UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Biochemical and functional characterization of MISTIC

Permalink

https://escholarship.org/uc/item/5b594287

Author

Lundberg, Matthew Evan

Publication Date

2013

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Biochemical and Functional Characterization of MISTIC

A dissertation submitted in partial satisfaction of the

requirements for the degree of Doctor of Philosophy

in

Biology

by

Matthew Evan Lundberg

Committee in charge:

Professor Senyon Choe, Chair
Professor Kit J. Pogliano, Co-chair
Professor Pieter C. Dorrestein
Professor Elizabeth A. Komives
Professor Milton H. Saier

2013

Copyright

Matthew Evan Lundberg, 2013

All rights reserved

The dissertation of Matthew Evan Lundberg is approved, and it is acceptable in quality and form for publication on microfilm:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2013

DEDICATION

To my mother, for making all of this possible.

TABLE OF CONTENTS

Signature page iii

Dedication iv

Table of Contents v

List of Abbreviations viii

Lists of Figures ix

Lists of Tables xi

Acknowledgements xii

Vita xiv

Abstract of the Dissertation xvi

Chapter I Introduction 1

A. Membrane protein biogenesis in bacteria 2

B. MISTIC and membrane protein expression 5

C. Biochemical studies of Mistic 7

D. Functional Characterization of MstX in *Bacillus subtilis* 8

E. References 9

Chapter II Mistic: Cellular Localization, Solution Behavior, Polymerization, and Fibril Formation 15

A. Abstract 16

B. Introduction 16

C. Results and Discussion 18

1. Mistic Orthologs express in the cytoplasm 18

2. The Shorter Mistic Proteins are Soluble Multimers 19

3. Crystals of M2 20

4. Fibril Formation by M2 21

D. Materials and Methods 23

1. Cloning and Purification 24
2. Light Scattering 25
3. Crystallization 25
4. X-Ray Data Analysis 26
5. Electron Microscopy 26
6. Thioflavin T Binding 26
7. Circular Dichroism Spectroscopy 27
E. References 27

Chapter III MstX and a putative potassium channel facilitate biofilm formation in *Bacillus subtilis* 40
A. Abstract and Introduction 41
B. Methods and Materials 44
1. Strains, media, and culture conditions 44
2. Biofilm growth and crystal violet assay 45
3. Chromatin Immunoprecipitation 46
4. RT-PCR and quantitative RT-PCR analysis 48
C. Results 49
1. MstX and YugO promote biofilm formation 49
2. A *sinR* mutation restores biofilm formation 52
3. The *mstX-yugO* operon is expressed during biofilm assembly 54
4. SinR directly regulates the *mstX-yugO* promoter 55
5. Potassium and a kinC mutant abrogate biofilm formation in a MstX-YugO expression strain 56
D. Discussion 58
E. References 64

Chapter IV Mistic-tagged expression and purification of Claudin family members 82
A. Introduction and Background 83
1. Claudin Structure and Function 85
2. Claudins and their contribution to inherited human disease 89
3. Claudins as potential chemotherapeutic targets 90

4. Method for expression, purification, and crystallization of claudin homologues.....92

B. Results and Discussion.....94

1. Overexpression of claudin homologues in *E. coli*.....94

2. Claudin homologues are variably solubilized by multiple detergents.....95

3. Claudin expression in a cell-free expression system.....96

C. References.....98

Chapter V Conclusions and Perspectives.....111

A. Future Directions.....112

1. Directed Evolution of Mistic.....112

2. Potassium and Osmotic Pressure in the regulation of *B. subtilis* biofilm formation..116

B. References.....118

Appendix.....123

A. Results.....124

LIST OF ABBREVIATIONS

β ME	Beta-mercaptoethanol
$C_{12}E_8$	Octaethylene glycol monododecyl ether
D	Dalton
DM	Decyl-beta-D-maltopyranoside
DDM	Dodecyl-beta-D-maltopyranoside
FPLC	Fast Protein Liquid Chromatography
GFP	Green Fluorescent Protein
IM	Integral Membrane
IMP	Integral Membrane Protein
NG	Nonyl-beta-D-glucopyranoside
NMR	Nuclear Magnetic Resonance
OG	Octyl-beta-D-glucopyranoside
SDS	Sodium dodecyl sulfate

LIST OF FIGURES

Figure 1.1 Diagram of the genetic organization of the chromosomal regions surrounding *mstX* and *yugO* 14

Figure 2.1 Sequence alignment of Mistic proteins from *Bacillus* species 33

Figure 2.2 Target expression of *wt* and mutant Mistic proteins 34

Figure 2.3 Oligomeric state measured by light scattering 35

Figure 2.4 Stereographic projection of the self-rotation function 36

Figure 2.5 EM micrographs showing fibril structures of M2 37

Figure 2.6 CD spectra of Mistic 38

Figure 3.1 Diagram of the genetic organization of the chromosomal regions surrounding *mstX* and *yugO* 72

Figure 3.2 Alterations in colony morphology and biofilm formation related to *mstX* expression in the domesticated strain PY79 73

Figure 3.3 The *mstX* and *yugO* genes regulate biofilm formation in undomesticated *B. subtilis* 74

Figure 3.4 MstX negatively regulates parallel antirepressors involved in biofilm formation 75

Figure 3.5 Temporal expression of *mstX* and regulation by the transcriptional repressor SinR 76

Figure 3.6 Potassium chloride or a *kinC* deletion abrogate pellicle formation in a *mstX* expression strain 77

Figure 3.7 Diagram of Proposed Autoregulatory Loop 78

Figure 3.8 Growth curve of *B. subtilis* wild-type and mutant strains 79

Figure 3.9 Growth curve of *B. subtilis* in potassium chloride 80

Figure 4.1 Expression and Ni-NTA purification Claudin proteins 104

Figure 4.2 Detergent extraction test of Claudin-10 with DDM, DM, NG, OG, LDAO, and $C_{12}E_8$ 105

Figure 4.3 Claudin-10 detergent extraction measuered by GFP fluorescence 106

Figure 4.4 Size-exclusion chromatogram of Claudin-10 107

Figure 4.5 Workflow for cloning, cell-free expression, and screening of Claudin proteins for structural studies.....108

Figure 4.6 Cell-free expression and purification of Claudin family members 109

Figure A.1 Mistic compartmentalization assay using a c-terminal GFP tag.....128

Figure A.2 SDS-PAGE and Coomassie stain of unpurified and purified extracts of a strain expressing Mistic-GFP-6xhis.....129

LIST OF TABLES

Table 2.1 X-ray diffraction data..... 31

Table 3.1 Supplementary Strain List..... 69

Table 3.2 Supplementary Primer List 70

ACKNOWLEDGEMENTS

I would like to thank my advisors, Senyon Choe and Kit Pogliano, for their patience and generous support. My current advisor and co-chair, Kit Pogliano, is responsible for helping me become the scientist I am today. She runs an open and helpful lab and encourages her students to interact with each other and with her. I also appreciate all my committee members for their support and for their continued confidence in my research abilities, in spite of the obstacles encountered.

I want to thank my fellow and former lab colleagues. First, Eric C. Becker and Dan H. Broder, who taught me everything about *Bacillus subtilis* in the Pogliano Lab. I owe a significant amount of my early success to Kent A. Baker, Mario Meng-Chiang Kuo, Christian Klammt, and Scott D. Pegan, primarily for their many helpful discussions and technical assistance in the Choe Lab. It was unexpected but greatly appreciated. Thanks to Mingming Liu, for her perennial support behind the scenes. Finally, Alan Derman, Joe Pogliano, Anne Lamsa, Ranmali Perera, Julieta Aguilar, Diana Quach, Javier Lopez Garrido, Casey Johnson, Chris Dickson, , Christian Klammt, Hay Dvir, and Samir Maji made the lab a pleasant place to work and helped provide for my intellectual growth.

Chapter II, in full, is a reprint that the dissertation author significantly contributed to as both a researcher and an author. The material appears in *Protein Sci*.

Dvir, D., **Lundberg, M.E.,** Maji, S.K., Riek, R., and Choe, S. (2009). Mistic: Cellular localization, solution behavior, polymerization, and fibril formation. *Protein Sci,* **18**, 1564-1570.

Chapter III, in full, is a manuscript under revision that the dissertation author was the principal researcher and author of. The material is in review for *PLOS ONE*.

Lundberg, M.E., Becker, E.C., and Choe, S. (2013). MstX and a putative potassium channel facilitate biofilm formation in *Bacillus subtilis*. *PLOS ONE, in revision*.

Chapter IV, in part, was a collaboration with Christian Klammt involving cell-free expression.

VITA

2005 Bachelors of Arts, Biological Chemistry, Grinnell College

2013 Doctor of Philosophy, Biology, University of California, San Diego

PUBLICATIONS

Lundberg, M.E., Becker, E.C., and Choe, S. (2013). MstX and a putative potassium channel facilitate biofilm formation in *Bacillus subtilis*. *PLOS ONE, in revision*.

Dvir, D., **Lundberg, M.E.,** Maji, S.K., Riek, R., and Choe, S. (2009). Mistic: Cellular localization, solution behavior, polymerization, and fibril formation. *Protein Sci,* **18**, 1564-1570.

FIELDS OF STUDY

Major Field: Biochemistry

Studies in Membrane Protein Expression and Purification

Professor Senyon Choe

The Salk Institute for Biological Studies

Major Field: Microbiology:

Studies in *Bacillus subtilis* biofilm formation

Professor Kit Pogliano

University of California, San Diego

HONORS AND AWARDS

2007 Cell and Molecular Genetics Training Grant, trainee

Chapman Foundation Grant

2007 TA Excellence in Teaching Award, UCSD

ABSTRACT OF THE DISSERTATION

Biochemical and Functional Characterization of MISTIC

by

Matthew Evan Lundberg

Doctor of Philosophy in Biology

University of California, San Diego, 2013

Professor Senyon Choe, Chair

Professor Kit Pogliano, Co-chair

Membrane proteins are essential for basic biological functions and are implicated in many disease processes. Organisms have evolved a plethora of pathways for proper targeting and insertion of membrane proteins. Mistic is a recently identified protein in *Bacillus subtilis* that has been shown to facilitate the expression and targeting of integral membrane proteins. It has been proposed that Mistic autonomously targets fused proteins to the membrane, thereby bypassing the canonical secretory apparatus. However, little is known about how Mistic promotes integral membrane targeting to the membrane or its native function in *Bacillus subtilis*. This thesis investigates the *in vitro* biochemical properties and the *in vivo* function of MISTIC in *Bacillus subtilis*.

Chapter II investigates the solution behavior, polymerization, and fibril formation of Mistic and related homologs from other *Bacillus* species. We show that *Bacillus* Mistic homologs exist as multimeric, α-helical oligomers in solution when they are purified in the absence of detergent. In addition, the M2 homolog appears to form multimers in a crystalline array. Upon further characterization, Mistic homolog multimers polymerize and form fibrils Mutations that disrupted the membrane-chaperoning properties of Mistic homologs also abrogated polymerization and fibril formation. We propose that Mistic polymerization and fibril formation contributes to the membrane protein expression properties of Mistic.

Chapter III explores the function of Mistic in *Bacillus subtilis* and its physiological role in bacterial development. We identified Mistic and an associated potassium channel, YugO, as novel regulators of biofilm formation in *Bacillus subtilis*. The *mstX* gene is regulated by the master regulator for biofilm formation, SinR. Mistic and YugO functionality depended upon extracellular potassium levels and a potassium-sensing histidine kinase. We propose that Mistic serves in a positive autoregulatory loop controlling the initial stages of biofilm assembly through potassium leakage.

Chapter IV presents data from *Escherichia coli* and cell-free expression trials of Mistic fused to Claudin family proteins. Claudins form the tight junction that determine the paracellular permeability properties of cells. Little is known about how these permeability properties are maintained or regulated. We demonstrate that Mistic fusion constructs enable the expression and purification of Claudin proteins in

quantities sufficient for structural characterization by solution NMR or x-ray crystallography.

CHAPTER I

Introduction

A. Membrane protein biogenesis in bacteria

Membrane proteins are ubiquitous in eukaryotic and prokaryotic cells, constituting approximately 20% to 30% of total proteins encoded in the genome [1, 2]. Membrane protein targeting, folding, and assembly are critical processes for all life. Prokaryotes have evolved several pathways to enable transfer of proteins across the inner membrane, periplasm, and outer membrane. The primary route for protein translocation in and across the cytoplasmic membrane is the general secretion pathway (Sec-pathway), with the alternative YidC system facilitating proper folding and membrane insertion of some proteins [2-4]. Membrane protein biogenesis requires both coordinated targeting to the membrane and proper folding for assembly into their protein complexes in order to function properly in the cell [5]. These processes collectively serve to produce the membrane proteins that enable cells to divide, convert energy, selectively transport molecules, signal, and move [4].

The prevailing paradigm for integral membrane protein (IMP) biogenesis can be divided into three distinct steps. First, the signal-recognition particle (SRP) interacts with hydrophobic sequences in the nascent polypeptide chain and mediates cotranslational targeting[6]. Second, the nascent protein is transferred to the Sec-translocon, a membrane-embedded protein channel that inserts or transports other nascent proteins or peptides across the phospholipid bilayer of the inner membrane. Third, the nascent chain moves laterally from the Sec-translocon channel into the lipid bilayer where it folds into its native conformation and often forms multimeric complexes [7]. This intricate pathway serves to address the challenge of ribosomal

protein synthesis in the cytosol, an environment ill-suited for proper folding and function of the long, hydrophobic transmembrane segments inherent to all IMPs

The SRP is a universal ribonucleoprotein found in all domains of life and is essential for expression of most IMPs in bacteria [8, 9]. The mammalian SRP comprises six protein subunits and an 7SL RNA scaffold, whereas the *E. coli* SRP contains only one protein subunit, Ffh, and a 4.5S RNA component. Although eukaryotic and prokaryotic SRP homologues vary in complexity and some functions, all serve to bind a conserved targeting signal of an IMP polypeptide chain as it emerges from the exit tunnel in a ribosome, arresting elongation and chaperoning the nascent IMP to the membrane-associated SRP receptor (SR or FtsY in *E. coli*) and membrane-embedded Sec-translocon channel [9-11]. By localizing at the exit site of the ribosome, SRP and the Sec-pathway serve to prevent the aggregation of hydrophobic portions of IMPs in the cytoplasm.

SRP-mediated protein targeting is regulated by GTP binding and hydrolysis through SRP and SR [9, 12]. SRP binds to the targeting signal in a GTP-free state as it emerges from the ribosome [13]. Upon binding the signal sequence, conformational changes increase the GTP-binding affinity of the Ffh protein subdomain of SRP. Ffh bound to GTP and a targeting signal at the ribosome acts as a prime substrate for interactions with SR. After association with SR, the GTPases of SRP and SR become active and hydrolyze GTP after the targeting signal has been released by SRP. GTP-free SRP dissociates from the SRP/SR complex as it loses structural affinity for SR in the GTP-free state [8, 14]. The release of SRP enables the continued translation of the

ribosome nascent chain complex into the Sec-translocon where it is translocated across or inserted into the membrane.

Unlike the mammalian SRP, which serves to target secretory and membrane proteins to the ER, the *E. coli* SRP is predominantly involved in the targeting of IMPs [15]. This substrate selectivity is attributed to the preference of the SRP protein subunit Ffh for hydrophobic signals in nascent IMPs transmembrane sequences. Secretory proteins can be routed to the SRP pathway solely through increasing the hydrophobicity of leading sequences or by creating structural affinity in the leading sequence for the N-terminal region of SRP [6, 16, 17]. Individual proteins vary in their affinity for SRP, which therefore is able to more broadly participate in the translation and export of non-IMPs such as the periplasmic protein DsbA, the secreted protein SecM, or the lipoprotein bacteriocin.

The presence of a functional SRP pathway is essential for normal growth in most bacteria. However, it has become apparent that bacteria differ in their dependence on SRP-mediated targeting. Although the SRP subcomponents Ffh and 4.5S RNA are not abundant in the cell, *E. coli* cells are able to cope with a reduction in the existing availability of SRP with minimal impairment, implying the existence of redundant pathways or a natural affinity of the ribosome and nascent IMPs to the Sec-translocon apparatus [4, 18]. In *Bacillus subtilis*, analysis of the extracellular proteome under conditional SRP expression reveals that most secreted proteins rely upon SRP for export [19]. In contrast, the bacteria *Streptococcus mutans* and

Streptococcus pyogenes remain viable during in vitro growth even when Ffh is rendered non-functional [20, 21]. Thus, SRP is not essential in all bacteria.

Recent evidence appears to contradict the widely held conception that SRP is necessary for targeting of the ribosome-nascent chain to the membrane and for translation of IMP. First,depletion of Ffh affects the assembly of membrane proteins but not their membrane targeting or expression *in vivo*; instead, ribosomal localization to the membrane and membrane protein expression appears to depend upon the expression of the SRP-receptor FtsY [22, 23]. *E. coli* cells contain ten times more FtsY SR than Ffh molecules, a feature that belies the role of FtsY as merely an ancillary player in targeting. Although SRP has been convincingly demonstrated to target ribosomes to the membrane *in vitro*, it is challenging to reconcile these recent *in vivo* observations with prevailing models of SRP-mediated ribosomal docking. An intriguing recent paper suggests that ribosomal association to the membrane occurs independently of SRP through pretranslational binding of the uracil-rich mRNAs that encode IMPs to FtsY [24, 25]. Once the original mRNA is translated and released, FtsY could remain docked to membrane with a ribosome, where it could then promote the recruitment of SRP and translation of other IMPs [26]. The potential for mRNA localization and SRP-independent targeting of ribosome-nascent chain complexes opens up new possibilities for IMP translation mechanisms.

