huber, T., D. Steiner, D. Rothlisberger, and A. Pluckthun, *In vitro selection and characterization of darpins and fab fragments for the co-crystallization of*

- membrane proteins: The na⁺-citrate symporter cas as an example.* J Struct Biol, 2007. **159**(2): p. 206-221.
110. Milovnik, P., D. Ferrari, C.A. Sarkar, and A. Plückthun, *Selection and characterization of darpins specific for the neurotensin receptor 1.* Protein Eng Des Sel, 2009. **22**(6): p. 357-366.
 111. Binz, H.K., M.T. Stumpp, P. Forrer, P. Amstutz, and A. Pluckthun, *Designing repeat proteins: Well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins.* J Mol Biol, 2003. **332**(2): p. 489-503.
 112. Forrer, P., M.T. Stumpp, H.K. Binz, and A. Plückthun, *A novel strategy to design binding molecules harnessing the modular nature of repeat proteins.* FEBS Letters, 2003. **539**(1-3): p. 2-6.
 113. Kohl, A., H.K. Binz, P. Forrer, M.T. Stumpp, A. Plückthun, and M.G. Grütter, *Designed to be stable: Crystal structure of a consensus ankyrin repeat protein.* Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1700-1705.
 114. Lluís, M.W., J.I. Godfroy, and H. Yin, *Protein engineering methods applied to membrane protein targets.* Protein Eng Des Sel, 2013. **26**(2): p. 91-100.
 115. Sennhauser, G., P. Amstutz, C. Briand, O. Storchenegger, and M.G. Grutter, *Drug export pathway of multidrug exporter acrb revealed by darpin inhibitors.* PLoS biology, 2007. **5**(1): p. e7.
 116. Koldobskaya, Y., E.M. Duguid, D.M. Shechner, N.B. Suslov, J. Ye, S.S. Sidhu, D.P. Bartel, S. Koide, A.A. Kossiakoff, and J.A. Piccirilli, *A portable rna sequence whose recognition by a synthetic antibody facilitates structural determination.* Nat Struct Mol Biol, 2011. **18**: p. 100-106.
 117. Piccirilli, J.A. and Y. Koldobskaya, *Crystal structure of an rna polymerase ribozyme in complex with an antibody fragment.* Phil Trans R Soc B, 2011. **366**: p. 2918-2928.
 118. Roosild, T.P., S. Castronovo, and S. Choe, *Structure of anti-flag m2 fab domain and its use in the stabilization of engineered membrane proteins.* Acta. Cryst., 2006. **F62**: p. 835-839.
 119. Kalyoncu, S., J. Hyun, J.C. Pai, J.L. Johnson, K. Entzminger, A. Jain, D.P. Heaner, Jr., I.A. Morales, T.M. Truskett, J.A. Maynard, and R.L. Lieberman, *Effects of protein engineering and rational mutagenesis on crystal lattice of single chain antibody fragments.* Proteins, 2014. **82**(9): p. 1884-1895.
 120. Johnson, J.L., K.C. Entzminger, J. Hyun, S. Kalyoncu, D.P. Heaner, Jr, I.A. Morales, A. Sheppard, J.C. Gumbart, J.A. Maynard, and R.L. Lieberman, *Structural and biophysical characterization of an epitope-specific engineered fab*

- fragment and complexation with membrane proteins: Implications for co-crystallization.* Acta Crystallogr D, 2015. **71**(4): p. 896-906.
121. Wong, J., A. Chilkoti, and V.T. Moy, *Direct force measurements of the streptavidin–biotin interaction.* Biomol Eng, 1999. **16**(1–4): p. 45-55.
 122. Lim, K.H., H. Huang, A. Pralle, and S. Park, *Engineered streptavidin monomer and dimer with improved stability and function.* Biochemistry, 2011. **50**(40): p. 8682-8691.
 123. Mancina, F. and J. Love, *High-throughput expression and purification of membrane proteins.* J Struct Biol, 2010. **172**(1): p. 85-93.
 124. Wagner, S., M.L. Bader, D. Drew, and J.-W. de Gier, *Rationalizing membrane protein overexpression.* Trends Biotechnol, 2006. **24**(8): p. 364-371.
 125. Yamashita, A., S.K. Singh, T. Kawate, Y. Jin, and E. Gouaux, *Crystal structure of a bacterial homologue of na⁺/cl⁻-dependent neurotransmitter transporters.* Nature, 2005. **437**(7056): p. 215-223.
 126. Yernool, D., O. Boudker, Y. Jin, and E. Gouaux, *Structure of a glutamate transporter homologue from pyrococcus horikoshii.* Nature, 2004. **431**(7010): p. 811-818.
 127. Corringer, P.J., M. Baaden, N. Bocquet, M. Delarue, V. Dufresne, H. Nury, M. Prevost, and C. Van Renterghem, *Atomic structure and dynamics of pentameric ligand-gated ion channels: New insight from bacterial homologues.* J Physiol, 2010. **588**(Pt 4): p. 565-572.
 128. Weihofen, A., K. Binns, M.K. Lemberg, K. Ashman, and B. Martoglio, *Identification of signal peptide peptidase, a presenilin-type aspartic protease.* Science, 2002. **296**(5576): p. 2215-2218.
 129. Haass, C. and H. Steiner, *Alzheimer disease γ -secretase: A complex story of gxd-type presenilin proteases.* Trends Cell Biol, **12**(12): p. 556-562.
 130. Kilic, A., S. Klose, B. Dobberstein, E. Knust, and K. Kapp, *The drosophila crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase.* Eur J Cell Biol, 2010. **89**(6): p. 449-461.
 131. Wu, C.M. and M.D. Chang, *Signal peptide of eosinophil cationic protein is toxic to cells lacking signal peptide peptidase.* Biochem Biophys Res Commun, 2004. **322**(2): p. 585-592.
 132. El Hage, F., V. Stroobant, I. Vergnon, J.F. Baurain, H. Echchakir, V. Lazar, S. Chouaib, P.G. Coulie, and F. Mami-Chouaib, *Preprocalcitonin signal peptide generates a cytotoxic t lymphocyte-defined tumor epitope processed by a*