B. MISTIC and membrane protein expression

In spite of the importance of eukaryotic integral membrane proteins in many biological and disease processes, our understanding of their structures and functions has traditionally lagged that of soluble proteins [27]. This deficiency is largely attributable to experimental bottlenecks in producing sufficient quantities of protein and analyzing them in functional form. Recombinant expression of eukaryotic membrane proteins in *Escherichia coli* presents major limitations in that prokaryotes lack the cellular translocon machinery necessary for proper membrane integration of some IM proteins; also, high-level expression can ultimately prove toxic by overloading the cellular machinery, competitively excluding the production of vital host membrane proteins [28]. Fusion partner proteins that serve to facilitate expression often fail to target the construct to the membrane, thus preventing structural study of its membrane bound conformation. Furthermore, there are few replacements for *Escherichia coli* in high-level expression of membrane proteins: host systems based on SF9 insect? (what are these) cells, yeast, or mammalian cell lines prove cumbersome in that they are difficult to genetically manipulate or transfect and they usually generate low yields of the desired membrane protein. These technical obstacles are of seminal importance, effectively hampering studies that might otherwise elucidate fundamental cellular processes and provide insight into the function of these membrane proteins that are often desirable drug targets.

A new method emerged to address this problem with the discovery of Mistic. Mistic and its homologues are unique proteins found in a few *Bacillus* species that enable high-level expression of properly folded integral membrane proteins when

fused to the n-terminus of a desired IM construct [29, 30]. The mechanism for how this event is achieved remains unclear. Mistic family members are small, negatively-charged, hydrophilic proteins containing 84 moderately conserved residues sufficient for membrane targeting [30-34]. It has been proposed that Mistic is able to autonomously integrate other membrane proteins, in a Sec-independent manner, into the lipid bilayer. This proposal stems from the absence of a stretch of hydrophobic amino acids that could be interpreted as a signal sequence by the bacterial translocon machinery. Moreover, high level expression of IMs via Mistic fusion proteins does not induce the cellular toxicity normally associated with saturation of the secretory apparatus. Mutational studies revealed that some conserved residues are essential for Mistic's chaperoning properties, thus implying that the structure plays an important role [29, 31]. Other researchers have shown that Mistic expression and folding are dependent upon lipid composition, a surprising discovery in light of Mistic's highly negatively charged surface, which one would expect to be repelled from the membrane [35]. These results suggest that the hydrophilic surface of Mistic must achieve substantial conformational transitions in order to associate with the membrane.

C. Biochemical studies of MISTIC polymerization and solution behavior

In order to investigate the underlying biochemical properties that enable Mistic to promote expression of associated IMP constructs, we studied Mistic homologs found in other *Bacillus* species that have been shown to retain the membrane

chaperoning properties of the originally identified *Bacillus subtilis* protein (Chapter 2). These Mistic orthologs from *B. licheniformis* (M2), *B. mojavensis* (M3), and *B. atrophaeus* (M4) lack the first 26 amino acid residues corresponding to the first α-helix of the solution NMR structure. An additional advantage of working with these orthologs is that they are significantly more soluble in the absence of detergent yet retain the functionality of the *B. subtilis*. We were able to obtain diffracting crystals from the more soluble M2 variant that provide insight into Mistic's oligomerization properties. Moreover, during our purification trials of the M2 otholog, we observed α-helical fibril formation at high concentrations (>20 mg/ml). We analyzed these fibrils both by EM and CD spectroscopy. The introduction of mutations that are known to abrogate the chaperoning properties of Mistic also disrupted the ability of the M2 ortholog to form fibrils. Our finding (described in reference and Chapter 2) uncovers a potential mechanistic role for oligomerization in the enhanced expression of IMP by Mistic.

D. Functional Characterization of MstX in *Bacillus subtilis*

To further elucidate the mechanistic underpinnings of Mistic, we decided to functionally characterize Mistic (*mstX*) in the native Gram-positive bacterium, *Bacillus subtilis*. Prior to my studies in *Bacillus*, all experiments with Mistic were conducted in the heterologous Gram-negative host *Escherichia coli*, a poorly suited platform for studying the natural function of MstX in bacterial physiology. The MstX homologues are limited to *Bacillus atrophaeus*, *Bacillus mojavensis*, *Bacillus*

licheniformis, and a few closely related *Bacilli* species. In all cases, the *mstX* genes precede a putative potassium ion channel *yugO*, suggesting that MstX might be involved in targeting YugO (Figure 1). The MstX protein does not align with any protein of known function, raising the question as to what function MstX might serve in *Bacillus subtilis*.

My studies identified MstX as a novel regulator of biofilm formation in *Bacillus subtilis*. The *mstX* promoter is regulated in part by SinR, the master regulator for biofilm formation. The function of MstX depended upon the presence of the downstream potassium ion channel YugO. Supplementation of media with potassium, disruption of YugO, or disruption of the potassium-sensing kinase KinC abrogated MstX-mediated biofilm formation. We proposed that MstX operates through a potassium efflux-driven positive autoregulatory loop that enhances biofilm formation in *B. subtilis*.

E. References

1. Facey, S.J. and A. Kuhn, *Biogenesis of bacterial inner-membrane proteins.* Cell Mol Life Sci, 2010. **67**(14): p. 2343-62.

2. Egea, P.F., R.M. Stroud, and P. Walter, *Targeting proteins to membranes: structure of the signal recognition particle.* Curr Opin Struct Biol, 2005. **15**(2): p. 213-20.

3. Kuhn, A., *From the Sec complex to the membrane insertase YidC.* Biol Chem, 2009. **390**(8): p. 701-6.

4. Luirink, J., et al., *Biogenesis of inner membrane proteins in Escherichia coli.* Annu Rev Microbiol, 2005. **59**: p. 329-55.

5. von Heijne, G., *Introduction to theme "membrane protein folding and insertion".* Annu Rev Biochem, 2011. **80**: p. 157-60.

6. Valent, Q.A., et al., *The Escherichia coli SRP and SecB targeting pathways converge at the translocon.* EMBO J, 1998. **17**(9): p. 2504-12.

7. Dalal, K. and F. Duong, *The SecY complex: conducting the orchestra of protein translocation.* Trends Cell Biol, 2011. **21**(9): p. 506-14.

8. Keenan, R.J., et al., *The signal recognition particle.* Annu Rev Biochem, 2001. **70**: p. 755-75.

9. Luirink, J. and I. Sinning, *SRP-mediated protein targeting: structure and function revisited.* Biochim Biophys Acta, 2004. **1694**(1-3): p. 17-35.

10. Herskovits, A.A., E.S. Bochkareva, and E. Bibi, *New prospects in studying the bacterial signal recognition particle pathway.* Mol Microbiol, 2000. **38**(5): p. 927-39.

11. Koch, H.G., M. Moser, and M. Muller, *Signal recognition particle-dependent protein targeting, universal to all kingdoms of life.* Rev Physiol Biochem Pharmacol, 2003. **146**: p. 55-94.

12. Bacher, G., et al., *Regulation by the ribosome of the GTPase of the signal-recognition particle during protein targeting.* Nature, 1996. **381**(6579): p. 248-51.

13. Miller, J.D., H.D. Bernstein, and P. Walter, *Interaction of E. coli Ffh/4.5S ribonucleoprotein and FtsY mimics that of mammalian signal recognition particle and its receptor.* Nature, 1994. **367**(6464): p. 657-9.

14. Connolly, T., P. Rapiejko, and R. Gilmore, *Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor.* Science, 1991. **252**(5009): p. 1171-3.

15. de Gier, J.W. and J. Luirink, *Biogenesis of inner membrane proteins in Escherichia coli.* Mol Microbiol, 2001. **40**(2): p. 314-22.

16. Beck, K., et al., *Discrimination between SRP- and SecA/SecB-dependent substrates involves selective recognition of nascent chains by SRP and trigger factor.* EMBO J, 2000. **19**(1): p. 134-43.

17. Valent, Q.A., et al., *Nascent membrane and presecretory proteins synthesized in Escherichia coli associate with signal recognition particle and trigger factor.* Mol Microbiol, 1997. **25**(1): p. 53-64.

18. Bernstein, H.D. and J.B. Hyndman, *Physiological basis for conservation of the signal recognition particle targeting pathway in Escherichia coli.* J Bacteriol, 2001. **183**(7): p. 2187-97.

19. Hirose, I., et al., *Proteome analysis of Bacillus subtilis extracellular proteins: a two-dimensional protein electrophoretic study.* Microbiology, 2000. **146 (Pt 1)**: p. 65-75.

20. Crowley, P.J., et al., *An ffh mutant of Streptococcus mutans is viable and able to physiologically adapt to low pH in continuous culture.* FEMS Microbiol Lett, 2004. **234**(2): p. 315-24.

21. Hasona, A., et al., *Streptococcal viability and diminished stress tolerance in mutants lacking the signal recognition particle pathway or YidC2.* Proc Natl Acad Sci U S A, 2005. **102**(48): p. 17466-71.

22. Herskovits, A.A., et al., *Accumulation of endoplasmic membranes and novel membrane-bound ribosome-signal recognition particle receptor complexes in Escherichia coli.* J Cell Biol, 2002. **159**(3): p. 403-10.

23. Herskovits, A.A. and E. Bibi, *Association of Escherichia coli ribosomes with the inner membrane requires the signal recognition particle receptor but is independent of the signal recognition particle.* Proc Natl Acad Sci U S A, 2000. **97**(9): p. 4621-6.

24. Nevo-Dinur, K., et al., *Translation-independent localization of mRNA in E. coli.* Science, 2011. **331**(6020): p. 1081-4.

25. Prilusky, J. and E. Bibi, *Studying membrane proteins through the eyes of the genetic code revealed a strong uracil bias in their coding mRNAs.* Proc Natl Acad Sci U S A, 2009. **106**(16): p. 6662-6.

26. Bibi, E., *Is there a twist in the Escherichia coli signal recognition particle pathway?* Trends Biochem Sci, 2012. **37**(1): p. 1-6.

27. Loll, P.J., *Membrane protein structural biology: the high throughput challenge.* J Struct Biol, 2003. **142**(1): p. 144-53.

28. Laage, R. and D. Langosch, *Strategies for prokaryotic expression of eukaryotic membrane proteins.* Traffic, 2001. **2**(2): p. 99-104.

29. Roosild, T.P., et al., *NMR structure of Mistic, a membrane-integrating protein for membrane protein expression.* Science, 2005. **307**(5713): p. 1317-21.

30. Roosild, T.P., et al., *Characterization of the family of Mistic homologues.* BMC Struct Biol, 2006. **6**: p. 10.

31. Dvir, H., et al., *Mistic: cellular localization, solution behavior, polymerization, and fibril formation.* Protein Sci, 2009. **18**(7): p. 1564-70.

32. Dvir, H. and S. Choe, *Bacterial expression of a eukaryotic membrane protein in fusion to various Mistic orthologs.* Protein Expr Purif, 2009. **68**(1): p. 28-33.

33. Blain, K.Y., W. Kwiatkowski, and S. Choe, *The functionally active Mistic-fused histidine kinase receptor, EnvZ.* Biochemistry, 2010. **49**(42): p. 9089-95.

34. Kefala, G., et al., *Application of Mistic to improving the expression and membrane integration of histidine kinase receptors from Escherichia coli.* J Struct Funct Genomics, 2007. **8**(4): p. 167-72.

35. Debnath, D.K., et al., *The role of membrane properties in Mistic folding and dimerisation.* Protein Eng Des Sel, 2011. **24**(1-2): p. 89-97.

FIGURES

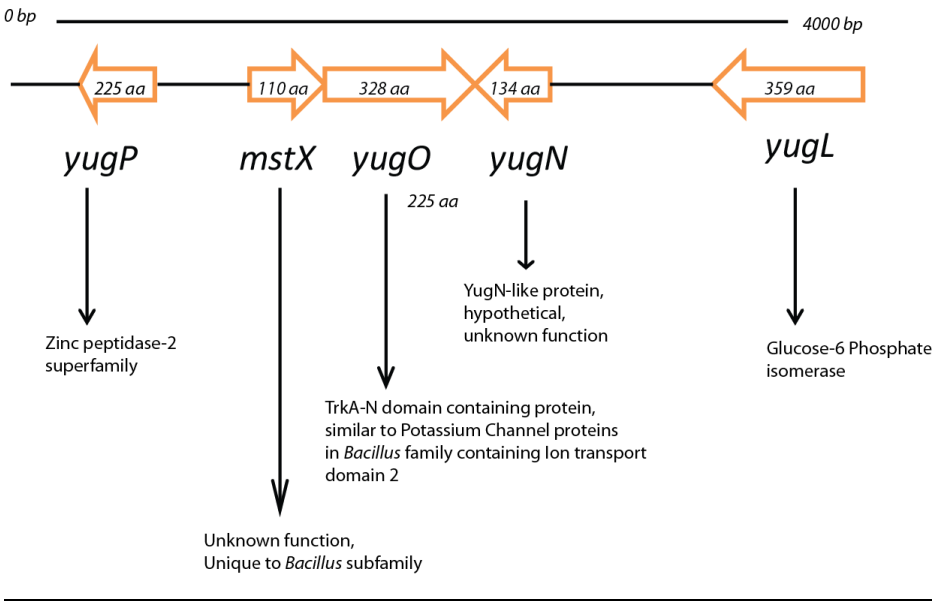


Figure 1.1 Diagram of the genetic organization of the chromosomal regions surrounding *mstX* and *yugO*. Sequence homology was identified by BLAST.

CHAPTER II

Mistic: Cellular Localization, Solution Behavior, Polymerization, and Fibril Formation

A. Abstract

Mistic represents a family of unique membrane-associating proteins originally found in *Bacillus subtilis* (M110). As a fusion partner, it has been shown to assist overexpression of foreign integral membrane proteins in *E. coli.* We have expressed shorter Mistic homologs from other *Bacillus* species and surprisingly, unlike M110, found them abundant in the cytoplasm. These Mistic homologs including the corresponding shorter sequence (amino acids 27 through 110 of M110) exit as multimeric assemblies in solution in the absence of detergent. Crystals of Mistic from *B. leicheniformis* (M2) diffracted to 3.2 Å resolution, indicating that it exists as a multimer in the crystalline state as well. Moreover, we show that although M2 is mostly α-helical, it tends to polymerize and form fibrils. Such oligomerization could potentially mask the charged surface of the monomeric Mistic to assist membrane integration.

B. Introduction

Integral membrane proteins (IMPs) constitute up to 30% of all genes encoded by sequenced genomes [1, 2]. They mediate various activities such as nerve impulses, hormone action or allow pathogens to sense their environment, and elicit a virulence response, all of which makes them the largest groups of drug target proteins [3]. Unfortunately, structural determination of IMPs remains an immense challenge with only ~1.5% of the protein structures in the Protein Data Bank (PDB) are of IMPs, thus limiting our understanding of their function at the molecular level.

Most membrane proteins do not exist in abundance naturally, and therefore heterologous overexpression of IMPs is a prerequisite challenge for structural studies. Heterologous expressions of eukaryotic IMPs in bacteria has limited success because of the toxicity associated with saturating the secretory system, incompatibility in translation and traffic to the membrane, rapid degradation of the foreign protein, lack of post-translational modifications, and other misfolding problems of directing the protein to inclusion bodies.

Recently, a novel membrane-associating protein discovered in *Bacillus subtilis*, Mistic, has been shown to enhance the expression levels of a large number of foreign IMPs in their inner membrane of *E. coli*, when used as a fusion partner linked to their N-termini [4-8]. Interestingly, Mistic lacks any identified signal sequence and its overexpression, whether alone or in fusion to other IMP, lacks the toxicity issues associated with overloading the protein translation machinery and thus is likely to bypass the *E. coli* translocon complex. The solution-NMR structure of Mistic, in detergent micelle of lauryl-dimethylamine-*N*-oxide (LDAO), shows a compact four-helix-bundle fold.[4] Structurally, Mistic differs from most of the membrane proteins as its surface area is not very hydrophobic but rather made up of many acidic residues, and none of its helices are predicted to form hydrophobic transmembrane segments. Therefore, the resident orientation of Mistic within the membrane has not been established, and the relationship of the four-helix bundle structure to its function is still to be investigated by other techniques.

Mistic requires detergent for solubility, and it has been so far unsuccessful to crystallize. To bypass this challenge, we have studied Mistic homologs found in other *Bacillus* species [5]. These Mistic orthologs from *B. licheniformis* (M2), *B. mojavensis* (M3), and *B. atrophaeus* (M4) all lack the N-terminal 26 amino acids of Mistic from *B. subtilis*, corresponding to the first α-helix of its solution-NMR structure. For nomenclature, we refer to the originally identified Mistic from *B. subtilis* (110 amino acids long) as M110, and to its shorter sequence starting at Met27 as M1 (Fig. 1). We have expressed these Mistic proteins, M1, M2, M3, and M4, in *E. coli* independently, and found that, unlike M110, the shorter forms are significantly more soluble in the absence of detergent. In light of the crystals of M2 we report here and our data on the formation of fibrils by Mistic proteins, we discuss the correlation between its crystalline, soluble, and aggregated states and the possible implications for membrane insertion.

C. Results and Discussion

1. Mistic Orthologs Express in the Cytoplasm

The facts that M110 expresses primarily in the bacterial membrane, and that the shorter Mistic proteins have comparable capacity to enhance expression of IMP fused to their C-termini, stimulated us to determine whether they too express at their inner membrane of *E. coli [4, 5]*. Figure 2 shows that M110 is essentially insoluble and associates tightly with the membrane. Surprisingly, M1, which corresponds to the amino acid sequence 27-110 of M110, is mainly soluble with some level of

association with the membrane. M2, the most distant sequence from M1, also exists mainly in the soluble fraction but shows an increased association with the membrane. Interestingly, M3 and M4, which are more closely related to the sequence of M1, interact significantly with the membrane, with M4 showing the highest existence in the pellet, comparable with that of M110.

The mutation M75A had been shown to dramatically reduce the membrane association of M110 leading to cytoplasmic protein expression and lack of enhanced IMP expression as a fusion partner [4]. Similarly, we find that the mutation corresponding to this position of M1, M49A, yields a highly soluble protein product with practically no existence at the membrane (see Fig. 2). M2 has Ile instead of Met as position 49; however, it has a Met at position 48. In an attempt to increase its solubility, we generated two M2 mutants, M48A and I49A. Our data shows that M48A, but not I49A, reduces membrane association and increased soluble expression levels of the wt M2.

2. The Shorter Mistic Proteins Are Soluble Multimers

As seen earlier, the shorter Mistic forms, M1 through M4, are all abundant in the bacterial cytoplasm and thus do not require detergent for their purification. Moreover, expression can be readily upscaled to obtain 30-40 mg/L of pure protein from either form, which makes them potentially less challenging for crystallization than M110 that requires detergent for solubility.

All the Mistic proteins migrate as monomers on an SDS gel. However, it can be seen in Figure 2 (Soluble gel) that there are noticeable differences in their

migration rates that cannot be accounted for by the difference in their mass (~10,500 Da for all of them). Such differences in the migration rates are indicative of variation in the monomeric form, even in the presence of SDS, particularly exemplified by the mutation M48A which slows the migration of M2.

Interestingly, in the absence of detergent, all the short Mistic forms exist as multimers in solution (including M1, M2, M3, and M4), as typified by the migration of M2 on a size-exclusion chromatography (SEC) column shown in Figure 3. Static-light scattering analysis shows that there is a relatively stable oligomeric form with ~173 kDa, corresponding to about 16 M2 subunits (each monomer of 10,516 Da). However, higher order aggregates exist and their proportion increases daily when the proteins are stored at 4°C or -20°C. In contrast, M1 M49A exists as a discrete and stable tetramer with a measured molar mass of ~43 kDa (see Fig. 3). Interestingly, although the mutation M48A increases the solubility of M2, it neither changes the aggregation of M2 dramatically nor gives rise to discrete tetramerization of M2.

3. Crystals of M2

Data from protein crystals were collected to a resolution of 3.2 Å. Initial indexing of the intensities indicated the possibility of a primitive orthorhombic Bravais lattice (since the β angle is close to 90°, see later); however, the merging statistics were only reasonable for a monoclinic lattice. Inspection of the systematic absences clearly revealed (o,*k*,o) with $k = 2n$ as the reflection condition, consistent with the P21 space group. The reflection data statistics for the M2 crystals are shown in Table I.

Solvent-content analysis indicates eight molecules in the asymmetric unit (ASU) for crystals containing 54.7% solvent volume [9]. The self-rotation function calculated with MOLREP[10] has its third strongest peak at χ =90°, indicating the presence of fourfold noncrystallographic symmetry (NCS) in the ASU [Fig. 4(A)]. Moreover, the native Patterson map has a strong peak (~39% of the origin peak) at 0.5, 0.5, 0.5, indicative of translational element of NCS, or pseudo-centering symmetry [Fig. 4(B)]. Taken together, it seems likely that the ASU contains two tetramers related by pseudo-centering symmetry, and thus the whole unit cell consisting of four tetramers. Interestingly, the size of such as assembly is consistent with the size we measured in solution by light scattering, although the symmetry might not be similar.

The solution-MNR structure of Mistic did not provide significant molecular replacement solutions, and the ab initio or structure-based theoretical models for M2 produced by the Rosetta server (Rosetta.org) as well. It is likely that there are variations in the structures among the eight putative molecules of M2 of ASU. Regardless, it is evident that the large number of molecules in ASU prohibits structure determination by molecular replacement. Translational element does not contribute to NCS averaging, thus provides no help in phase improvement. For the purpose of crystal structure determination, we are searching for heavy-atom derivative crystals and new crystal forms with smaller number of ASU and with better diffracting power.

4. Fibril Formation by M2

Interestingly, M2 turns into a gel-like state once concentrated down to ~20 mg/mL at 25°C, a process that is inhibited at 4°C or on ice. M1, although more soluble than M2, can still turn into gel-like entity at around 20 mg/mL, whereas the mutant M1 M49A is highly soluble and does not form a gel up to a concentration of ~100 mg/mL. This form of polymerization is common among proteins that form amyloid fibrils. We have therefore focused on the analysis of concentrated (>20 mg/mL) samples of M2, M1, and M1 M49A by electron microscopy (EM). The EM data show that M2 forms fibrils, M1 forms short and bended protofibrils, which is a form of premature fibril, while M1 M49A does not aggregate at all (Fig. 5). Moreover, it can be seen that M2 exists primarily in the fibrillar state and no globule-like structures were observed, indicating the practical absence of low-order assemblies. In contrast, M1 M49A shows only globule-like structures typified by low-order assemblies.

Unlike amyloid fibrils, which are formed by stacking of β–sheets, the fibrils of M2 and M1 are likely composed of α–helical structure as indicated by CD [Fig. 6(A)] and by the solution-NMR structure of M110 [4]. Additional support for this notion comes from the failure of M2 fibrils to bind the traditional amyloid-specific dye, Thioflavin-T, suggesting that these fibrils are not amyloid fibrils, thus probably not made up of β-sheets. An alternate way to form highly organized aggregates has been proposed by the "runaway domain swapping" model. It is different from the commonly observed "close-ended" domain swapping, in which monomers swap small domains to form symmetric oligomers (such as dimers, trimmers, tetramers, etc.) [10-

15]. In the case of runaway domain swapping for the polymerization to continue to a higher order, it requires "open-ended" domain swapping, in which each monomer swaps to a domain with the complementary domain of an adjacent monomer along the fibril allowing it to grow on both ends.

If the polymerization observed in the solution is similar to that in the multimeric crystal of M2 we herein report, then the three-dimensional packing of such a structure may shed light into the formation of runaway domain swapping. In addition, since the mutation M49A of M1 appears to inhibit fibril formation and eliminates higher-order aggregates, it is possible that the tetramer is the minimal domain-swapping unit in this case. Studying such a structure in comparison with its *wt* would be highly valuable to our understanding of α–helical fibril formation. We hope that the data we present here will facilitate such analysis in the near future.

Finally, both the highly charged surface of the solution-NMR structure of Mistic and the high solubility of the short Mistic orthologs make it difficult to imagine how it inserts into the hydrophobic lipid-bilayer membrane. Polymerization, as observed for all the soluble Mistic proteins, may help to mask the charged surface of the monomers and thereby assist membrane insertion. Intriguingly, it has been shown that fibrils tend to associate with lipid membranes therefore structural insight into the polymerization of Mistic may provide insight into the unknowns biophysical process of membrane insertion independent of the translocon machinery [16-19].

D. Materials and Methods

1. Cloning and Purification

All Mistic homologs were cloned into the plasmid pET21b (Novagen, Madison, WI) under the control of a T7 promoter, with a C-terminal thrombin cleavage site followed by a hexa-His tag. Freshly transformed BL21(DE3) *E. coli* cells (Invitrogen, Carlsbad, CA) were grown overnight in Terrific Broth (TB) supplemented with 200 μg/mL ampicillin. The cultures (20 mL) were inoculated with 1 mL of the aforementioned starter cultures, and then the growth proceeded to an approximate optical density of 1.0. Protein production was induced with 0.5 m*M* isopropyl β–D-1-thiogalactopyranoside (IPTG) for 5 h at 37°C. The cultures were harvested by centrifugation at 3000*g* for 15 min, and then the cell pellets were resuspended in 10 mL of lysis buffer (25 m*M* Tris pH 8.0, 20 m*M* imidazole, 200 m*M* NaCl, and 1 m*M* phenylmethylsulfonyl fluoride [PMSF]). Cells were lysed with a microfluidizer (Microfluidics, Newton, MA), and the membranes and inclusion bodies were spun down by centrifugation at 100,000*g* for 2 h. The supernatant was incubated with 200 μL buffer-equilibrated NiNTA beads (Qiagen, Hilden, Germany) for 2 h at 4°C. The mixture was columned and washed fives times by gravity with 1 column volume of wash buffer (lysis buffer without PMSF). Bound protein was eluted with 1 mL elution buffer (wash buffer with 0.5*M* imidazole). Protein extraction from the membranes was done by resuspending the pellets in 1 mL lysis buffer containing 0.5% (17.5 m*M*) SDS and shaking for 2 h at 25°C. The SDS-insoluble fraction (inclusion bodies) was spun down by centrifugation at 20,000*g* for 15 min.

Large-scale expressions of soluble Mistic proteins were carried out similarly from cultures of 1-L TB, using the soluble fraction only for purification. The His-tags were cleaved off from the proteins with bovine thrombin (Sigma-Aldrich, St. Louis, MO) digestion during dialysis against 50 m*M* NaCl for O/N at 4°C. The protein samples were analyzed for cleavage with NuPAGE 4-12% Bis-Tris gel with MES running buffer (Invitrogen, Carlsbad, CA). The completely cleaved proteins were concentrated using 50K MWCO Vivaspin concentrators (Vivascience AG, hannover, Germany) and further purified by SEC with Superdex-200 (16/60) column (Amersham Pharmacia Biotech, Piscataway, NJ) pre-equilibrated with 10 m*M* Tris pH8.0, 100 m*M* NaCl.

2. Light Scattering

Static light-scattering measurements coupled with SEC were performed using Wyatt miniDAWN (Wyatt Technology, Santa Barbara, CA) and G3000PWXL column (Tosoh Biosciene, Montgomeryville, PA). Molar masses were derive by the Zimm fit method in ASTRA V (Wyatt Technology, Santa Barbara, CA) using concentrations obtained from UV at 280 nm and protein-specific refractive index increment (d*n*/d*c*) of 0.186.

3. Crystallization

Before crystallization, M2 was concentrated to 10-15 mg/mL in the SEC buffer. The crystals of M2 grow (up to ~0.35 mm in length) via both the hanging and the sitting-drop vapor-diffusion methods, by mixing equal volumes of protein stock solution and the reservoir (38% MPD, 0.1*M* Tris pH 8.8 and 0.2*M* MgAc). The

protein crystals grow in 2-4 days at 15°C. Before X-ray data collection, protein crystals were transferred to a fresh mother liquor solution containing 40% of MPD.

4. X-ray Data Analysis

Diffraction data from the native protein crystal were collected on an ADSC Q210 CCD detector at 100 K on beamline 8.2.1 (ALS, Berkeley, CA). All X-ray data were indexed and integrated using the program MOSFLM v.6.2.4 [20] and merged and scaled with SCALA [11] from the CCP4 program package [21].

5. Electron Microscopy

Protein samples were concentrated to at least 20 mg/mL and kept at 4°C. The concentrated samples were diluted in Buffer A (10 m*M* Tris, 1 m*M* EDTA, 25 m*M* dithiothreitol, 7.5) to a final concentration of 50 μM and spotted on a glow-discharged, carbon-coated Formvar grid (Electron Microscopy Sciences), incubated for 5 min, washed with distilled water, and then stained with 1% (w/v) aqueous uranyl formate solution. Uranyl formate solutions were filtered through 0.2-μm sterile syringe filters (Corning, Corning, NY) before use. EM analysis was performed suing a JEOL JEM-100CXII electron microscope at 80 kV with nominal magnifications between 36,000 and 72,000. The images were recorded digitally by using the SIS Megaview III imaging system. At least two independent experiments were carried out for each sample.

6. Thioflavin T Binding

The concentrated samples were diluted with Buffer A to a final concentration of 50 μ*M*, and then the solution was mixed with 2 μL of 1 m*M* Thio T prepared in the

same buffer. Fluorescence was measured immediately after the addition of Thio T. The experiment was measured on a spectrofluorimeter (Photon Technology International, Lawrenceville, NJ) with excitation at 450 nm and emission at 482 nm. A rectangular 10-mm quartz microcuvette was used. Three independent experiments were performed for each sample.

7. Circular Dichroism Spectroscopy

Protein samples were concentrated to at least 20 mg/mL and kept at 4°C. The concentrated samples were diluted into Buffer A to a final concentration of 25 μM. The solution was then placed into a 0.1-cm path-length quartz cell (Hellma, Forest Hills, NY). Spectra were acquired using an instrument of BioLogic MOS-450 (Molecular Kinetics, Pullman, WA). All measurements were doen at 23°C. Spectra were generally recorded over the wavelength range of 196-260 nm. Three independent experiments were performed with each sample. Raw data were manipulated by smoothing and subtraction of buffer spectra, according to the manufacturer's instructions.

E. References

1. 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. 1029-38.

2. Boyd, D., C. Schierle, and J. Beckwith, *How many membrane proteins are there?* Protein Sci, 1998. **7**(1): p. 201-5.

3. Speers, A.E. and C.C. Wu, *Proteomics of integral membrane proteins--theory and application.* Chem Rev, 2007. **107**(8): p. 3687-714.

4. Roosild, T.P., et al., *NMR structure of Mistic, a membrane-integrating protein for membrane protein expression.* Science, 2005. **307**(5713): p. 1317-21.

5. Roosild, T.P., et al., *Characterization of the family of Mistic homologues.* BMC Struct Biol, 2006. **6**: p. 10.

6. Kefala, G., et al., *Application of Mistic to improving the expression and membrane integration of histidine kinase receptors from Escherichia coli.* J Struct Funct Genomics, 2007. **8**(4): p. 167-72.

7. Liu, X., et al., *Characterization of mitochondrial trifunctional protein and its inactivation study for medicine development.* Biochim Biophys Acta, 2008. **1784**(11): p. 1742-9.

8. Lee, K.E., et al., *Regulation of CD40 reconstitution into a liposome using different ratios of solubilized LDAO to lipids.* Colloids Surf B Biointerfaces, 2008. **62**(1): p. 51-7.

9. Matthews, B.W., *Solvent content of protein crystals.* J Mol Biol, 1968. **33**(2): p. 491-7.

10. Vagin, A. and A. Teplyakov, *Molecular replacement with MOLREP.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 22-5.

11. Potterton, E., et al., *The CCP4 molecular-graphics project.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 11): p. 1955-7.

12. Naiki, H., et al., *Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavin T1.* Anal Biochem, 1989. **177**(2): p. 244-9.

13. Bennett, M.J., M.P. Schlunegger, and D. Eisenberg, *3D domain swapping: a mechanism for oligomer assembly.* Protein Sci, 1995. **4**(12): p. 2455-68.

14. Bennett, M.J., S. Choe, and D. Eisenberg, *Refined structure of dimeric diphtheria toxin at 2.0 A resolution.* Protein Sci, 1994. **3**(9): p. 1444-63.

15. Guo, Z. and D. Eisenberg, *Runaway domain swapping in amyloid-like fibrils of T7 endonuclease I.* Proc Natl Acad Sci U S A, 2006. **103**(21): p. 8042-7.

16. McLean, L.R. and A. Balasubramaniam, *Promotion of beta-structure by interaction of diabetes associated polypeptide (amylin) with phosphatidylcholine.* Biochim Biophys Acta, 1992. **1122**(3): p. 317-20.

17. Mirzabekov, T.A., M.C. Lin, and B.L. Kagan, *Pore formation by the cytotoxic islet amyloid peptide amylin.* J Biol Chem, 1996. **271**(4): p. 1988-92.

18. Anderluh, G., et al., *Interaction of human stefin B in the prefibrillar oligomeric form with membranes. Correlation with cellular toxicity.* FEBS J, 2005. **272**(12): p. 3042-51.

19. Kagan, B.L., et al., *The channel hypothesis of Alzheimer's disease: current status.* Peptides, 2002. **23**(7): p. 1311-5.

20. Battye, T.G., et al., *iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 271-81.

21. Collaborative Computational Project, N., *The CCP4 suite: programs for protein crystallography.* Acta Crystallogr D Biol Crystallogr, 1994. **50**(Pt 5): p. 760-3.

TABLES

Table 2.1 X-ray Data Collection Statistic for M2 Crystals

Table I. *X-ray Data Collection Statistics for M2 Crystals*

X-ray source	ALS bl8.2.1
Wavelength (Å)	1.0781
Space group	P21
Unit cell dimensions	80.00, 48.23, 114.41,
a, *b*, *c* (Å) and β (°)	and 90.52
Resolution range (Å)	114.7–3.2 (3.37–3.2)
R_{sym} [a]	0.079 (0.410)
Unique reflections	14,685 (2089)
Completeness (%)	99.5 (99.2)
$I/\sigma(I)$	3.9 (1.9)
Multiplicity	4.0 (3.7)

[a] $\Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_h\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th measurement and $\langle I(hkl)\rangle$ is the mean of all measurements of $I(hkl)$ for Miller indices hkl.