- proteasome-independent pathway*. Proc Natl Acad Sci U S A, 2008. **105**(29): p. 10119-10124.
133. Lemberg, M.K., F.A. Bland, A. Weihofen, V.M. Braud, and B. Martoglio, *Intramembrane proteolysis of signal peptides: An essential step in the generation of hla-e epitopes*. J Immunol, 2001. **167**(11): p. 6441-6446.
 134. Martoglio, B. and T.E. Golde, *Intramembrane-cleaving aspartic proteases and disease: Presenilins, signal peptide peptidase and their homologs*. Hum Mol Genet, 2003. **12 Spec No 2**: p. R201-R206.
 135. Sato, T., A.C. Nyborg, N. Iwata, T.S. Diehl, T.C. Saido, T.E. Golde, and M.S. Wolfe, *Signal peptide peptidase: Biochemical properties and modulation by nonsteroidal antiinflammatory drugs*. Biochemistry, 2006. **45**(28): p. 8649-8656.
 136. Erez, E., D. Fass, and E. Bibi, *How intramembrane proteases bury hydrolytic reactions in the membrane*. Nature, 2009. **459**(7245): p. 371-378.
 137. Narayanan, S., T. Sato, and M.S. Wolfe, *A c-terminal region of signal peptide peptidase defines a functional domain for intramembrane aspartic protease catalysis*. J Biol Chem, 2007. **282**: p. 20172-20179.
 138. Nyborg, A.C., A.Y. Kornilova, K. Jansen, T.B. Ladd, M.S. Wolfe, and T.E. Golde, *Signal peptide peptidase forms a homodimer that is labeled by an active site-directed gamma-secretase inhibitor*. J Biol Chem, 2004. **279**(15): p. 15153-15160.
 139. Weihofen, A., M.K. Lemberg, E. Friedmann, H. Rueeger, A. Schmitz, P. Paganetti, G. Rovelli, and B. Martoglio, *Targeting presenilin-type aspartic protease signal peptide peptidase with gamma-secretase inhibitors*. J Biol Chem, 2003. **278**(19): p. 16528-16533.
 140. Torres-Arancivia, C., C.M. Ross, J. Chavez, Z. Assur, G. Dolios, F. Mancina, and I. Ubarretxena-Belandia, *Identification of an archaeal presenilin-like intramembrane protease*. PLoS One, 2010. **5**(9).
 141. Serrano-Vega, M.J., F. Magnani, Y. Shibata, and C.G. Tate, *Conformational thermostabilization of the β 1-adrenergic receptor in a detergent-resistant form*. Proc Natl Acad Sci U S A, 2008. **105**(3): p. 877-882.
 142. Boulter, J.M. and D.N. Wang, *Purification and characterization of human erythrocyte glucose transporter in decylmaltoside detergent solution*. Protein Expr Purif, 2001. **22**(2): p. 337-348.
 143. Johnson, J.L., S. Kalyoncu, and R.L. Lieberman, *Lessons from an alpha-helical membrane enzyme: Expression, purification, and detergent optimization for biophysical and structural characterization*. Methods Mol Biol (Clifton, N.J.), 2015.

144. Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman, *Basic local alignment search tool*. J Mol Biol, 1990. **215**(3): p. 403-410.
145. Shen, H.B. and K.C. Chou, *Signal-3l: A 3-layer approach for predicting signal peptides*. Biochem Biophys Res Commun, 2007. **363**(2): p. 297-303.
146. Petersen, T.N., S. Brunak, G. von Heijne, and H. Nielsen, *Signalp 4.0: Discriminating signal peptides from transmembrane regions*. Nat Methods, 2011. **8**(10): p. 785-786.
147. Tropea, J., S. Cherry, and D. Waugh, *Expression and purification of soluble his6-tagged tev protease*, in *High throughput protein expression and purification*, S. Doyle, Editor. 2009, Humana Press. p. 297-307.
148. Sambrook, J. and D.W. Russell, *Molecular cloning : A laboratory manual*. 3rd ed. 2001, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.
149. Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, and A. Bairoch, *Protein identification and analysis tools on the expasy server*. Proteomics Prot Handbook, 2005: p. 571-607.
150. Sievers, F., A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, and D.G. Higgins *Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega*. Mol Syst Biol, 2011. **7**, 539
151. Herrmann, K.W., *Non-ionic-cationic micellar properties of dimethyldodecylamine oxide*. J Phys Chem, 1962. **66**(2): p. 295-300.
152. Strop, P. and A.T. Brunger, *Refractive index-based determination of detergent concentration and its application to the study of membrane proteins*. Protein Science, 2005. **14**(8): p. 2207-2211.
153. Malawski, G.A., R.C. Hillig, F. Monteclaro, U. Eberspaecher, A.A.P. Schmitz, K. Crusius, M. Huber, U. Egner, P. Donner, and B. Müller-Tiemann, *Identifying protein construct variants with increased crystallization propensity—a case study*. Protein Sci, 2006. **15**(12): p. 2718-2728.
154. Drew, D., M. Lerch, E. Kunji, D.-J. Slotboom, and J.-W. de Gier, *Optimization of membrane protein overexpression and purification using gfp fusions*. Nature Methods, 2006. **3**(4): p. 303-313.
155. Newstead, S., H. Kim, G. von Heijne, S. Iwata, and D. Drew, *High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in saccharomyces cerevisiae*. Proc Natl Acad Sci U S A, 2007. **104**(35): p. 13936-13941.

156. Zerbs, S., S. Giuliani, and F. Collart, *Chapter eleven - small-scale expression of proteins in e. Coli*, in *Method enzymol*, L. Jon, Editor. 2014, Academic Press. p. 117-131.
157. Raman, P., V. Cherezov, and M. Caffrey, *The membrane protein data bank*. *Cell Mol Life Sci*, 2006. **63**(1): p. 36-51.
158. Kawate, T. and E. Gouaux, *Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins*. *Structure*, 2006. **14**(4): p. 673-681.
159. Tomasiak, T.M., B.P. Pedersen, S. Chaudhary, A. Rodriguez, Y.R. Colmanares, Z. Roe-Zurz, S. Thamminana, M. Tessema, and R.M. Stroud, *General qpcr and plate reader methods for rapid optimization of membrane protein purification and crystallization using thermostability assays*. *Curr Protoc Protein Sci*, 2014. **77**: p. 29.11.1-29.11.14.
160. Sonoda, Y., S. Newstead, N.-J. Hu, Y. Alguel, E. Nji, K. Beis, S. Yashiro, C. Lee, J. Leung, A.D. Cameron, B. Byrne, S. Iwata, and D. Drew, *Benchmarking membrane protein detergent stability for improving throughput of high-resolution x-ray structures*. *Structure*. **19**(1): p. 17-25.
161. Alexandrov, A.I., M. Mileni, E.Y. Chien, M.A. Hanson, and R.C. Stevens, *Microscale fluorescent thermal stability assay for membrane proteins*. *Structure*, 2008. **16**(3): p. 351-359.
162. Engel, C.K., L. Chen, and G.G. Prive, *Stability of the lactose permease in detergent solutions*. *Biochim Biophys Acta*, 2002. **1564**(1): p. 47-56.
163. Lemieux, M.J., R.A. Reithmeier, and D.N. Wang, *Importance of detergent and phospholipid in the crystallization of the human erythrocyte anion-exchanger membrane domain*. *J Struct Biol*, 2002. **137**(3): p. 322-332.
164. Prince, C. and Z. Jia, *Detergent quantification in membrane protein samples and its application to crystallization experiments*. *Amino Acids*, 2013. **45**(6): p. 1293-1302.
165. Veessler, D., S. Blangy, C. Cambillau, and G. Sciara, *There is a baby in the bath water: Acrb contamination is a major problem in membrane-protein crystallization*. *Acta Crystallogr F Struct Biol Cryst Commun*, 2008. **64**(Pt 10): p. 880-885.
166. Pedelacq, J.-D., E. Piltch, E.C. Liong, J. Berendzen, C.-Y. Kim, B.-S. Rho, M.S. Park, T.C. Terwilliger, and G.S. Waldo, *Engineering soluble proteins for structural genomics*. *Nat. Biotechnol.*, 2002. **20**(9): p. 927-932.

167. Yang, J.K., M.S. Park, G.S. Waldo, and S.W. Suh, *Directed evolution approach to a structural genomics project: Rv2002 from mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. USA, 2003. **100**(2): p. 455-460.
168. Keenan, R.J., D.L. Siehl, R. Gorton, and L.A. Castle, *DNA shuffling as a tool for protein crystallization*. Proc. Natl. Acad. Sci USA, 2005. **102**(25): p. 8887-8892.
169. Argos, P., M.G. Rossmann, U.M. Grau, H. Zuber, G. Frank, and J.D. Tratschin, *Thermal stability and protein structure*. Biochemistry, 1979. **18**(25): p. 5698-5703.
170. Vedadi, M., F.H. Niesen, A. Allali-Hassani, O.Y. Fedorov, P.J. Finerty, G.A. Wasney, R. Yeung, C. Arrowsmith, L.J. Ball, H. Berglund, R. Hui, B.D. Marsden, P. Nordlund, M. Sundstrom, J. Weigelt, and A.M. Edwards, *Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination*. Proc Natl Acad Sci USA, 2006. **103**(43): p. 15835-15840.
171. Connelly, P.R., *Acquisition and use of calorimetric data for prediction of the thermodynamics of ligand-binding and folding reactions of proteins*. Current opinion in biotechnology, 1994. **5**(4): p. 381-388.
172. Laganowsky, A., M. Zhao, A.B. Soriaga, M.R. Sawaya, D. Cascio, and T.O. Yeates, *An approach to crystallizing proteins by metal-mediated synthetic symmetrization*. Protein Sci, 2011. **20**(11): p. 1876-1890.
173. Banatao, D.R., D. Cascio, C.S. Crowley, M.R. Fleissner, H.L. Tienson, and T.O. Yeates, *An approach to crystallizing proteins by synthetic symmetrization*. Proc. Natl. Acad. Sci USA, 2006. **103**(44): p. 16230-16235.
174. Derewenda, Z.S., *Rational protein crystallization by mutational surface engineering*. Structure, 2004. **12**(4): p. 529-535.
175. Caffrey, M., *Membrane protein crystallization*. J Struct Biol, 2003. **142**(1): p. 108-132.
176. Prive, G.G., *Lipopeptide detergents for membrane protein studies*. Curr Opin Struct Biol, 2009. **19**(4): p. 379-385.
177. Howell, S.C., N.J. Fraser, R. Mittal, L. Huang, B. Travis, R.M. Breyer, and C.R. Sanders, *Chobimalt: A cholesterol-based detergent*. Biochemistry, 2010. **49**(44): p. 9572-9583.
178. Wang, X., K. Corin, P. Baaske, C.J. Wienken, M. Jerabek-Willemsen, S. Duhr, D. Braun, and S. Zhang, *Peptide surfactants for cell-free production of functional g protein-coupled receptors*. Proc Natl Acad Sci USA, 2011. **108**(22): p. 9049-9054.