FIGURES

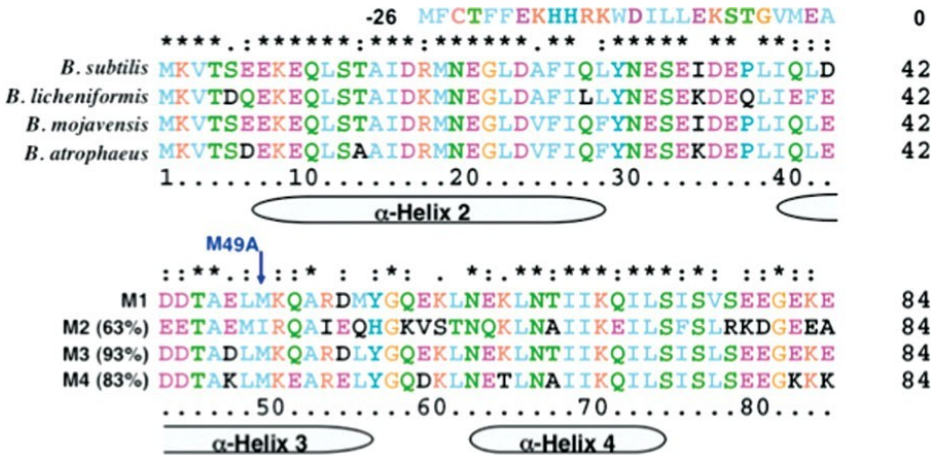


Figure 2.1 Sequence alignment of Mistic proteins from *Bacillus* species. Sequence identity with M1 is given in brackets. α-helices 2, 3, and 4 are aligned according to their positions in the solution-NMR structure of Mistic from *Bacillus subtilis* (M110). The first helix of M110, which is not found in the shorter Mistic homologs, is numbered -26 to 0. The M49A (blue) mutation of M1 corresponds to position 75 of M110. Sequences were aligned using ClustalX. Amino acids with significantly different residues are marked in black.

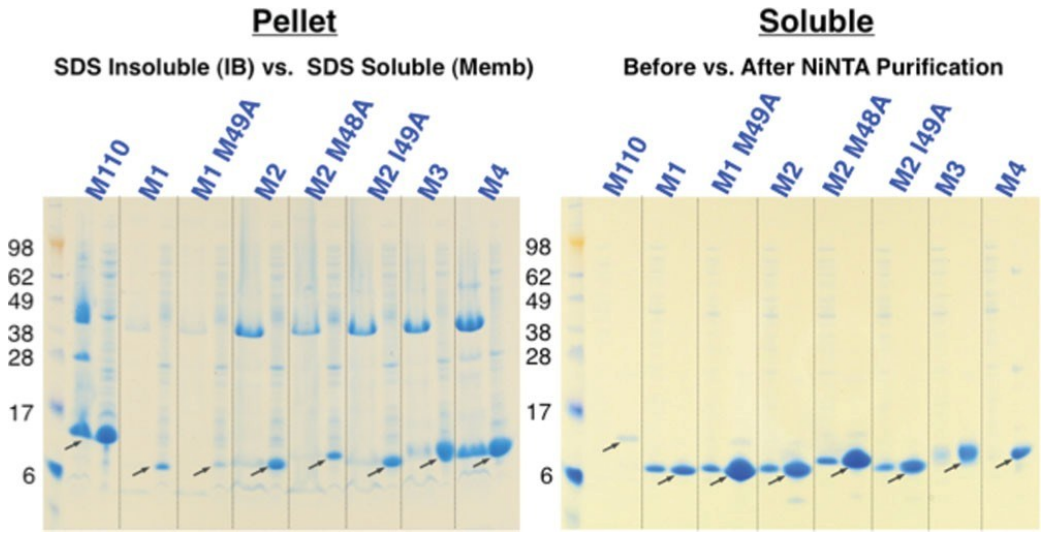


Figure 2.2 Target expression of *wt* and mutant Mistic proteins. Soluble proteins (right gel) were separated from the pellet (left gel) containing membrane (Memb) and inclusion bodies (IB) proteins. Membrane-associated proteins were solubilized by SDS, after which SDS-insoluble proteins (IB) were spun down (See Materials and Methods section). On the pellet gel, 10 µL samples were loaded of IB on the left next to SDS-solubilized pellet (out of 1mL total) on the right (Memb). On the soluble gel, 10 µL samples (out of 10 mL total soluble fraction, or 1 mL of the Ni^{2+} elution) were loaded before (left) and after (right) Ni^{2+} purification. Proteins bands are identified by Coomassie stain. The IB bands that show up for most samples at around 40 kDa does not correspond to Mistic protein, but rather to an *E. coli* outer membrane pore protein, as revealed by its N-terminal sequence (AEIYNKDGNKMD). Interestingly, the intensity for that band is correlated with the intensity for the membrane-associated Mistic proteins. We have detected this band in overexpression of some bacterial IMP fused to Mistic (unpublished data), and it appears that the expression of this porin may get induced upon population of the membrane by Mistic proteins. However, this possible connection would need to be further confirmed and investigated to understand its mechanism.

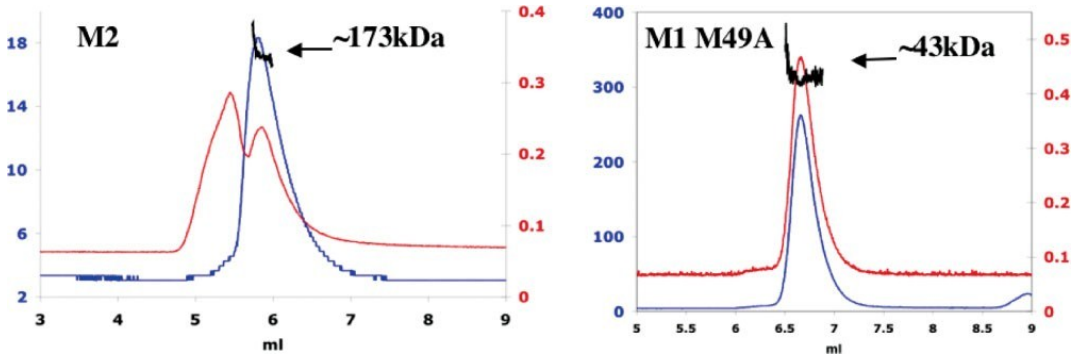


Figure 2.3 Oligomeric state measured by light scattering. Purified M2 (left) and M1 M49A (right) were injected on the SEC column G3000PWxl in line with a miniDAWN light scattering device. The blue trace shows the UV absorbance in mOD, the red trace shows the light scattering (Rayleigh ration at 90°), and the black trace shows the calculated (in the ASTRA software using the Zimm model with a protein d*n*/d*c* value of 0.186) molar mass distribution around the main peaks. The chromatogram for M2 is typical for all the soluble Mistic proteins M1 through M4 with various degrees of high-order aggregation.

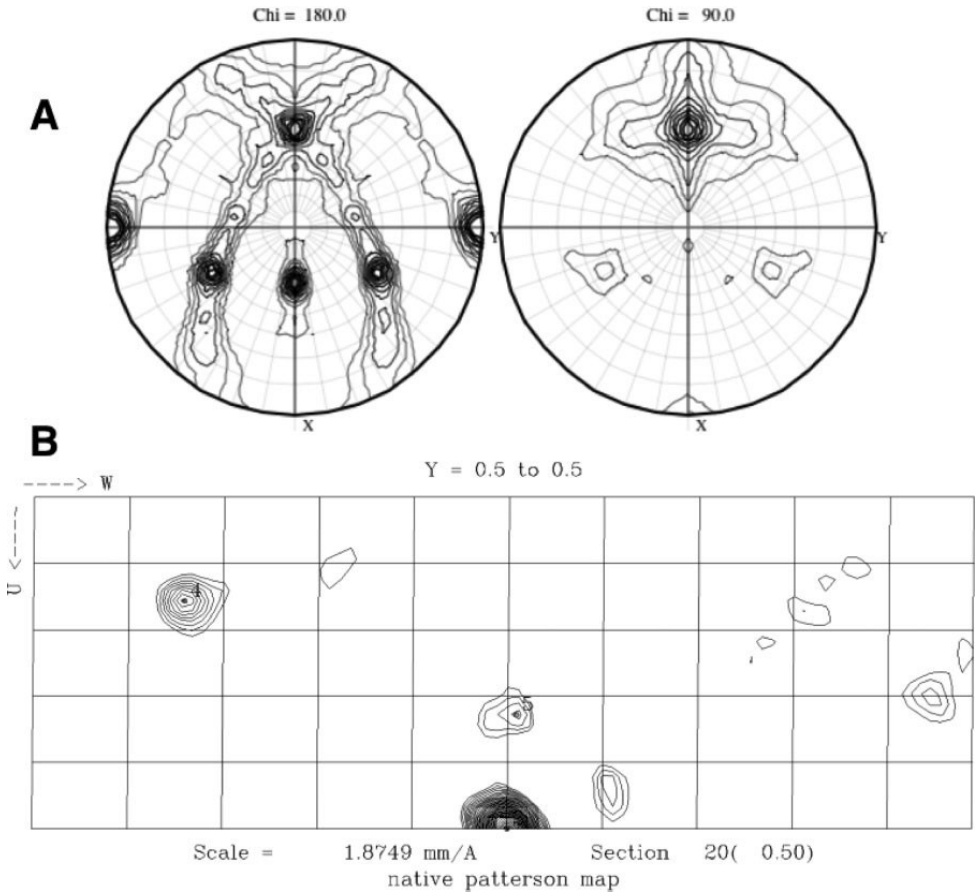


Figure 2.4 (A) Stereographic projection of the self-rotation function. The values of the self-rotation function (calculated from data in the resolution range 40-3.8 Å) at χ = 180° (left) or χ = 90° (right) are plotted with the polar angles (θ, ϕ, χ) convention looking down the *z*-axis. These projections have their τ = 0° or 180° at the center and τ = 90° at the edges, whereas the ϕ angle goes from 0° to 360° around the sphere starting from *x* in the direction of *y*. The χ = 180° projection has its twofold crystallographic symmetry peaks along the *y*-axis at the edges of the sphere (θ = 90°, ϕ = 90° or -90°). The other off-edge peak indicates the presence of twofold NCS, whereas the χ = 90° projection has a single peak that indicates the presence of fourfold NCS in the ASU; calculated using MOLREP.[10] (B) Native Patterson map. The Patterson map was calculated using data from the native M2 crystal in the 40-3.8 Å resolution range excluding the reflections $F_{obs} < 3\sigma$ (F_{obs}). The Harker section at u = 0.5 shows a strong peak (about 39% the intensity of the origin peak) at u = 0.5, v = 0.5, w = 0.5, indicating a pseudo-centering symmetry. The calculation was done using the Patterson routine from the CCP4.

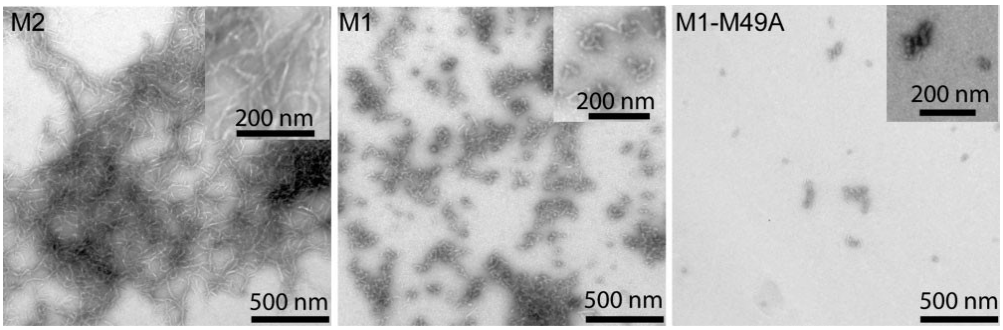


Figure 2.5 EM micrographs showing fibril structures of M2.

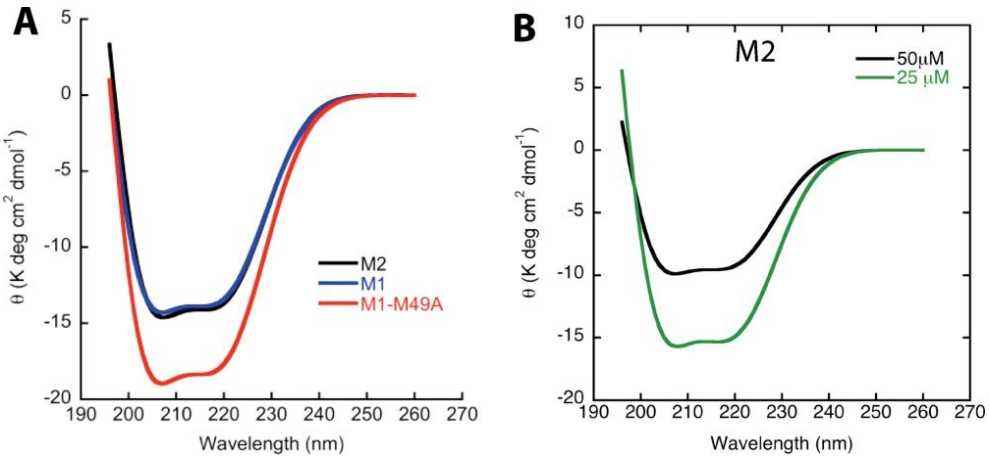


Figure 2.6 CD spectra of Mistic. In Panel A, the CD spectra of M2, M1, and M1M49A shows helical structure in solution. The negative ellipticity (θ) of M1M49A at 222 and 208 nm is greater than that of M2 and M1, suggesting more helical content of this protein in solution. In Panel B, the CD spectrum of M2 shows decreased helicity at higher protein concentration; however, the presence of helical secondary structure is still evident.

ACKNOWLEDGEMENTS

Chapter II, in full, is a reprint that the dissertation author significantly contributed to as both a researcher and an author. The material appears in *Protein Sci*.

Dvir, D., **Lundberg, M.E.,** Maji, S.K., Riek, R., and Choe, S. (2009). Mistic: Cellular localization, solution behavior, polymerization, and fibril formation. *Protein Sci,* **18**, 1564-1570

CHAPTER III

MstX and a putative potassium channel facilitate biofilm formation in *Bacillus subtilis*

A. Abstract

Biofilms constitute the predominant form of microbial life and a potent reservoir for innate antibiotic resistance in systemic infections. In the spore-forming bacterium *Bacillus subtilis*, the transition from a planktonic to sessile state is mediated by mutually exclusive regulatory pathways controlling the expression of genes required for flagellum or biofilm formation. Here, we identify *mstX* and *yugO* as novel regulators of biofilm formation in *B. subtilis*. We show that expression of *mstX* and the downstream putative K+ efflux channel, *yugO*, is necessary for biofilm development in *B. subtilis*, and that overexpression of *mstX* activation induces biofilm assembly. Transcription of the *mstX-yugO* operon is under the negative regulation of SinR, a transcription factor that governs the switch between planktonic and sessile states. Furthermore, *mstX* regulates the activity of Spo0A through a positive autoregulatory loop involving the KinC histidine kinase that is activated by potassium leakage. The addition of potassium abrogated *mstX*-mediated biofilm formation. Our findings expand the role of Spo0A and potassium homeostasis in the regulation of bacterial development.

B. Introduction

Nearly all bacteria are capable of forming multicellular communities through complex signaling events that lead to differentiation into a myriad of cell types. These sessile, surface-attached bacterial populations, referred to as biofilms, create an elaborate extracellular matrix comprised of protein and exopolysaccharides that enhance survival in a nutrient-depleted state and mediate attachment to surfaces [1, 2].

Cells lacking either component of the extracellular matrix form flat, featureless colonies devoid of complex architecture and they fail to adhere to surfaces [3]. A signature feature of biofilm communities is their increased resistance to antibiotics and environmental stresses; both features make them particularly problematic in clinical and industrial settings. For instance, biofilms constitute over 65% of bacterial infections and represent a formidable source of innate multidrug resistance [4]. A small percentage of bacteria in a biofilm give rise to dormant, persistent cells that are recalcitrant to antibiotic treatment. The molecular mechanisms by which biofilms acquire antibiotic resistance have only recently begun to emerge [5-7].

The Gram-positive bacterium *B. subtilis* is capable of forming thin, floating biofilms at the liquid-air interface (pellicles) with sporulation at apical tips that project into the air [8]. Previous studies have identified multiple transcriptional regulatory networks that give rise to multicellular communities in *B. subtilis*. These pathways include the regulatory proteins Spo0A, σ^H, the transition state regulator AbrB, the master regulator for biofilm formation SinR, and Slr [8-12]. Biofilm assembly requires expression of the 15-gene operon *epsA-O,* which encodes enzymes that synthesize the exopolysaccharide layer, and the *tapA-sipW-tasA* operon, which encodes the protein component. SinR is a transcription factor that directly represses these genes and the flagellar motor inhibitor EpsE during exponential growth [13]. It also inhibits Slr, a transcriptional factor that activates biofilm genes while repressing motility [14]. The balance between SinR and Slr activity depends on Spo0A-P accumulation, which allows production of SinI, an inhibitor of SinR, which therefore

turns on matrix production and turns off motility [12]. The switch between motility and biofilm formation therefore critically depends on the phosphorylation state of Spo0A, which is controlled by a variety of kinases and phosphatases that respond to different stimuli including oxidative stress, K+ leakage, osmotic pressure, and malic acid ([15-17]. These kinases (KinA, KinB, KinC, and KinD) help facilitate biofilm formation through spatial regulation but can be partially redundant through signaling overlap [18].

Mistic (MstX) is a unique protein found in a small number of *Bacillus* species, including *B. subtilis*, that enables high-level, heterologous expression and targeting of integral membrane protein sequences to cell membranes when fused to the N-terminus of a cargo protein construct [19]. In spite of its small and highly acidic nature, MstX associates with the membrane, presumably through autonomous association with the phospholipid bilayer, thereby bypassing or facilitating the traditional secretory apparatus [19]. The MstX homologues in *Bacillus atrophaeus*, *Bacillus mojavensis*, and *Bacillus licheniformis,* like *B. subtilis,* facilitate heterologous integral membrane protein expression when used as part of a fusion construct [20]. Furthermore, in all cases, *mstX* homologues precede a putative potassium ion channel *yugO,* suggesting that the MstX protein might be involved in membrane insertion of YugO (Figure 1). No similar sequence with a known function exists, raising the question as to what function MstX might serve in *Bacillus subtilis*.

The initial goal of the present work was to elucidate the function of *mstX* in the Gram-positive bacterium, *Bacillus subtilis*. During the course of this investigation, we

discovered novel roles for *mstX* in cell-fate determination and morphogenesis. We show that *mstX* is necessary for robust biofilm formation. The *mstX* promoter is regulated by SinR, the master regulator for biofilm formation, and induces biofilm formation at least partially through KinC mediated phosphorylation of Spo0A, and correspondent increases in expression of the regulators *abbA* and *sinI*. A mutation in SinR proved epistatic to the *mstX* biofilm film defect, restoring both colony morphology and pellicle formation in a double mutant. Supplementation of media with potassium or disruption of the downstream putative potassium ion channel abrogated *mstX*-mediated biofilm formation, illustrating the importance of KinC activation and potassium in biofilm development. These data suggest that *mstX* operates through a potassium efflux-driven positive feedback loop that enhances biofilm formation in *B. subtilis*.

C. Methods and Materials

1. Strains, Media, and Culture conditions

The parent strains for all experiments were either *B. subtilis* strains PY79 or NCIB3610 wild strain [8, 21]. Deletion mutants were created by long-flanking homology PCR or by standard cloning procedures [22]. A *loxP-kan-loxP* cassette was used to construct the *mstX* deletion, and after integration into the *B. subtilis* chromosome, the cassette was removed by Cre-mediated excision [23]. The IPTG-inducible expression strain, P_{spac}-*mstX*, was constructed by cloning a full-length copy of the *mstX* coding region downstream of the P_{spac} promoter and subsequent integration at the *lacA* locus. P_{spac}-*mstX M75A* was obtained through site-directed

mutagenesis of the resulting plasmid as described [24]. Additional details pertaining to strain construction can be found in the Supplemental Methods and Materials section (Supplemental Table S1).

2. Biofilm growth and crystal violet assay

Biofilm growth and crystal violet assays for PY79 strains were performed essentially as described by Hamon and Lazazzera [9]. *Bacillus subtilis* starter cultures were grown to OD_{600} 0.3 at 37°C and added to polyvinylchloride microtitre plates (Fisher scientific) at a final OD_{600} of 0.01. Biofilm growth media was Luria-Bertani medium in addition to 0.15 ammonium sulfate, 100 mM potassium phosphate pH 7.0, 30 mM sodium citrate, 1 mM $MgSO_4$ and 0.1% glucose or MsGG [8]. Samples of 100 µl diluted cells were added to 96-well PVC microtitre plates and incubated under stationary conditions at 30°C. 24 h after inoculation, we mixed the cultures by pipetting up and down as a means of oxygenating the cells. In addition, spent growth medium was exchanged for fresh biofilm growth medium. 72 h after inoculation and growth at 30°C, liquid medium was removed and wells were washed with fresh biofilm growth medium. Cells that had adhered to the wells were stained with 0.1% crystal violet at room temperature for 20 min. Excess crystal violet was then removed and adherent cells were washed with biofilm growth medium. The crystal violet that had stained the cells was solubilized in 200 ml of 80% ethanol and 20% acetone. Biofilm formation for each well was quantified by measurement of OD_{570} using a spectrophotometer. For characterizing biofilm growth in the NCIB3610 strain, cells were grown to OD_{600} 0.8 in LB and spotted with 2 µl on MSGG plates or inoculated in

MSGG-containing wells [8]. Plates were grown for three days at 22°C and pellicles were then photographed with a Nikon Coolpix S4300 camera under special lighting. Wells measured approximately 3 cm in diameter

3. Chromatin Immunoprecipitation (ChIP)

B. subtilis strains were grown in Luria-Bertani medium to OD_{600} 0.5. Cultures were incubated for 30 min at room temperature. All chromatin immunoprecipitations were performed essentially as described [25]. Formaldehyde at final concentration of 1% and $NaPO_4$ at final concentration 10 mm were added to 50 ml of cells grown in LB and 0.1 mM IPTG at OD_{600} 0.5. A *B. subtilis* strain (MEL102) containing a *sinR-FLAG* gene fusion was created by cloning the complete sequence of *SinR-FLAG* into pMUTIN4 and by single-crossover recombination. The cross-linking reaction was terminated by the addition of glycine to a final concentration of 200 mM. Following cross-linking, cells were collected by centrifugation and were washed with TBS 50 mM Tris-HCl pH 7.5, 150 mM NaCl). The cells were then suspended in 1 ml ice cold lysis buffer [50 mM Tris-HCl with 1mM EDTA, 5 mg/ml lysozyme and protease inhibitor cocktail (Roche)] and incubated for 30 min at 37°C. After lysis, Triton X-100 at final concentration 1% and sodium deoxycholate at final concentration 0.1% were added. The DNA was sheered by sonication to 500-1000 base-pairs, as determined by agarose gel electrophoresis. Insoluble cell debris was removed by centrifugation and the 50 µl supernatant was removed and added to 200 µl TES buffer (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 1% SDS). DNA extracted from the insoluble cell debris was used as the "total DNA" control.

Protein and protein-DNA complexes were incubated (4°C O/N) with 5 µg monoclonal rabbit anti-FLAG antibody (Sigma-Aldrich) followed by incubation for 1 h at 4°C with Protein A-Sepharose beads (Sigma-Aldrich). Complexes were collected by centrifugation and washed 4 times (5 min at room temperature) with 1.5 ml Wash Buffer A (5 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, Roche protease inhibitors), twice (5 min at room temperature) with 1.5 ml of Wash Buffer B (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM LiCl, 0.5% NP-40, 0.5% sodium deoxycholate) and once (5 min at room temperature) with TE Buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA). Protein and protein-DNA complexes were eluted from the beads by the addition of 100 µl of TES Buffer (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 1% SDS) and incubation at 65°C for 15 min. The beads were removed by centrifugation and the eluate was transferred to a fresh tube and re-extracted with 150 µl of TES. The eluates as well as the "total DNA" sample were incubated overnight at 65°C. The DNA was subsequently extracted with phenol chloroform extraction and ethanol precipitation, with the resulting DNA resuspended in 100 µl TE buffer.

PCR with the listed primers (see Supplementary Table 2) were carried out w/*Taq* DNA polymerase using standard PCR reaction conditions, with 1 µl of the DNA from the precipitation used as template. 1/100 µl of the "total DNA" was used for comparison. 25 amplification cycles were performed and the resulting PCR products were analyzed by agarose gel electrophoresis and ethidium bromide staining.

4. RT-PCR and quantitative RT-PCR analysis

We used standard methods for analyzing gene expression via RT-PCR and quantitative RT-PCR, with some modifications [26]. For qRT-PCR analysis of *sinI*, *abbA*, *mstX*, *yqxM* and *sigA*, we used the primers listed in Supplementary Table 2. The constitutively expressed *veg* gene (also known as BSU00440) was used as a positive control for quantification gene expression, as described by Hamon and colleagues [10]. For RT-PCR analysis of early log, stationary growth and biofilm growth cultures, we collected cells grown in LB medium at OD_{600} 0.3-0.5, OD_{600} 1.0, and microtitre plates. Cells grown as a biofilm were collected and characterized according to the growth protocol described above. RNA were isolated using RNAeasy miniprep columns (Qiagen) with the resulting RNA subjected to DNAse I digestion for 1 hr in order to remove contaminating chromosomal DNA. At this point, a "no reverse transcriptase" negative control was aliquoted, diluted 1/1000, and stored at -80°C. Following DNAse I digestion, cDNA were synthesized with Superscript II reverse transcriptase and a random hexamer primer according to standard protocol (Invitrogen). Following reverse transcription, the mixture cDNA and RNA were digested with RNAse to remove contaminating RNA. The solution was diluted by 1/1000 in Tris-HCl pH 8.5 buffer and stored at -80°C prior to use.

For RT-PCR, we added 1 µl cDNA or 1 µl from the no RT negative control to 100 µl GoTaq green master mix (Promega), applied 23 cycles, and analyzed by agarose gel electrophoresis. For real-time PCR, *abbA, sinI, kinC,* and *yqxM* transcription levels were standardized using *sigA* transcription and the $^{\Delta\Delta}$Ct method for

quantification [27, 28]. 1 µl of cDNA preparation were added to 25 µl SYBRgreen mastermix (Invitrogen) and cycled 40 times in an ABI7900HT thermocycler. Data were averaged across three independent trials. Primer sequences are shown in Supplementary Table 2.

C. Results

1. MstX and YugO promote biofilm formation in domesticated and undomesticated *B. subtilis* strains

In order to elucidate the function of *mstX* in *Bacillus subtilis*, we first examined the effects of *mxtX* overexpression and an *mstX* deletion (□*mstX*) on growth. To do so, we constructed a strain expressing *mstX* under an IPTG-inducible promoter (P_{spac}), integrated this construct into the chromosome of the domesticated strain PY79, and plated the strain in the presence and absence of 200 µM IPTG, on LB agar plates (1.5%). Induction of *mstX* did not impair growth (Supplementary Figure 1) but gave rise to architecturally complex colonies (Figure 2C), whereas the corresponding wild-type and □*mstX* mutant strains formed colonies devoid of the thick exopolysaccharide layer that typifies biofilm formation (Figure 2A and Figure 2B). . The *mstX* colonies showed a slight increase in motility but were otherwise indistinguishable from the wild-type strain (Figure 2B). These observations indicate that expression of *mstX* promotes the formation of architecturally complex colonies in a domesticated strain (PY79) that is otherwise incapable of forming these colonies or in forming robust biofilms.

To ensure that the observed enhancement of colony architecture was a product of *mstX* functionality as opposed to an unintended consequence of protein overexpression, we created a strain bearing an alanine substitution at a methionine residue (M75A) that has previously been shown to be essential for the ability of MstX to support high levels of membrane protein expression [19]. When expressed and purified heterologously in *E. coli*, MstX (M75A) mutant variants form more soluble oligomers that fail to associate with the membrane or enhance expression of cargo proteins [29]. In agreement with our *E. coli* expression results, the resulting P_{spac}-*mstX (M75A) B. subtilis* mutant failed to form rough colonies in the presence of IPTG, in stark contrast to the P_{spac}-*mstX* strain (Figure 2D). However, the P_{spac}-*mstX* (*M75A) B. subtilis* mutant did not restore the wild-type colony morphology, as some roughness persisted. It is probable that that the *mstX* (M75A) remained at least partially functional relative to the deletion strain. Functional expression of the *mstX* gene appears to play a significant role in producing the complicated colony architecture observed when it is overexpressed in the domesticated strain PY79.

We postulated that *mstX* might also be necessary for biofilm formation, a process that is associated with the ability to form architecturally complex colonies. We therefore quantified biofilm formation in a domesticated strain via a polyvinyl-chloride crystal violet assay [9, 30]. The assay measures the ability of bacterial cells to adhere and retain to a surface after washing, thereby approximating overall biofilm mass. The □*mstX* mutant and a P_{spac}-*mstX(M75A)* mutant decreased biofilm formation over three-fold relative to the wild-type strain (Figure 2E). In the P_{spac}-

mstX strain, IPTG induction significantly enhanced biofilm formation in excess of the wild-type. These results show that *mstX* is both importantfor biofilm formation, and that it appears to stimulate biofilm formation when expressed at high levels. MstX is a unique and relatively small protein with few orthologues in other species [20], and it is improbable that a protein with minimal genetic conservation is responsible for biofilm formation. However, *mstX* is immediately upstream of *yugO* in a two gene operon. YugO is a putative potassium efflux channel that contains a highly conserved sequence motif shared among many other prokaryotic potassium channels. Due to its location adjacent to *mstX* and the previously described role for potassium efflux in biofilm assembly [15], we questioned if *yugO* is also involved in biofilm formation.

These findings led us to hypothesize that *mstX* and *yugO* play a role in *B. subtilis* biofilm formation in an undomesticated strain that produces more robust biofilms than PY79. We transferred the aforementioned mutationsand a *yugO* deletion (□*yugO*) into the undomesticated strain NCIB 3610 to see if we could replicate our results from the PY79 strain. Indeed, mutations in *mstX* and *yugO* decreased biofilm formation in both the colony architecture and pellicle formation assays (Figure 3). The resulting mutants formed thin and detached surface pellicles that failed to recapitulate the observed phenotype in the wild-type strain. The introduction of an extra copy of *mstX* or *yugO* at the *amyE* locus under the control of the xylose-inducible promoter (P_{xyl}) rescued the □*mstX or* □*yugO* strain for its biofilm defect in the presence of 0.5% xylose. However, both *mstX* and *yugO* were required in order to restore biofilm formation to wild-type levels (Figure 3), so it is likely MstX acts in

conjunction with YugO in order to promote biofilm formation. In agreement with this conclusion, a P_{xyl}-*mstX* failed to rescue a *yugO* mutation and behaved the same as a strain missing *mstX*. Similarly, a P_{xyl}-*yugO* failed to rescue a □*mstX* mutation and behaved the same as a strain missing *yugO* (Figure 3). MstX has previously been shown to enhance integral membrane protein expression when fused to a diverse number of heterologous proteins, including those of eukaryotic origin. We suggest that MstX likely plays a similar role in promoting the membrane insertion or expression of YugO, a putative potassium channel downstream of the *mstX* open-reading frame.

2. A *sinR* mutation restores biofilm formation and gene expression in *mstX* and *yugO* mutants

SinR is a key regulatory protein that represses biofilm assembly during growth and under non-biofilm promoting conditions [31]. The failure in biofilm assembly in the Δ*mstX* and Δ*yugO* mutants could either be due to a structural defect in biofilm assembly or to continued SinR activity during biofilm promoting conditions. To test this hypothesis, we examined the effect of the *sinR* mutation on colony architecture and pellicle formation in the *mstX* and *yugO* mutant strains. In agreement with previous observations, the wild-type NCIBI3610 strain and sinR mutant strain exhibited complexy colony architecture or wrinkled colonies, respectively. In contrast, smooth nonstructured colonies were detected for the *mstX* mutant and *yugO* mutants (Figure 3). Introduction of a *sinR* mutation into strains with a *mstX* or *yugO* mutation restored biofilm formation. These data suggest that the *mstX* and *yugO*

mutants fail to initiate biofilm assembly because SinR remains active, thus significantly repressing the genes responsible for assembly.

Continued SinR activity in the *mstX* mutant should result in decreased expression of genes required for biofilm formation, such as *eps* and *tasA,* which are involved in matrix assembly, and this defect in gene expression should be rescued by the disruption of *sinR* repression. To test this hypothesis, we used q-RT PCR in the NCIB3610 background strain to examine the effects of MstX and SinR on expression of these genes and on the key regulatory genes *abbA*, *sinI* and *kinC*. We harvested total RNA from the designated strains grown in MsGG media (OD_{600} 0.5) in order to mimic biofilm forming conditions. The disruption of *mstX* expression significantly reduced the expression of the biofilm structural components, *epsE* and *tasA*, and the regulatory antirepressor genes, *abbA* and *sinI* (Figure 4, 5-fold). However, the upstream kinase *kinC* experienced only marginally decreased expression (less than 2-fold), indicating that most of the observed deficiency in biofilm formation is a byproduct of decreased antirepression of SinR and AbrB by lower levels of SinI and AbbA. Decreased antirepression of SinR would thus directly lead to decreased expression of downstream targets of SinR, such as *epsE* and *tasA*.

The expression of *sinI* and *abbA* are controlled by the sporulation and biofilm regulatory factor, Spo0A. Spo0A is a bistable regulator that is activated heterogeneously across the cell population. Although SinR predominates over SinI even in media that promotes the derepression of the *eps* matrix operon, *sinI* is selectively overexpressed in a distinct subpopulation of cells that ultimately gives rise

to a biofilm. Increased phosporylation of Spo0A in this subpopulation contributes to increased expression of SinI and consequent SinR repression. These results suggest that the *yugO* and *mxtX* mutants affect the levels of Spo0A-P, most probably through increased activity of one of the upstream kinases, KinA, KinB, KinC, or KinD (Figure 4B).

3. The *mstX-yugO* operon is selectively expressed during biofilm assembly

If MstX and YugO are involved in biofilm assembly, then they should be produced during conditions that promote biofilm production. We therefore assessed temporal expression of *mstX* in *B. subtilis*. We used RT-PCR analysis to determine *mstX* transcript levels during logarithmic growth phase and stationary cultures grown as a biofilm on PVC plates. *B. subtilis* cultures grown in LB media or under biofilm-forming conditions were collected and analyzed by RT-PCR. Primers were selected for amplifying ~200 bp regions of *mstX* and *veg*, a constituitively expressed gene that functioned as an internal control [10]. RT-PCR revealed low expression levels in early and mid-logarithmic growth. However, *mstX* transcription was significantly upregulated in RNA harvested from *B. subtilis* biofilms (Figure 5A). Thus, *mstX* is expressed during biofilm formation, consistent with the role it plays promoting biofilm formation.