179. Cherezov, V., D.M. Rosenbaum, M.A. Hanson, S.G. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka, and R.C. Stevens, *High-resolution crystal structure of an engineered human β 2-adrenergic g protein-coupled receptor*. Science, 2007. **318**(5854): p. 1258-1265.
180. Liu, W., E. Chun, A.A. Thompson, P. Chubukov, F. Xu, V. Katritch, G.W. Han, C.B. Roth, L.H. Heitman, A.P. IJzerman, V. Cherezov, and R.C. Stevens, *Structural basis for allosteric regulation of gpcrs by sodium ions*. Science, 2012. **337**(6091): p. 232-236.
181. Zou, Y., W.I. Weis, and B.K. Kobilka, *N-terminal t4 lysozyme fusion facilitates crystallization of a g protein coupled receptor*. PLoS One, 2012. **7**(10): p. e46039.
182. Kovari, L.C., C. Momany, and M.G. Rossmann, *The use of antibody fragments for crystallization and structure determinations*. Structure, 1995. **3**(12): p. 1291-123.
183. Ostermeier, C., S. Iwata, B. Ludwig, and H. Michel, *Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase*. Nat Struct Biol, 1995. **2**(10): p. 842-846.
184. Zhou, Y., J.H. Morais-Cabral, A. Kaufman, and R. MacKinnon, *Chemistry of ion coordination and hydration revealed by a k^+ channel-fab complex at 2.0 a resolution*. Nature, 2001. **414**(6859): p. 43-48.
185. Stura, E.A., M. Graille, N.G. Housden, and M.G. Gore, *Protein l mutants for the crystallization of antibody fragments*. Acta Crystallogr, 2002. **D58**(Pt 10 Pt 1): p. 1744-178.
186. Rasmussen, S.G., H.J. Choi, D.M. Rosenbaum, T.S. Kobilka, F.S. Thian, P.C. Edwards, M. Burghammer, V.R. Ratnala, R. Sanishvili, R.F. Fischetti, G.F. Schertler, W.I. Weis, and B.K. Kobilka, *Crystal structure of the human β 2 adrenergic g-protein-coupled receptor*. Nature, 2007. **450**(7168): p. 383-387.
187. Tereshko, V., S. Uysal, A. Koide, K. Margalef, S. Koide, and A.A. Kossiakoff, *Toward chaperone-assisted crystallography: Protein engineering enhancement of crystal packing and x-ray phasing capabilities of a camelid single-domain antibody (vhh) scaffold*. Protein Sci, 2008. **17**(7): p. 1175-1187.
188. Uysal, S., V. Vasquez, V. Tereshko, K. Esaki, F.A. Fellouse, S.S. Sidhu, S. Koide, E. Perozo, and A. Kossiakoff, *Crystal structure of full-length kcsa in its closed conformation*. Proceedings of the Proc. Natl. Acad. Sci USA, 2009. **106**(16): p. 6644-6649.
189. Hassaine, G., C. Deluz, L. Grasso, R. Wyss, M.B. Tol, R. Hovius, A. Graff, H. Stahlberg, T. Tomizaki, A. Desmyter, C. Moreau, X.D. Li, F. Poitevin, H. Vogel, and H. Nury, *X-ray structure of the mouse serotonin 5-ht3 receptor*. Nature, 2014. **512**(7514): p. 276-281.

190. Huber, T., D. Steiner, D. Rothlisberger, and A. Plückthun, *In vitro selection and characterization of darpins and fab fragments for the co-crystallization of membrane proteins: The na(+)-citrate symporter cits as an example*. J Struct Biol, 2007. **159**(2): p. 206-221.
191. Sennhauser, G. and M.G. Grutter, *Chaperone-assisted crystallography with darpins*. Structure, 2008. **16**(10): p. 1443-1453.
192. Pai, J.C., J.A. Culver, J.E. Drury, R.S. Motani, R.L. Lieberman, and J.A. Maynard, *Conversion of scfv peptide-binding specificity for crystal chaperone development*. Protein Eng Des Sel, 2011. **24**(5): p. 419-428.
193. Pai, J.C., J.A. Culver, J.E. Drury, R.S. Motani, R.L. Lieberman, and J.A. Maynard, *Conversion of scfv peptide-binding specificity for crystal chaperone development*. Protein Eng Des Sel, 2011. **24**(5): p. 419-28.
194. Kalyoncu, S., J. Hyun, J.C. Pai, J.L. Johnson, K. Entzminger, A. Jain, D.P. Heaner, Jr., I.A. Morales, T.M. Truskett, J.A. Maynard, and R.L. Lieberman, *Effects of protein engineering and rational mutagenesis on crystal lattice of single chain antibody fragments*. Proteins: Struct, Funct, Bioinf, 2014. **82**(9): p. 1884-1895.
195. Koldobskaya, Y., E.M. Duguid, D.M. Shechner, N.B. Suslov, J. Ye, S.S. Sidhu, D.P. Bartel, S. Koide, A.A. Kossiakoff, and J.A. Piccirilli, *A portable rna sequence whose recognition by a synthetic antibody facilitates structural determination*. Nat Struct Mol Biol, 2011. **18**: p. 100-106.
196. Röthlisberger, D., A. Honegger, and A. Plückthun, *Domain interactions in the fab fragment: A comparative evaluation of the single-chain fv and fab format engineered with variable domains of different stability*. J Mol Biol, 2005. **347**: p. 773-789.
197. Levy, R., R. Weiss, G. Chen, B.L. Iverson, and G. Georgiou, *Production of correctly folded fab antibody fragment in the cytoplasm of escherichia coli trxb gor mutants via the coexpression of molecular chaperones*. Protein Expr Purif, 2001. **23**(2): p. 338-347.
198. Maynard, J., E.J. Adams, M. Krogsgaard, K. Petersson, C.W. Liu, and K.C. Garcia, *High-level bacterial secretion of single-chain alphabeta t-cell receptors*. J Immunol Methods, 2005. **306**(1-2): p. 51-67.
199. Minor, W., M. Cymborowski, Z. Otwinowski, and M. Chruszcz, *Hkl-3000: The integration of data reduction and structure solution--from diffraction images to an initial model in minutes*. Acta Crystallogr, 2006. **D62**(Pt 8): p. 859-866.
200. McCoy, A., R. Grosse-Kunstleve, P. Adams, M. Winn, L. Storoni, and R. Read, *Phaser crystallographic software*. J Appl Cryst, 2007. **40**: p. 658-674.

201. Stein, N., *Chainsaw: A program for mutating pdb files used as templates in molecular replacement*. J Appl Cryst, 2008. **41**: p. 641-643.
202. Bourhis, E., W. Wang, C. Tam, J. Hwang, Y. Zhang, D. Spittler, Oscar W. Huang, Y. Gong, A. Estevez, I. Zilberleyb, L. Rouge, C. Chiu, Y. Wu, M. Costa, Rami N. Hannoush, Y. Franke, and Andrea G. Cochran, *Wnt antagonists bind through a short peptide to the first β -propeller domain of lrp5/6*. Structure, 2011. **19**(10): p. 1433-1442.
203. Emsley, P., B. Lohkamp, W. Scott, and K. Cowtan, *Features and development of coot*. Acta Crystallogr, 2010. **D66**: p. 486-501.
204. Adams, P.D., P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, and P.H. Zwart, *Phenix: A comprehensive python-based system for macromolecular structure solution*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.
205. Lovell, S.C., I.W. Davis, W.B. Arendall, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson, and D.C. Richardson, *Structure validation by ϕ , ψ and $c\beta$ deviation*. Proteins: Struct, Funct, Bioinf, 2003. **50**(3): p. 437-450.
206. Krissinel, E. and K. Henrick, *Inference of macromolecular assemblies from crystalline state*. J Mol Biol, 2007. **372**(3): p. 774-797.
207. Link, A.J. and J. LaBaer, *Trichloroacetic acid (tca) precipitation of proteins*. Cold Spring Harb. Protoc., 2011. **2011**(8):
208. Krissinel, E. and K. Henrick, *Secondary-structure matching (ssm), a new tool for fast protein structure alignment in three dimensions*. Acta Crystallogr, 2004. **D60**(12 Part 1): p. 2256-2268.
209. Adams, G.P. and R. Schier, *Generating improved single-chain fv molecules for tumor targeting*. J. Immunol Methods, 1999. **231**: p. 249-260.
210. Maynard, J.A., C.B.M. Maassen, S.H. Leppla, K. Brasky, J.L. Patterson, B.L. Iverson, and G. Georgiou, *Protection against anthrax toxin by recombinant antibody fragments correlates with antigen affinity*. Nat Biotechnol, 2002. **20**: p. 597-601.
211. Krebs, B., R. Rauchenberger, S. Reiffert, C. Rothe, M. Tesar, E. Thomassen, M. Cao, T. Dreier, D. Fischer, A. Höß, L. Inge, A. Knappik, M. Marget, P. Pack, X. Meng, R. Schier, P. Söhlemann, J. Winter, J. Wölle, and T. Kretzschmar, *High-throughput generation and engineering of recombinant human antibodies*. J Immunol Methods, 2001. **254**: p. 67-84.

212. Lim, H.-H., Y. Fang, and C. Williams, *High-efficiency screening of monoclonal antibodies for membrane protein crystallography*. PLoS One, 2011. **6**: p. e24653.
213. Ericsson, U.B., M. Hallberg, G.T. DeTitta, N. Dekker, and P. Nordlund, *Thermofluor-based high-throughput stability optimization of proteins for structural studies*. Anal Biochem, 2006. **357**: p. 289-298.
214. Fairman, J.W., N. Dautin, D. Wojtowicz, W. Liu, N. Noinaj, T.J. Barnard, E. Udho, T.M. Przytycka, V. Cherezov, and S.K. Buchanan, *Crystal structures of the outer membrane domain of intimin and invasins from enterohemorrhagic *E. coli* and enteropathogenic *E. coli* pseudotuberculosis*. Structure, 2012. **20**(7): p. 1233-1243.
215. Costa, G.L. and M.P. Weiner, *Colony pcr*. Cold Spring Harb. Protoc., 2006.
216. O'Malley, M.A., T. Lazarova, Z.T. Britton, and A.S. Robinson, *High-level expression in *Saccharomyces cerevisiae* enables isolation and spectroscopic characterization of functional human adenosine A_{2A} receptor*. J Struct Biol, 2007. **159**: p. 166-178.
217. Cole, C., J.D. Barber, and G.J. Barton, *The jpred 3 secondary structure prediction server*. Nucleic Acids Res, 2008. **36**(suppl 2): p. W197-W201.
218. Buchan, D.W., F. Minneci, T.C. Nugent, K. Bryson, and D.T. Jones, *Scalable web services for the psipred protein analysis workbench*. Nucleic Acids Res, 2013. **41**(Web Server issue): p. W349-W357.
219. van den Ent, F. and J. Löwe, *Rf cloning: A restriction-free method for inserting target genes into plasmids*. J Biochem Biophys Methods, 2006. **67**(1): p. 67-74.
220. Jo, S., T. Kim, V.G. Iyer, and W. Im, *Charmm-gui: A web-based graphical user interface for charmm*. J Comp Chem, 2008. **29**(11): p. 1859-1865.
221. Humphrey, W., A. Dalke, and K. Schulten, *Vmd: Visual molecular dynamics*. J Mol Graphics, 1996. **14**(1): p. 33-38.
222. Phillips, J.C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, and K. Schulten, *Scalable molecular dynamics with namd*. J Comput Chem, 2005. **26**(16): p. 1781-1802.
223. Comeau, S.R., D.W. Gatchell, S. Vajda, and C.J. Camacho, *Cluspro: An automated docking and discrimination method for the prediction of protein complexes*. Bioinformatics, 2004. **20**(1): p. 45-50.
224. Best, R.B., X. Zhu, J. Shim, P.E.M. Lopes, J. Mittal, M. Feig, and A.D. MacKerell, *Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles*. J Chem Theory Comput, 2012. **8**(9): p. 3257-3273.