4. SinR directly regulates the *mstX-yugO* promoter

Investigation of the *mstX-yugO* promoter region indicated that it possesses a putative SinR binding site (GTTCTTT) at -65 relative to the likely *mstX* translational

start codon, suggesting that the operon might be regulated by SinR. We therefore tested if SinR represses *mstX* expression. We constructed a □*sinR* mutant strain and examined *mstX* expression in the absence of SinR *in vivo*. RNA was harvested at late stages of logarithmic growth for analysis (OD_{600} ~0.8) from a wild-type strain and a □*sinR::neoR* deletion strain (Materials and Methods). RT-PCR demonstrated that *mstX* was expressed during logarithmic growth phase in the *sinR* mutant but not in the wild-type strain (Figure 5B). In the wild-type strain, no *mstX* transcripts were detected during logarithmic growth. These results indicate that SinR represses *mstX* expression during growth and that the alleviation of SinR repression induces *mstX* expression during biofilm formation.

To determine if SinR directly binds the *mstX* promoter, we constructed a strain with a FLAG-tagged copy of SinR for chromosome immunoprecipitation experiments (ChIP). A copy of *sinR-FLAG* was cloned in pMUTIN4 and integrated into the chromosome through single recombination event and erythromycin selection. ChIP experiments were performed with samples harvested during logarithmic growth, when *mstX* is expressed at low levels. Samples were formaldehyde crosslinked and purified on a FLAG affinity column and the regions of interest amplified by PCR (~200 bp target sequence). The PCR primers were designed to amplify the *mstX-yugO* promoter region, as well as two flanking genes, *yugP* and *yugR* (Figure 5C). PCR failed to detect an interaction between *yugP* and SinR, but did reveal an interaction near the *mstX-yugO* promoter region and within *yugR*. The inability of SinR-FLAG to bind

yugP sequence demonstrates the specific binding of our protein construct. Thus, SinR appears to have multiple binding sites near the *mstX-yugO* region.

5. Potassium and a *kinC* mutant abrogate biofilm formation in a MstX-YugO expression strain

Potassium efflux has previously been shown to induce biofilm formation in *B. subtilis* by activating the Spo0A kinase KinC, suggesting that MstX and the YugO K+ efflux pump might induce biofilm formation by activating KinC. To test this hypothesis, we first examined if KinC is necessary for biofilm formation when *mstX* is overexpressed in LB plates. The wild-type NCIB3610 strain does not form robust biofilms when grown in LB media (Figure 6). However, when *mstX* is introduced, the strain can form a solid pellicle when induced. Deletion of *kinC* or addition of KCl disrupted the biofilm phenotype of a *mstX* overexpression strain grown in microtitre plates (Figure 6). Minimal pellicle was formed compared to the wild-type strain or the strain without the *kinC* mutation (Figure 6). This result indicates that KinC and KCl are importantfor *mstX*-mediated biofilm formation and that *mstX* principally promotes biofilm formation through this pathway.

After examining the influence of *kinC* and KCl on *mstX*-mediated biofilm formation, we then sought to determine their interactions with SinR under limiting biofilm growth conditions. The introduction of *sinR* nullified the impact of a *kinC* mutation or the addition of KCl, implying that these effectors of biofilm growth act at least in part through direct or indirect inhibition of SinR. SinR is the master regulator

for biofilm formation and a dominant player in the transition from planktonic to sessile states in *B. subtilis* [11]. Thus, it is probable that *mstX* acts at least in part by derepressing SinR during biofilm formation.

The ability of certain small molecules to induce biofilm formation by stimulating K+ efflux is abrogated by high concentrations of extracellular potassium. Lopez and colleagues (2009) identified the kinase KinC as a critical component to potassium efflux biofilm assembly when subjected to small molecules that contribute to potassium leakage [15]. We proposed that biofilm induction by overexpression of *mstX* might also depend upon extracellular concentrations of potassium and that the addition of potassium would suppress the biofilm phenotype in an *mstX*-induced biofilm strain. To test this hypothesis, we grew an *mstX* overexpression strain in the presence of 150 mM KCl. In microtitre plates, both the wild-type and *mstX* overexpression strains failed to form robust pellicles in the presence of potassium (Figure 6). KCl at 150 mM KCl does not inhibit inhibit the growth of B. subtilis but does perturb biofilm formation (Supplementary Figure 2) Thus, our results suggest that the effect of MstX and YugO on biofilm assembly is mediated by their ability to activate KinC, which in turn activates Spo0A to mediate the inhibition of SinR and the onset of biofilm gene expression. These findings further suggest that MstX promotes the assembly of active YugO. In this model, YugO would mediate K+ efflux, thereby activating KinC and leading to Spo0A-P production and biofilm assembly (Figure 7).

D. Discussion

The data presented here indicate that the *mstX-yugO* operon, which was identified in a screen for putative potassium channels in *B. subtilis*, participate in biofilm formation. We previously noted that including *mstX* in overexpression plasmids greatly facilitated expression of the downstream integral membrane protein YugO [32] and that *mstX* family members displayed the ability to increase expression of multiple heterologous membrane proteins in *E. coli* when expressed as an N-terminal fusion protein to a cargo membrane protein [20]. MstX possesses no homology to proteins of known function, leaving open the question as to how MstX might function in *B. subtilis.* Our findings suggest that MstX and the putative potassium channel YugO positively regulate biofilm assembly, in a pathway that depends on the key regulatory protein SinR.

The transcription factor SinR plays a critical role in regulating the transition from motile, logarithmic growth to sessile biofilms or pellicle structures in *B. subtilis* populations. SinR manages this switch in cellular fate by repressing the expression of genes required for the synthesis of the biofilm matrix. The activity of SinR is coordinated by antagonistic regulatory interactions between SinR and the antirepressors SinI or SlrR, with *slrR* also representing a target for repression by SinR (Figure 7). SlrR serves an additional function as a repressor of σ^D-dependent genes for autolysis and motility, including the *hag* gene [33, 34]. Expression of SinR and its antagonist, SinI, is activated by the phosphorylated master regulator of sporulation, Spo0A. Spo0A is bistable and only phosphorylated in a fraction of the cell

population, thereby influencing motility and biofilm formation by regulating the activity of SinR through an ON and OFF switch [33]. As conditions trigger an accumulation of phosphorylated Spo0A, an increasing number of cells repress SinR activity via antagonistic interactions with the antirepressors SinI and SlrR [35]. The consequent cellular inhibition of SinR enables the expression of genetic components that ultimately give rise to a biofilm.

Our findings implicate MstX and YugO in a positive autoregulatory cascade involving potassium, KinC activation, Spo0A phosphorylation and the alleviation of SinR mediated repression of biofilm formation in *B. subtilis*. Specifically, we demonstrate that *mstX* overexpression induces biofilm formation in a manner dependent on K+, YugO and KinC functionality. Furthermore, disruption of *mstX* or *yugO* abrogated biofilm formation in a manner that was suppressed by a *sinR* mutation, suggesting that MstX and YugO stimulate biofilm formation by inhibiting SinR, which represses biofilm formation. The *mstX* mutation also caused decreased expression of the *abbA* and *sinI* antirepressors. This suggests that in the absence of MstX, cells fail to accumulate sufficient SinI to overcome the SinR-mediated repression of biofilm formation.

This pathway is likely dependent upon potassium signaling either through MstX or, more likely, through a putative potassium ion channel encoded by the gene that lies downstream of *mstX*, *yugO* (Figure 1). The following lines of evidence support the hypothesis that MstX and YugO work together in potassium regulation: (i) disruption of *yugO* also generates a strain with similar biofilm phenotype as the *mstX*

deletion strain, (ii) a *yugO* knockout, a *kinC* knockout, or elevated KCl concentrations ablate biofilm formation in an xylose-induced P_{xyl}-*mstX* strain (Figure 6), (iii) *yugO* encodes a putative TrkA-domain containing potassium ion channel that was initially identified in a bioinformatic screen for potential potassium ion channels in *B. subtilis* [12] and it possesses sequence similarities to KefC, a widely distributed glutathione-S-conjugate gated K+ efflux system [36-39]. It is unclear whether MstX contributes to enhanced KinC activation through non-surfactant-mediated K+ ion leakage or via allosteric activation of the kinase through production of a small-molecule enhancer [40]. However, the minimal decrease in expression of *kinC* in a *mstX* deletion mutant and the phenotypic dependence on *yugO* point to a connection with its downstream gene, *yugO*. We therefore propose that MstX primarily operates by enhancing the functional expression of *yugO* (likely at the stage of membrane insertion [19]), facilitating potassium ion leakage through the putative YugO ion channel, activating KinC and biofilm formation.

Two previous reports have indicated the important role played by potassium in biofilm assembly in *B. subtilis*. First, Lopez and colleagues [15] identified potassium leakage and KinC activation as a critical regulator of biofilm formation in *B. subtilis.* In this pathway, Surfactin contributes to membrane permeabilization, potassium ion leakage and KinC activation responsible for biofilm formation. Deletion of the *srfA* genes required for surfactin production or the *kinC* gene that encodes the KinC kinase abrogated biofilm formation in *B. subtilis*. Second, Lopez and colleagues identified the potassium ion transport regulator KtrC as being instrumental in biofilm regulation

included as part of a supplemental finding [15]. Deletion of *ktrC* gave rise to increased biofilm formation in a microtitre plate assay, as one would expect with elevated intracellular concentrations of K+ and decreased KinC phosphorylation of Spo0A (see supplementary section, [15]).

Potassium ion channels are widely distributed among prokaryotic species and have been used as models for understanding potassium signaling in the excitation of nerves or muscles of multicellular organisms [30]. Core features of potassium channels, including gating principles and the structure of the selectivity filter, are highly conserved over evolutionary space. The structures and activities of potassium ion channels evolved long before the emergence of complex multicellular organisms and their use in neurophysiology. However, the roles of selective cation channels in microbial physiology remain largely unknown. Our finding ascribes a known physiological function to a potassium ion channel in a prokaryote. It is unlikely that *B. subtilis* is unique in utilizing potassium efflux for intracellular signaling. Potassium uptake and efflux systems have also been implicated as critical regulators for biofilm formation and pathogenesis in *Pseudomonas aeruginosa* [41]. In addition, the KefC system couples glutathione-adduct formation to cytoplasmic acidification by potassium transport, thereby protecting *Escherichia coli* from electrophilic attack [39]. The potential for ligand-gated ion channel transport to serve as a means for coordinating complex and spontaneous signal transduction throughout a cell without the necessity of protein synthesis provides a tantalizing explanation for the prevalence of cation channels in both prokaryotic and eukaryotic species.

The data presented in this study demonstrate that MstX and YugO influence the expression of key antirepressors involved in modulating the activity of the master regulator for biofilm formation, SinR and that the *mstX-yugO* operon is also regulatd by SinR. We find that *mstX-yugO* is selectively expressed during biofilm formation and that SinR directly represses *mstX* transcription during growth. We also show that KinC and YugO are required for facilitating biofilm assembly via MstX-mediated SinR derepression. Thus, the ability of MstX and YugO to induce biofilm formation via KinC and SinR together with the repression of MstX by SinR would provide the basis for positive autoregulation and enhancement of the regulatory cascade involved in biofilm formation (Figure 7).

Previously it has been shown that bistable gene expression can be influenced by any mutation that disrupts the regulators that control the transcription of the bistable target genes. One example is increased expression of the *sigD* gene results in the accumulation of active σ^D. This is likely a product of an inadequate amount of the anti-sigma factor FlgM to inactivate the σ^D due to the shift in stoichiometry. As a result, expression of the σ^D regulon no longer exhibits dispersed bistability but occurs in all cells of the population [28, 34]. Similarly, a mutation in the putative phosphodiesterase *ymdB* shifts the stoichiometry of c-di-GMP toward a motile state and impairs biofilm formation [28]. MstX expression might operate in a similar fashion by increasing the expression of the antagonist SinI, thus decreasing the activity of the bistable regulator SinR.

The application of MstX to membrane protein overexpression in *Escherichia coli* has enabled the production of many eukaryotic membrane proteins and other challenging bacterial proteins [19, 42, 43]. Nevertheless, a complete explanation describing the mechanism that enables MstX-tagged overexpression of otherwise toxic membrane proteins in *Escherichia coli* or MstX-mediated biofilm induction in *B. subtilis* remains elusive. Preliminary structural analysis indicates that the oligomeric state of MstX may be instrumental for its membrane protein chaperoning properties [29]. MstX homologues in *Bacillus mojavensis*, *Bacillus licheniformis*, and *Bacillus atrophaeus* present different solubility profiles than *Bacillus subtilis* yet still facilitate membrane protein overexpression in *E. coli* [20, 29, 32]. In all cases, MstX homologues precede a putative potassium ion channel (*yugO*). It is plausible that MstX facilitates expression of the putative ion channel in these strains, thereby triggering the events that lead to biofilm formation.

In summary, our results identify *mstX* as a novel determinant of biofilm formation and antibiotic resistance in *B. subtilis* that is positively autoregulated through a genetic loop involving potassium, KinC, Spo0A activation, and SinR-mediated derepression. *MstX* expression was necessary and capable of inducing biofilm formation under nutrient conditions which would normally inhibit biofilm formation, rescuing an atavistic response in a domesticated *B. subtilis* strain. The addition of potassium or deletion of *yugO*, a putative potassium ion channel, abrogated *mstX*-mediated biofilm formation, highlighting the importance of potassium homeostasis for initiation of regulatory networks involved in development. Our

finding raises the prospect for a broader role for potassium ion channel signaling in microbial physiology.

References

1. O'Toole, G.A., *Microbiology: Jekyll or hide?* Nature, 2004. **432**(7018): p. 680-1.

2. Kolter, R. and E.P. Greenberg, *Microbial sciences: the superficial life of microbes.* Nature, 2006. **441**(7091): p. 300-2.

3. Branda, S.S., et al., *A major protein component of the Bacillus subtilis biofilm matrix.* Mol Microbiol, 2006. **59**(4): p. 1229-38.

4. Lewis, K., *Persister cells, dormancy and infectious disease.* Nat Rev Microbiol, 2007. **5**(1): p. 48-56.

5. Mah, T.F., et al., *A genetic basis for Pseudomonas aeruginosa biofilm antibiotic resistance.* Nature, 2003. **426**(6964): p. 306-10.

6. Schumacher, M.A., et al., *Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB.* Science, 2009. **323**(5912): p. 396-401.

7. Gefen, O., et al., *Single-cell protein induction dynamics reveals a period of vulnerability to antibiotics in persister bacteria.* Proc Natl Acad Sci U S A, 2008. **105**(16): p. 6145-9.

8. Branda, S.S., et al., *Fruiting body formation by Bacillus subtilis.* Proc Natl Acad Sci U S A, 2001. **98**(20): p. 11621-6.

9. Hamon, M.A. and B.A. Lazazzera, *The sporulation transcription factor Spo0A is required for biofilm development in Bacillus subtilis.* Mol Microbiol, 2001. **42**(5): p. 1199-209.

10. Hamon, M.A., et al., *Identification of AbrB-regulated genes involved in biofilm formation by Bacillus subtilis.* Mol Microbiol, 2004. **52**(3): p. 847-60.

11. Chu, F., et al., *Targets of the master regulator of biofilm formation in Bacillus subtilis.* Mol Microbiol, 2006. **59**(4): p. 1216-28.

12. Chu, F., et al., *A novel regulatory protein governing biofilm formation in Bacillus subtilis.* Mol Microbiol, 2008. **68**(5): p. 1117-27.

13. Blair, K.M., et al., *A molecular clutch disables flagella in the Bacillus subtilis biofilm.* Science, 2008. **320**(5883): p. 1636-8.

14. Kobayashi, K., *SlrR/SlrA controls the initiation of biofilm formation in Bacillus subtilis.* Mol Microbiol, 2008. **69**(6): p. 1399-410.

15. Lopez, D., et al., *Structurally diverse natural products that cause potassium leakage trigger multicellularity in Bacillus subtilis.* Proc Natl Acad Sci U S A, 2009. **106**(1): p. 280-5.

16. Chen, Y., et al., *A Bacillus subtilis sensor kinase involved in triggering biofilm formation on the roots of tomato plants.* Mol Microbiol, 2012. **85**(3): p. 418-430.

17. Rubinstein, S.M., et al., *Osmotic Pressure Can Regulate Matrix Gene Expression in Bacillus subtilis.* Mol Microbiol, 2012.

18. McLoon, A.L., et al., *Spatial regulation of histidine kinases governing biofilm formation in Bacillus subtilis.* J Bacteriol, 2011. **193**(3): p. 679-85.

19. Roosild, T.P., et al., *NMR structure of Mistic, a membrane-integrating protein for membrane protein expression.* Science, 2005. **307**(5713): p. 1317-21.

20. Roosild, T.P., et al., *Characterization of the family of Mistic homologues.* BMC Struct Biol, 2006. **6**: p. 10.

21. Youngman, P.J., J.B. Perkins, and R. Losick, *Genetic transposition and insertional mutagenesis in Bacillus subtilis with Streptococcus faecalis transposon Tn917.* Proc Natl Acad Sci U S A, 1983. **80**(8): p. 2305-9.

22. Wach, A., *PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in S. cerevisiae.* Yeast, 1996. **12**(3): p. 259-65.

23. Becker, E., et al., *DNA segregation by the bacterial actin AlfA during Bacillus subtilis growth and development.* EMBO J, 2006. **25**(24): p. 5919-31.

24. Zheng, L., U. Baumann, and J.L. Reymond, *An efficient one-step site-directed and site-saturation mutagenesis protocol.* Nucleic Acids Res, 2004. **32**(14): p. e115.