225. Jorgensen, W.L., J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein, *Comparison of simple potential functions for simulating liquid water*. J Chem Phys, 1983. **79**(2): p. 926-935.
226. Darden, T., D. York, and L. Pedersen, *Particle mesh ewald: An $n \cdot \log(n)$ method for ewald sums in large systems*. J Chem Phys, 1993. **98**(12): p. 10089-10092.
227. Rosenbaum, D.M., S.G. Rasmussen, and B.K. Kobilka, *The structure and function of g-protein-coupled receptors*. Nature, 2009. **459**(7245): p. 356-363.
228. Rosenbaum, D.M., V. Cherezov, M.A. Hanson, S.G.F. Rasmussen, F.S. Thian, T.S. Kobilka, H.-J. Choi, X.-J. Yao, W.I. Weis, R.C. Stevens, and B.K. Kobilka, *Gpcr engineering yields high-resolution structural insights into β 2-adrenergic receptor function*. Science, 2007. **318**(5854): p. 1266-1273.
229. Day, P.W., S.G.F. Rasmussen, C. Parnot, J.J. Fung, A. Masood, T.S. Kobilka, X.-J. Yao, H.-J. Choi, W.I. Weis, D.K. Rohrer, and B.K. Kobilka, *A monoclonal antibody for g protein-coupled receptor crystallography*. Nat Meth, 2007. **4**(11): p. 927-929.
230. Fairman, James W., N. Dautin, D. Wojtowicz, W. Liu, N. Noinaj, Travis J. Barnard, E. Udho, Teresa M. Przytycka, V. Cherezov, and Susan K. Buchanan, *Crystal structures of the outer membrane domain of intimin and invasin from enterohemorrhagic e. Coli and enteropathogenic y. Pseudotuberculosis*. Structure, 2012. **20**(7): p. 1233-1243.
231. Chaiet, L. and F.J. Wolf, *The properties of streptavidin, a biotin-binding protein produced by streptomycetes*. Arch Biochem Biophys, 1964. **106**(0): p. 1-5.
232. Green, N.M., *Avidin*, in *Advances in protein chemistry*, J.T.E. C.B. Anfinsen and M.R. Frederic, Editors. 1975, Academic Press. p. 85-133.
233. Weber, P.C., D.H. Ohlendorf, J.J. Wendoloski, and F.R. Salemme, *Structural origins of high-affinity biotin binding to streptavidin*. Science, 1989. **243**(4887): p. 85-88.
234. Dundas, C.M., D. Demonte, and S. Park, *Streptavidin-biotin technology: Improvements and innovations in chemical and biological applications*. Appl Microbiol Biotechnol, 2013. **97**(21): p. 9343-9353.
235. Chilkoti, A., B.L. Schwartz, R.D. Smith, C.J. Long, and P.S. Stayton, *Engineered chimeric streptavidin tetramers as novel tools for bioseparations and drug delivery*. Biotechnology (N Y), 1995. **13**(11): p. 1198-1204.
236. Howarth, M., D.J. Chinnapen, K. Gerrow, P.C. Dorrestein, M.R. Grandy, N.L. Kelleher, A. El-Husseini, and A.Y. Ting, *A monovalent streptavidin with a single femtomolar biotin binding site*. Nat Methods, 2006. **3**(4): p. 267-273.

237. Nordlund, H.R., V.P. Hytonen, J. Horha, J.A. Maatta, D.J. White, K. Halling, E.J. Porkka, J.P. Slotte, O.H. Laitinen, and M.S. Kulomaa, *Tetravalent single-chain avidin: From subunits to protein domains via circularly permuted avidins*. *Biochem J*, 2005. **392**(Pt 3): p. 485-91.
238. Aslan, F.M., Y. Yu, S.C. Mohr, and C.R. Cantor, *Engineered single-chain dimeric streptavidins with an unexpected strong preference for biotin-4-fluorescein*. *Proc Natl Acad Sci U S A*, 2005. **102**(24): p. 8507-12.
239. Laitinen, O.H., H.R. Nordlund, V.P. Hytonen, S.T. Uotila, A.T. Marttila, J. Savolainen, K.J. Airene, O. Livnah, E.A. Bayer, M. Wilchek, and M.S. Kulomaa, *Rational design of an active avidin monomer*. *J Biol Chem*, 2003. **278**(6): p. 4010-4014.
240. Wu, S.C. and S.L. Wong, *Engineering soluble monomeric streptavidin with reversible biotin binding capability*. *J Biol Chem*, 2005. **280**(24): p. 23225-23231.
241. Chilkoti, A., P.H. Tan, and P.S. Stayton, *Site-directed mutagenesis studies of the high-affinity streptavidin-biotin complex: Contributions of tryptophan residues 79, 108, and 120*. *Proc Natl Acad Sci U S A*, 1995. **92**(5): p. 1754-1758.
242. Sano, T. and C.R. Cantor, *Intersubunit contacts made by tryptophan 120 with biotin are essential for both strong biotin binding and biotin-induced tighter subunit association of streptavidin*. *Proc Natl Acad Sci U S A*, 1995. **92**(8): p. 3180-3184.
243. Helppolainen, S.H., K.P. Nurminen, J.A. Maatta, K.K. Halling, J.P. Slotte, T. Huhtala, T. Liimatainen, S. Yla-Herttuala, K.J. Airene, A. Narvanen, J. Janis, P. Vainiotalo, J. Valjakka, M.S. Kulomaa, and H.R. Nordlund, *Rhizavidin from rhizobium etli: The first natural dimer in the avidin protein family*. *Biochem J*, 2007. **405**(3): p. 397-405.
244. Li, Y. and R. Sousa, *Expression and purification of e. Coli bira biotin ligase for in vitro biotinylation*. *Protein Expr Purif*, 2012. **82**(1): p. 162-167.
245. Ke, J., K.G. Harikumar, C. Erice, C. Chen, X. Gu, L. Wang, N. Parker, Z. Cheng, W. Xu, B.O. Williams, K. Melcher, L.J. Miller, and H.E. Xu, *Structure and function of norrin in assembly and activation of a frizzled 4-lrp5/6 complex*. *Genes & Development*, 2013. **27**(21): p. 2305-2319.
246. Zhi, X., X.E. Zhou, Y. He, C. Zechner, K.M. Suino-Powell, S.A. Kliewer, K. Melcher, D.J. Mangelsdorf, and H.E. Xu, *Structural insights into gene repression by the orphan nuclear receptor shp*. *Proc Natl Acad Sci U S A*, 2014. **111**(2): p. 839-844.
247. Zanier, K., S. Charbonnier, A.O. Sidi, A.G. McEwen, M.G. Ferrario, P. Poussin-Courmontagne, V. Cura, N. Brimer, K.O. Babah, T. Ansari, I. Muller, R.H. Stote, J. Cavarelli, S. Vande Pol, and G. Trave, *Structural basis for hijacking of cellular*