25. Ben-Yehuda, S., D.Z. Rudner, and R. Losick, *RacA, a bacterial protein that anchors chromosomes to the cell poles.* Science, 2003. **299**(5606): p. 532-6.

26. Locke, J.C., et al., *Stochastic pulse regulation in bacterial stress response.* Science, 2011. **334**(6054): p. 366-9.

27. D'Angelo, M.A., et al., *Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells.* Cell, 2009. **136**(2): p. 284-95.

28. Diethmaier, C., et al., *A novel factor controlling bistability in Bacillus subtilis: the YmdB protein affects flagellin expression and biofilm formation.* J Bacteriol, 2011. **193**(21): p. 5997-6007.

29. Dvir, H., et al., *Mistic: cellular localization, solution behavior, polymerization, and fibril formation.* Protein Sci, 2009. **18**(7): p. 1564-70.

30. O'Toole, G.A. and R. Kolter, *Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development.* Mol Microbiol, 1998. **30**(2): p. 295-304.

31. Kearns, D.B., et al., *A master regulator for biofilm formation by Bacillus subtilis.* Mol Microbiol, 2005. **55**(3): p. 739-49.

32. Dvir, H. and S. Choe, *Bacterial expression of a eukaryotic membrane protein in fusion to various Mistic orthologs.* Protein Expr Purif, 2009. **68**(1): p. 28-33.

33. Kearns, D.B. and R. Losick, *Cell population heterogeneity during growth of Bacillus subtilis.* Genes Dev, 2005. **19**(24): p. 3083-94.

34. Cozy, L.M., et al., *SlrA/SinR/SlrR inhibits motility gene expression upstream of a hypersensitive and hysteretic switch at the level of sigma(D) in Bacillus subtilis.* Mol Microbiol, 2012. **83**(6): p. 1210-28.

35. Chai, Y., et al., *Bistability and biofilm formation in Bacillus subtilis.* Mol Microbiol, 2008. **67**(2): p. 254-63.

36. Ferguson, G.P., D. McLaggan, and I.R. Booth, *Potassium channel activation by glutathione-S-conjugates in Escherichia coli: protection against methylglyoxal is mediated by cytoplasmic acidification.* Mol Microbiol, 1995. **17**(6): p. 1025-33.

37. Martinac, B., Y. Saimi, and C. Kung, *Ion channels in microbes.* Physiol Rev, 2008. **88**(4): p. 1449-90.

38. Munro, A.W., et al., *The cloning and DNA sequence of the gene for the glutathione-regulated potassium-efflux system KefC of Escherichia coli.* Mol Microbiol, 1991. **5**(3): p. 607-16.

39. Roosild, T.P., et al., *KTN (RCK) domains regulate K+ channels and transporters by controlling the dimer-hinge conformation.* Structure, 2009. **17**(6): p. 893-903.

40. Burkholder, W.F., I. Kurtser, and A.D. Grossman, *Replication initiation proteins regulate a developmental checkpoint in Bacillus subtilis.* Cell, 2001. **104**(2): p. 269-79.

41. Ueda, A. and T.K. Wood, *Potassium and sodium transporters of Pseudomonas aeruginosa regulate virulence to barley.* Appl Microbiol Biotechnol, 2008. **79**(5): p. 843-58.

42. Liu, X., et al., *Characterization of mitochondrial trifunctional protein and its inactivation study for medicine development.* Biochim Biophys Acta, 2008. **1784**(11): p. 1742-9.

43. Kefala, G., et al., *Application of Mistic to improving the expression and membrane integration of histidine kinase receptors from Escherichia coli.* J Struct Funct Genomics, 2007. **8**(4): p. 167-72.

44. Kinsinger, R.F., et al., *Genetic requirements for potassium ion-dependent colony spreading in Bacillus subtilis.* J Bacteriol, 2005. **187**(24): p. 8462-9.

TABLES

Table 3.1 *Bacillus subtilis* strains

Strain name	**Genotype**
MEL63	PY79
MEL64	PY79 sac::(PspoIID-cre Δ*mstX::loxP*
MEL65	NCIB3610
MEL66	PY79 Δ*mstX, lacA*::P_{spac}-*mstX* ::*erm*
MEL67	PY79 Δ*mstX, lacA*::P_{spac}-*mstX* M75A::*erm*
MEL73	PY79 Δ*sinR::neo*
MEL102	PY79 *MUTIN4-sinR-FLAG ::erm*
MEL110	PY79 Δ*mstX::loxP-kan-loxP*
MEL111	PY79 Δ*yugO::kan*
MEL112	PY79 Δ*kinC::spc*
MEL212	PY79 Δ*sinR::spc*
MEL217	PY79 Δ*kinC::cam*
MEL218	PY79 Δ*sinR::spc*
MEL239	NCBI3610 Δ*yugO::kan*
MEL240	NCBI3610 Δ*mstX::kan*
MEL420	PY79 *amyE::*P_{xyl}*-mstX::spc* Δ*mstX::loxP*
MEL421	NCIB3610 *amyE::*P_{xyl}*-mstX:spc* Δ*yugO::kan*
MEL422	NCIB3610 *amyE::*P_{xyl}*-mstX::spc* Δ*mstX::kanR*
MEL423	NCIB3610 Δ*sinR::spc*
MEL424	NCIB3610 Δ*sinR::spc* Δ*yugO::kan*
MEL425	NCIB3610 Δ*sinR::spc* Δ*mstX::kan*
MEL426	NCIB3610 Δ*kinC::spc*
MEL427	NCIB3610 Δ*kinC::cam*
MEL428	NCIB3610 Δ*sinR::spc* Δ*kinC::cam*
MEL429	NCIB3610 *amyE::*P_{xyl}*-mstX::spc* Δ*mstX::kan* Δ*kinC::cm*
MEL430	NCIB3610 *amyE::*P_{xyl}*-yugO::spc* Δ*yugO::kan*
MEL431	NCIB3610 *amyE::*P_{xyl}*-yugO::spc* Δ*mstX::kan*
MEL432	PY79 *amyE::*P_{xyl}*-yugO::spc*

Table 3.2 Primers used in this study

Primer name	**Sequence (5'-3')**
ChIP mstX (-2000)-(-1773) forward	TGAATCCGCTCCTTCTGCATGCGGATTAAAAGC
ChIP mstX (-223)-(0) forward	GCCTGCTGAAAACAAATAAAAG
ChIP mstX (1752)-(2000) forward	GGCGTCAATCCGTTTGACCAGCCT
MUTIN4-SinR-FLAG forward	AAAGAATTCTCATTTATCATCATCATCTTTATAATCTTGATTGGCCAGCGTATTAAACAATACC
MUTIN4-SinR-FLAG reverse	AAAGGATCCCTCCTCTTTTTGGGATTTTCTCCA
ChIP mstX (-2000)-(-1773) reverse	AGCAGTCGCTTGAAGAGGCAATTCATC
ChIP mstX (-223)-(0) reverse	ATAAAACCGGATTTCTCCATA
ChIP mstX (1752)-(2000) reverse	ATCAAAGGAGATAAGGAAACTTAAATTAG
qPCR KinC Forward	ACTGCTAAAACCGCCTCTCA
qPCR KinC Reverse	TGCCTTTTCCAGCTTCACTT
qPCR SinI Forward	TGCAAAACAAGAGCACTTTGA
qPCR SinI Reverse	GGACCAGGATGAGCAGACTT
qPCR AbbA (ykzF) Forward	TGCCATCGAGCTATTAAGCA
qPCR AbbA (ykzF) Reverse	TTCAAATCGAAATCGGTCGT
qPCR yqxM Forward	AGGCGAAGACGAAACTGAAA
qPCR yqxM Reverse	TGAGCATACAGTGCCTTTCG
qPCR mstX Forward	ATCGACCGAATGAATGAAGG
qPCR mstX Reverse	GCTTTTCCTGGCCGTACATA
qPCR veg Forward	TCGCTTGATGGGAATTTAGG
qPCR veg Reverse	CGGCGTCATCATTAAATGTC

mstX-US-F1 5'-GTAAAACGACGGCCAGTGAATTCAATCCGTCTTGC-3'
mstX-US-R1 5'-caactgatcgatgaaccttct-3'
mstX-DS-F1 5'-gtaaagcttaaggaggaactactatgcgggaagagctaaaggg-3'
mstX-DS-R1 5'-CCCGGGTACCGAGCTCGAATTCTGCCGGTTTTTC-3'
yugO-US-F1 5'-
CGTTGTAAAACGACGGCCAGTGAATTCAATCCGTCTTGCCGTTTCTGCCCCTGT -3'
yugO-US-R1 5'-
TACTGTTTATGTAAGCAGACACTATTCTTTTTCTCCTTCTTCAGATACTGAGAT-3'
yugO-DS-F1 5'-
TTGATGCTCGATGAGTTTTTTCTAAGTGTAACAGAACATCCTCCAATTCTTTC -3'
yugO-DS-R1 5'-
GGATCCCCGGGTACCGAGCTCGAATTCTGCCGGTTTTTCTGCAGCTATTTTAAG -3'
kinC-US-F1 5'-GTATATTAGCTGTTCCTAAGTCAATCCCGATTTCAG -3'
kinC-US-R1 5'-TTCTCATATGACCACCTGCCGCTTGTGT -3'
kinC-DS-F1 5'-CAGCGAGAGGAGAAAAATAAAGTG -3'
kinC-DS-R1 5'-GCTGTCTGATTTTAAAGGCAAACCACTTGAAAG -3'
kanR-F1 5'-TGTCTGCTTACATAAACAGTAATAATACAAGGG -3'
kanR-R1 5'-TTAGAAAACTCATCGAGCATCAAATGAAACTG-3'
P_{xyl} SG1154 -F 5'-AAGCTTGATATCGAATTCCTGCAGATG -3'
P_{xyl} SG1154 -R 5'-CATCCTAGGAATCTCCTTTCTAGA -3'
mstX(M75A)-F1 5'- ATGATGATACAGCCGAGTTAGCCGAAGCAGGCCCGAGATAT-3'
mstX(M75A)-R1 5'- GTACATATCTCGGGCCTGCTTGGCTAACTCGGCTGTATCATCAT-3'

FIGURES

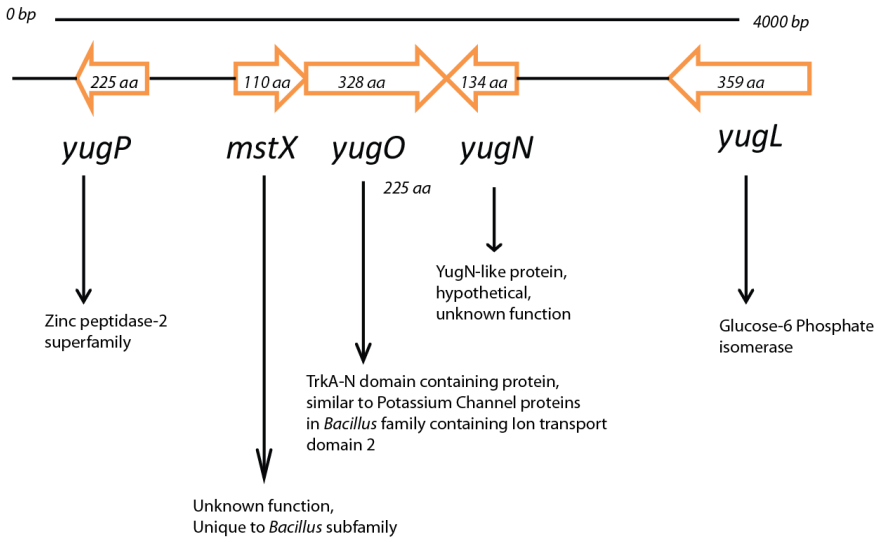


Figure 3.1 Diagram of the genetic organization of the chromosomal regions surrounding *mstX* and *yugO*. Sequence homology was identified by BLAST.

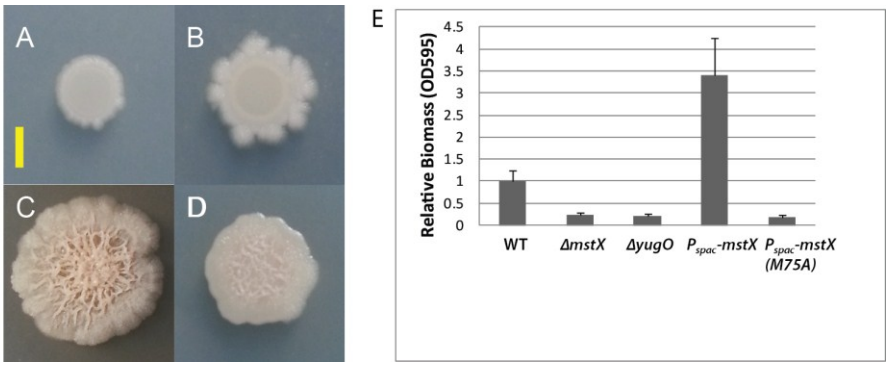


Figure 3.2 Alterations in colony morphology and biofilm formation related to *mstX* expression in the domesticated strain PY79. Images show colony morphology after 1 days of growth on MSgg medium at 30°C. Scale bar corresponds to approximately 3 mm.

(A) Colony morphology of *B. subtilis* PY79

(B) Colony morphology of *B. subtilis* PY79 Δ*mstX* (MEL64)

(C) Colony morphology of *B. subtilis* PY79 domesticated strain after IPTG induction of *mstX* (*lacA::*P_{spac}*-mstX::erm;* MEL66)

(D) Colony morphology of *Bacillus subtilis* PY79 after IPTG induction of *mstX M75A* (*lacA::*P_{spac}*-mstX M75A::erm*; MEL67)

(E) Microtitre crystal violet staining assay for WT, Δ*mstX*, *lacA::*P_{spac} *–mstX::erm*, and *lacA::*P_{spac} *-mstX (M75A)::erm* strains (strains PY79, MEL64, MEL66 and MEL67). Error bars represent standard error calculated from three independent experiments.

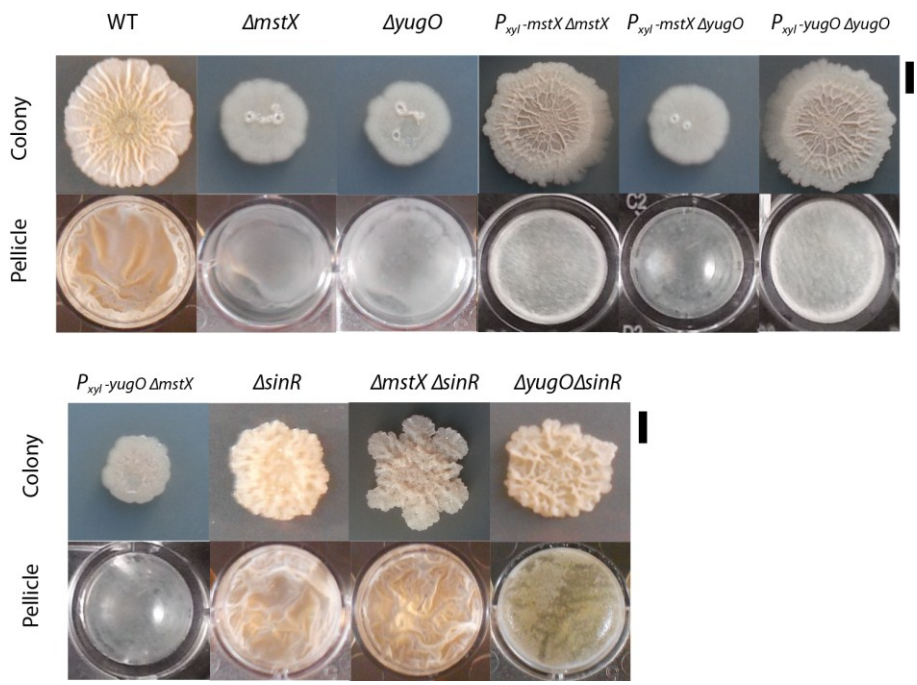


Figure 3.3 The *mstX* and *yugO* genes regulate biofilm formation in undomesticated *Bacillus subtilis* NCIB3610. Top rows of images show colony morphology after 3 days of growth on MSGG medium at 22°C; bottom rows show pellicle formation in MSGG medium after 72 hours of growth at 22°C. Strains used include NCIB3610 (wt), MEL240 (Δ*mstX::kan*), MEL239 (Δ*yugO::kan)*, MEL422 (Δ*mstX::kan, amyE::*P_{xyl} *–mstX::spc)*, MEL421 (Δ*yugO::kan, amyE::*P_{xyl} *–mstX::spc)*, MEL218 (Δ*sinR::spc*), MEL425 (Δ*mstX::kan*, Δ*sinR::spc*), MEL424 (Δ*yugO::kan*, Δ*sinR::spc*), MEL430 (NCIB3610 P_{xyl}*-yugO* Δ*yugO::kan ::spc*), and MEL431 (NCIB3610 *amyE::*P_{xyl}*-yugO::spc* Δ*mstX::kan*). The Δ*mstX* and Δ*yugO* mutations reduce colony architecture and pellicle formation, which is rescued by the *sinR* mutation. Microtitre wells measure approximately 3 cm in diameter.