- lxxll motifs by papillomavirus e6 oncoproteins*. Science, 2013. **339**(6120): p. 694-698.
248. Hewitt, S.N., R. Choi, A. Kelley, G.J. Crowther, A.J. Napuli, and W.C. Van Voorhis, *Expression of proteins in escherichia coli as fusions with maltose-binding protein to rescue non-expressed targets in a high-throughput protein-expression and purification pipeline*. Acta Crystallogr F, 2011. **67**(Pt 9): p. 1006-1009.
249. Cole, D.K., P.J. Rizkallah, F. Gao, N.I. Watson, J.M. Boulter, J.I. Bell, M. Sami, G.F. Gao, and B.K. Jakobsen, *Crystal structure of hla-a*2402 complexed with a telomerase peptide*. Eur J Immunol, 2006. **36**(1): p. 170-179.
250. Chappell, P., P. Roversi, M.C. Harrison, L.E. Mears, J.F. Kaufman, and S.M. Lea, *B21 chicken class i mhc molecules and promiscuous peptide binding*. To be published.
251. Shuman, H., M.J. Greenberg, A. Zwolak, T. Lin, C.V. Sindelar, R. Dominguez, and E.M. Ostap, *A vertebrate myosin-i structure reveals unique insights into myosin mechanochemical tuning*. Proc Natl Acad Sci U S A, 2014. **111**(6): p. 2116-2121.
252. Germain, D., *Fabry disease*. Orphanet Journal of Rare Diseases, 2010. **5**(1): p. 30.
253. Spada, M., S. Pagliardini, M. Yasuda, T. Tukel, G. Thiagarajan, H. Sakuraba, A. Ponzzone, and R.J. Desnick, *High incidence of later-onset fabry disease revealed by newborn screening*, Am J Hum Genet, 2006. **79**: p. 31-40
254. Hwu, W.L., Y.H. Chien, N.C. Lee, S.C. Chiang, R. Dobrovolny, A.C. Huang, H.Y. Yeh, M.C. Chao, S.J. Lin, T. Kitagawa, R.J. Desnick, and L.W. Hsu, *Newborn screening for Fabry disease in Taiwan reveals a high incidence of the later-onset FLA mutation c.963+919G>A*, Hum Mutat, 2009. **30**: p. 1397-1405.
255. Guce, A. and S. Garman, *The structure of human α -galactosidase a and implications for fabry disease*, in *Fabry disease*, D. Elstein, G. Altarescu, and M. Beck, Editors. 2010, Springer Netherlands. p. 21-38.
256. Lieberman, R.L., J.A. D'aquino, D. Ringe, and G.A. Petsko, *Effects of pH and iminosugar pharmacological chaperones on lysosomal glycosidase structure and stability*, Biochemistry, 2009. **48**: p. 4816-1427.
257. Guce, A.I., N.E. Clark, E.N. Salgado, D.R. Ivanen, A.A. Kulminskaya, H. Brumer, 3rd, and S.C. Garman, *Catalytic mechanism of human alpha-galactosidase*, J Biol Chem, 2010. **285**(6): p. 3625-3632.
258. Eng, C.M., N. Guffon, W.R. Wilcox, D.P. Germain, P. Lee, S. Waldek, L. Caplan, G.E. Linthorst, and R.J. Desnick, *Safety and efficacy of recombinant*

- human α -galactosidase A-- replacement therapy in Fabry's disease*, N. Eng. J. Med., 2001. **345**: p. 9-16.
259. Schiffmann, R., J.B. Kopp, H.A. Austin, S. Sabnis, D.F. Moore, T. Weibel, J.E. Balow, and R.O. Brady, *Enzyme replacement therapy in Fabry disease: a randomized controlled trial*, JAMA, 2001. **285**: p. 2743-2749.
260. Bénichou, B., S. Goyal, C. Sung, A.M. Norfleet, and F. O'Brien, *A retrospective analysis of the potential impact of IgG antibodies to agalsidase beta on efficacy during enzyme replacement therapy for fabry disease*, Mol Genet Metab, 2009. **96**: p. 4-12.
261. Fan, J.Q., S. Ishii, N. Asano, and Y. Suzuki, *Accelerated transport and maturation of lysosomal alpha-galactosidase A in Fabry lymphoblasts by an enzyme inhibitor*, Nat Med, 1999. **5**: p. 112-115.
262. Wu, X., E. Katz, M.C. Della Valle, K. Mascioli, J.J. Flanagan, J.P. Castelli, R. Schiffmann, P. Boudes, D.J. Lockhart, K.J. Valenzano, and E.R. Benjamin, *A pharmacogenetic approach to identify mutant forms of α -galactosidase A that respond to a pharmacological chaperone for Fabry disease*, Hum Mutat, 2011. **32**: p. 965-977.
263. Khanna, R., R. Soska, Y. Lun, J. Feng, M. Frascella, B. Young, N. Brignol, L. Pellegrino, S.A. Sitaraman, R.J. Desnick, E.R. Benjamin, D.J. Lockhart, and K.J. Valenzano, *The pharmacological chaperone 1-deoxygalactonojirimycin reduces tissue globotriaosylceramide levels in a mouse model of Fabry disease*, Mol Ther, 2010. **18**: p. 23-33.
264. Asano, N., S. Ishii, H. Kizu, K. Ikeda, K. Yasuda, A. Kato, O.R. Martin, and J.Q. Fan, *In vitro inhibition and intracellular enhancement of lysosomal alpha-galactosidase A activity in Fabry lymphoblasts by 1-deoxygalactonojirimycin and its derivatives*, Eur J Biochem, 2000. **267**: p. 4179-4186.
265. Guce, A.I., N.E. Clark, J.J. Rogich, and S.C. Garman, *The molecular basis of pharmacological chaperoning in human alpha-galactosidase*. Chem Biol, 2011. **18**(12): p. 1521-1526.
266. Yu, Y., T. Mena-Barragán, K. Higaki, J.L. Johnson, J.E. Drury, R.L. Lieberman, N. Nakasone, H. Ninomiya, T. Tsukimura, H. Sakuraba, Y. Suzuki, E. Nanba, C.O. Mellet, J.M. García Fernández, and K. Ohno, *Molecular basis of 1-deoxygalactonojirimycin arylthiourea binding to human α -galactosidase a: Pharmacological chaperoning efficacy on fabry disease mutants*. ACS Chemical Biology, 2014. **9**(7): p. 1460-1469.
267. Schitter, G., E. Scheucher, A.J. Steiner, A.E. Stütz, M. Thonhofer, C.A. Tarling, S.G. Withers, J. Wicki, K. Fantur, E. Paschke, D.J. Mahuran, B.A. Rigat, M. Tropak, and T.M. Wrodnigg, *Synthesis of lipophilic 1-deoxygalactonojirimycin derivatives as d-galactosidase inhibitors*. Beilstein J Org Chem, 2010. **6**: p. 21.

268. Kim, D.W., T. Uetsuki, Y. Kaziro, N. Yamaguchi, and S. Sugano, *Use of the human elongation factor 1 alpha promoter as a versatile and efficient expression system*, *Gene*, 1990. **91**: p. 217-223.
269. Higaki, K., L. Li, U. Bahrudin, S. Okuzawa, A. Takamura, K. Yamamoto, K. Adachi, R.C. Paraguison, T. Takai, H. Ikehata, L. Tominaga, I. Hisatome, M. Iida, S. Ogawa, J. Matsuda, H. Ninomiya, Y. Sakakibara, K. Ohno, Y. Suzuki, and E. Nanba, *Chemical chaperone therapy: chaperone effect on mutant enzyme and cellular pathophysiology in β -galactosidase deficiency*, *Hum Mutat*, 2011. **32**: p. 843-852.
270. Aguilar-Moncayo, M., T. Takai, K. Higaki, T. Mena-Barragan, Y. Hirano, K. Yura, L. Li, Y. Yu, H. Ninomiya, M.I. Garcia-Moreno, S. Ishii, Y. Sakakibara, K. Ohno, E. Nanba, C. Ortiz Mellet, J.M. Garcia Fernandez, and Y. Suzuki, *Tuning glycosidase inhibition through aglycone interactions: Pharmacological chaperones for fabry disease and gml gangliosidosis*. *Chem Comm*, 2012. **48**(52): p. 6514-6516.
271. Ishii, S., H.H. Chang, K. Kawasaki, K. Yasuda, H.L. Wu, S.C. Garman, and J.Q. Fan, *Mutant alpha-galactosidase A enzymes identified in Fabry disease patients with residual enzyme activity: biochemical characterization and restoration of normal intracellular processing by 1-deoxygalactonojirimycin*, *Biochem J*, 2007. **406**: p. 285-295.
272. Brumshtein, B., M. Aguilar-Moncayo, J.M. Benito, J.M. Garcia Fernandez, I. Silman, Y. Shaaltiel, D. Aviezer, J.L. Sussman, A.H. Futerman, and C. Ortiz Mellet, *Cyclodextrin-mediated crystallization of acid beta-glucosidase in complex with amphiphilic bicyclic nojirimycin analogues*. *Organic & biomolecular chemistry*, 2011. **9**(11): p. 4160-4167.
273. Otwinowski, Z. and W. Minor, *Processing of x-ray diffraction data collected in oscillation mode*. *Method Enzymol*, 1997. **276**: p. 307-326.
274. Moriarty, N.W., R.W. Grosse-Kunstleve, and P.D. Adams, *Acta Crystallogr D Biol Crystallogr*, 2009. **65**: p. 1074.
275. Schuttelkopf, A.W. and D.M. van Aalten, *Prodrgr: A tool for high-throughput crystallography of protein-ligand complexes*. *Acta Crystallogr D Biol Crystallogr*, 2004. **60**(Pt 8): p. 1355-1363.
276. Takai, T., K. Higaki, M. Aguilar-Moncayo, T. Mena-Barragán, Y. Hirano, K. Yura, L. Yu, H. Ninomiya, M.I. García-Moreno, Y. Sakakibara, K. Ohno, E. Nanba, C. Ortiz Mellet, J.M. García Fernández, and Y. Suzuki, *A bicyclic 1-deoxygalactonojirimycin derivative as a novel pharmacological chaperone for GM1 gangliosidosis*, *Mol Ther*, 2013. **21**: p. 526-532.
277. Butters, T.D., R.A. Dwek, and F.M. Platt, *Glycobiology*, 2005. **15**: p. 42R.

278. Settembre, C., A. Fraldi, D.C. Rubinsztein, and A. Ballabio, *Autophagy*, 2008. **4**: p. 113.
279. Chévrier, M., N. Brakch, L. Céline, D. Genty, Y. Ramdani, S. Moll, M. Djavaheri-Mergny, C. Brasse-Lagnel, A.L. Annie Laquerrière, F. Barbey, and S. Bekri, *Autophagosome maturation is impaired in Fabry disease*, *Autophagy*, 2010. **6**: p. 589-599.
280. Dixon, R.A., B.K. Kobilka, D.J. Strader, J.L. Benovic, H.G. Dohlman, T. Frielle, M.A. Bolanowski, C.D. Bennett, E. Rands, R.E. Diehl, R.A. Mumford, E.E. Slater, I.S. Sigal, M.G. Caron, R.J. Lefkowitz, and C.D. Strader, *Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin*. *Nature*, 1986. **321**(6065): p. 75-79.
281. Kefala, G., C. Ahn, M. Krupa, L. Esquivies, I. Maslennikov, W. Kwiatkowski, and S. Choe, *Structures of the ompf porin crystallized in the presence of foscholine-12*. *Protein Sci*, 2010. **19**(5): p. 1117-1125.

VITA

Jennifer was born in Snellville, Georgia and grew up just a few miles east in Loganville. She graduated in 2006 from Loganville High School where she served as Senior Class Treasurer and varsity cheerleader. Afterwards, she attended Valdosta State University (Go Blazers!) where she studied chemistry, participated in Student Affiliates of the American Chemical Society (SAACS), was a member of the coed (2006-2007) and all-girl (2007-2009) cheerleading squads, and served on the Student Athlete Advisory Committee (SACS). She then went on to pursue a doctorate in Biochemistry. Her graduate research was funded by grants from the NIH (DK091357 to RLL), NSF (0845445 to RLL), the Department of Education Graduate Assistance in Areas of National Need (GAANN, P200A090307 to JLJ), and Georgia Tech Senior Biophysics Training award (to JLJ).