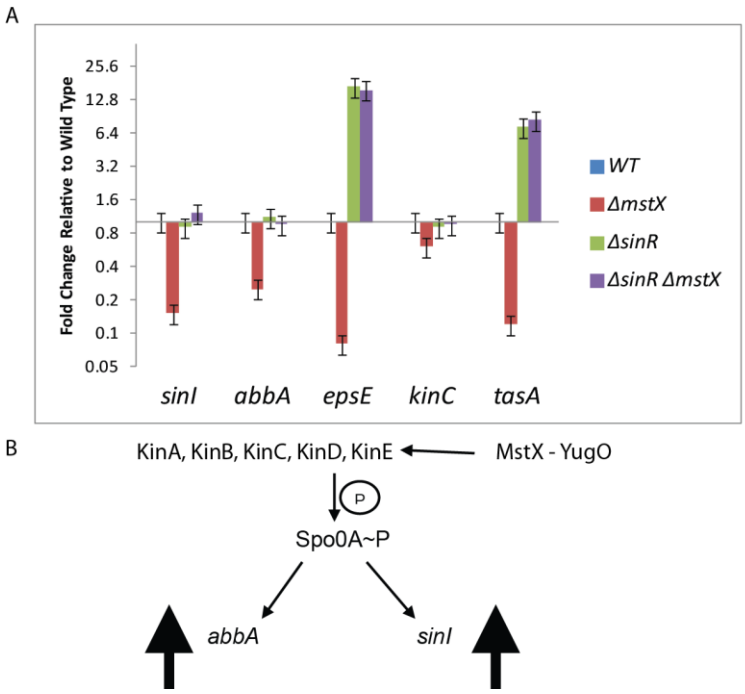


Figure 3.4 MstX negatively regulates parallel antirepressors involved in biofilm formation

(A) Quantitative RT-PCR analysis of *abbA, sinI, epsE, kinC,* and *tasA* in wild type, Δ*mstX*, Δ*sinR*, and Δ*sinR ΔmstX* double mutants (strains MEL65, MEL240, MEL423 and MEL 425) were grown in MSgg medium at 30°C and collected at 0.5-0.8 OD_{600}. Error bars represent standard error calculated from three independent experiments.

(B) Model for *mstX* activation of *kinC* with corresponding increases in *abbA* and *sinI* transcription.

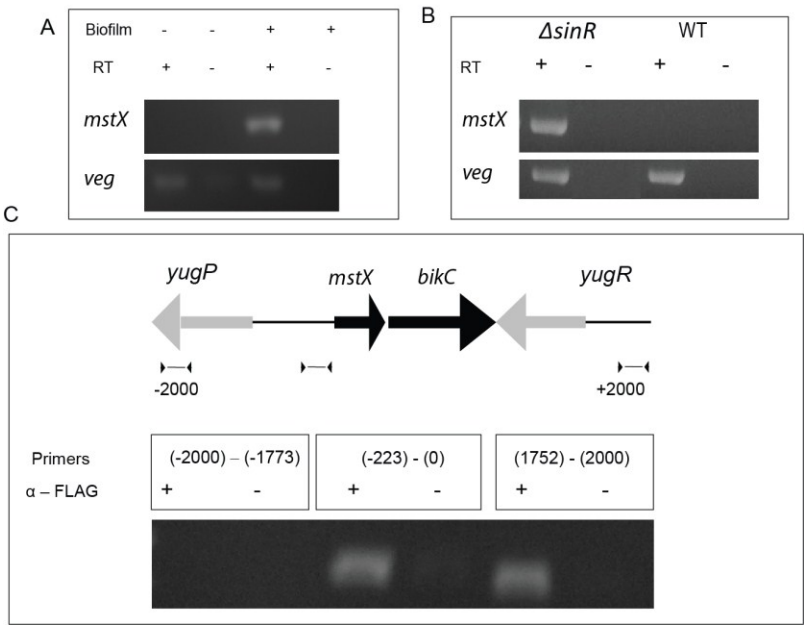


Figure 3.5 Temporal expression of *mstX* and regulation by the transcriptional repressor SinR.

(A) Comparison of *mstX* expression during log growth and biofilm growth using RT-PCR analysis in strain MEL63. The *veg* gene (BSU00440) is a constitutively expressed gene frequently used as a positive control for RT-PCR.

(B) Comparison of *mstX* gene expression in the presence and absence of the transcriptional repressor SinR using RT-PCR in strains MEL63 and MEL73. Cultures were collected during log growth.

(C) Chromatin immunoprecipitation assay of strain MEL102 shows that SinR-FLAG binds near the *mstX* promoter during exponential growth.

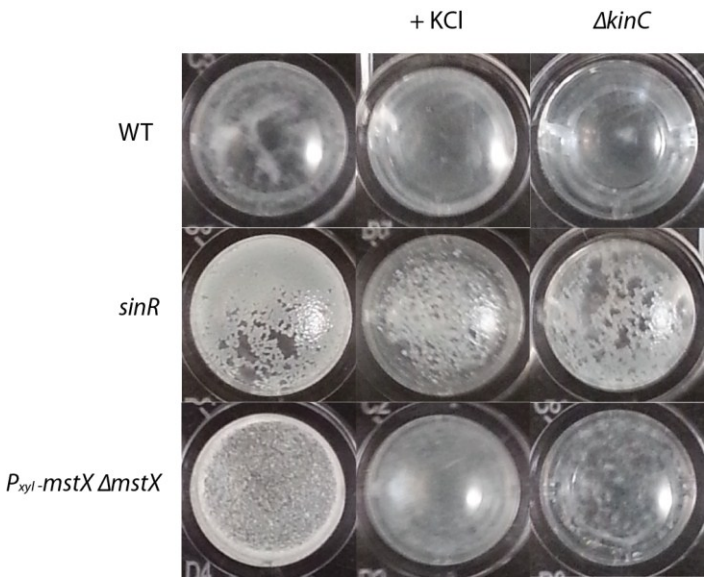


Figure 3.6 Addition of 150 mM potassium chloride or the *kinC* deletion abrogates pellicle formation after *mstX* overexpression and in the *sinR* strain in LB medium that does not normally support biofilm formation in the NCIB3610 background strain. Strains used includes MEL65 (wt), MEL423 (Δ*sinR*) and MEL428 (Δ*sinR, ΔkinC),* MEL422 *(amyE::*P_{xyl}*-mstX::spc)* and MEL429 (*amyE::*P_{xyl}*-mstX, ΔkinC::cm).* Xylose-inducible strains were grown in the presence of 0.5% xylose.

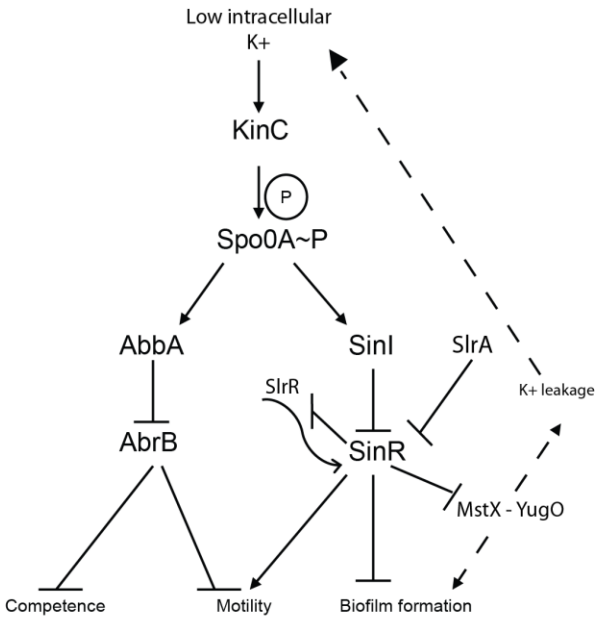
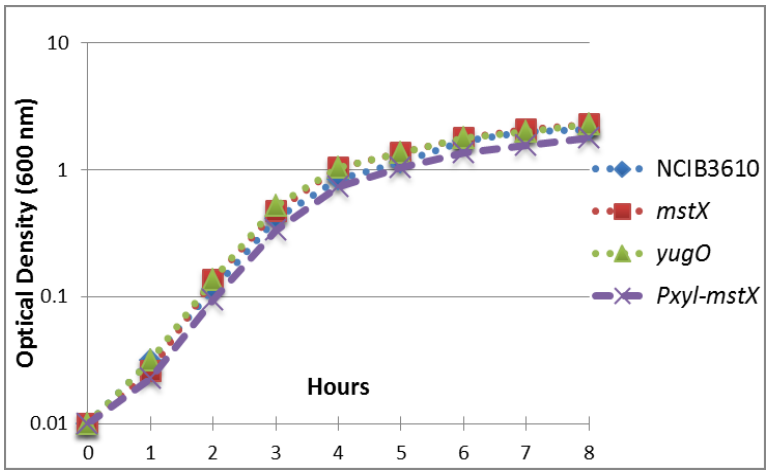


Figure 3.7 A positive autoregulatory loop involving MstX, YugO, potassium, and biofilm formation in *B. subtilis*. Our data suggest that MstX and YugO both positively regulate biofilm formation by inhibiting SinR (in a manner dependent on KinC and influenced by potassium) and that the expression of *mstX* and *yugO* are negatively regulated by SinR. This suggests that MstX and YugO participate in a positive feedback loop to lock a subpopulation of cells in the biofilm assembling state. We propose that MstX mediates the assembly of YugO, a putative potassium efflux channel, and that potassium leakage activates KinC [15, 35, 44].

A.



B.

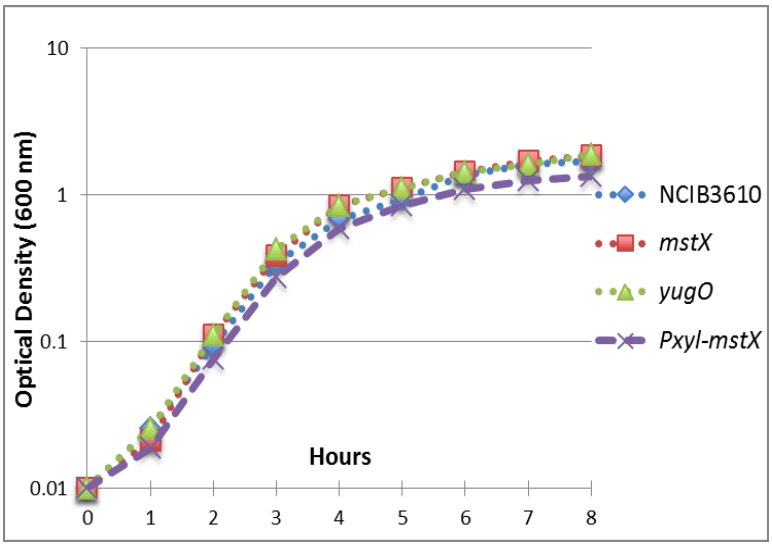


Figure 3.8 Growth of *B. subtilis* mutants in (a) LB or (b) MSgg media. Growth of the following strains was monitored: *B. subtilis* wt (NCIB3610), the *mstX*-deletion strain MEL (*mstX*), the *yugO*-deletion strain MEL (*yugO*), and the *mstX* overexpression strain MEL (P_{xyl}-*mstX*). The wild-type strain and deletion strains were also grown in 0.25% xylose. A volume of 5 ml of xylose-free MSgg or LB medium was inoculated with fresh colonies and incubated overnight at 37°C. Roller flasks containing 5 ml of xylose-containing MSGG or LB were inoculated with diluted aliquots of the overnight culture (start OD600 0.01) and incubated at 37°C. Growth was monitored by optical density measurements. Values represent the mean of three independent trials.

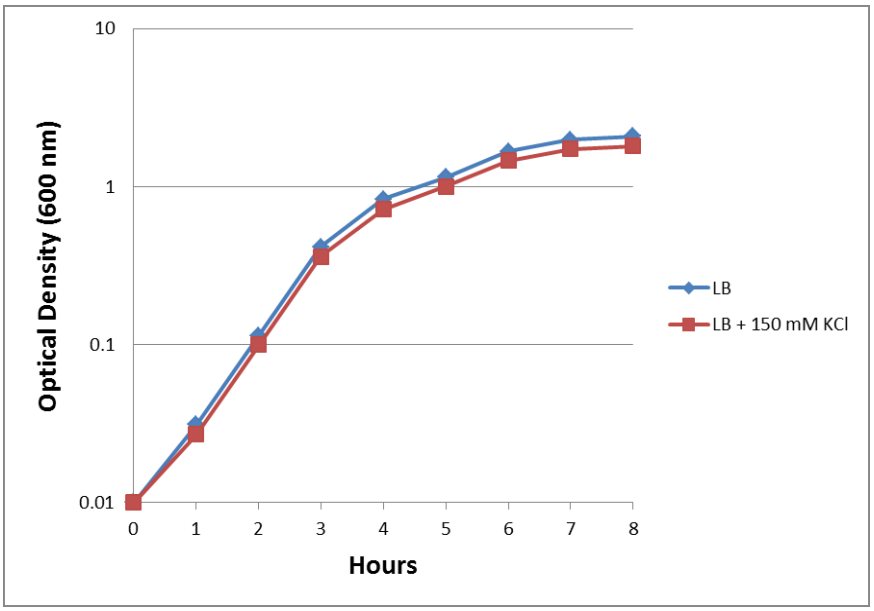


Figure 3.9 Growth of *B. subtilis* NCIB3610 wild-type strain in LB or LB media supplemented with 150 mM KCl. A volume of 5 ml of LB medium was inoculated with fresh colonies and incubated overnight at 37°C. Roller flasks containing 5 ml of LB were inoculated with diluted aliquots of the overnight culture (start OD600 0.01) and incubated at 37°C. Growth was monitored by optical density measurements. Values represent the mean of three independent trials.

ACKNOWLEDGEMENTS

Chapter III, in full, is a manuscript under revision that the dissertation author was the principal researcher and author of. The material is in review for *PLOS ONE*.

Lundberg, M.E., Becker, E.C., and Choe, S. (2013). MstX and a putative potassium channel facilitate biofilm formation in *Bacillus subtilis*. *PLOS ONE, in revision*

CHAPTER IV

Mistic-tagged expression and purification of Claudin family member

A. Introduction and Background

Although structural determination of soluble proteins has become routine, the structure of many integral membrane (IM) proteins have continued to elude researchers due to technical obstacles encountered in obtaining sufficient yields and in purification of the protein so that it may form ordered crystals for x-ray crystallography analysis. Recombinant expression of eukaryotic membrane proteins achieved in *Escherichia coli* presents major limitations in that prokaryotes lack the appropriate cellular translocation and protein modification machinery necessary for proper membrane integration. Furthermore, high-level expression can ultimately prove toxic by overloading the cellular machinery, competitively excluding the production of vital host membrane proteins [1, 2]. Bacterial homologues can serve as suitable models for understanding certain biological processes but are ultimately of limited utility for studying many disease states in humans [3].

These technical obstacles are of seminal importance, effectively hampering studies that might otherwise elucidate fundamental cellular processes in eukaryotic organisms. To underscore the importance of membrane proteins in dictating cellular events, approximately 60% of all approved drugs with human protein targets are specific for a member of the 7TM super family. Yet in stark contrast to soluble proteins (of which almost forty thousand three dimensional structures have been solved as of 2008, compared to two hundred membrane proteins), only a handful of structures are known for higher eukaryotic membrane proteins [4]. Structural determination of other membrane protein families in eukaryotes is essential for

rational drug discovery even with the exorbitant cost it entails (at least $2.5 million for a solved eukaryotic membrane protein structure). Possessing the structure of a known drug target can accelerate research and reduce costs by as much as fifty percent [5, 6]. Thus, expanding the realm of known eukaryotic membrane structures will directly contribute to the improvement of human health through decreased drug development costs and timetables.

Tight junctions represent a fundamental structure in cell biology that regulates the transepithelial movement of ions and noncharged solutes through a series of intercellular contacts mediated by integral membrane proteins. Disruptions in tight junctions can result in a myriad of diseases, including inflammation, renal failure, and cancer [7, 8]. Moreover, it is thought that many viruses and bacterial pathogens disrupt tight junctions as a means of invading host cells [9, 10]. In transcellular transport, movement across cell membranes generally occurs in an energy-dependent fashion through a large number of integral membrane pumps, channels, and transporters. In contrast, paracellular transport results from the passive movement of material down electrochemical gradients established either by the activity of the transcellular transporters or by externally imposed gradients from ingestion of solutes. Paracellular transport varies in its selectivity for ionic charge and solute size, or permselectivity, with discrimination between cations and anions varying in the range of 30-fold [11]. Permselective gradients establish a diffusion barrier for regulating epithelial flow, with overall magnitude estimated by the transepithelial electrical resistance (TER). Resistance exerted by permselectivity depends greatly on the tissue

type, with differences as high as ~10^5 in tight to leaky junctions [12]. Tight junctions display size discrimination, but it remains unclear whether the size varies among different tissue types.

Tights junctions are comprised of nearly 40 different proteins, including the multigene transmembrane familes claudin, occludin, and junctional adhesion molecules (JAMs). These proteins participate in fulfilling the diverse roles of scaffolding, cytoskeletal attachment, establishment of apical/basal polarity, signaling, and vesicle trafficking that are essential for cell development and survival. Many lines of evidence show that the 24-member protein family of claudins (20-27 kDa) is responsible for creating tight junction contacts and for conferring their permselective properties [13]. Since the discovery of the claudin protein family in 1998, the Tsukita group has shown that claudins are incorporated into tight junction strands when expressed in cultured epithelial cells, can form strands de novo when expressed in fibroblasts, and are cell-cell adhesion molecules [14-16]. Hence, claudins appear to be both sufficient and necessary for forming the paracellular barrier, with occludin and JAMs serving ancillary roles in cellular signaling.

1. Claudin structure and function

Claudin structure and charge polarity relate directly to their permselective properties in mediating the barrier of the transepithelial layer. Hydropathy plots suggest that claudins bear four transmemberane domains and contain a short intracellular N-terminal sequence (~7 residues), a large first extracellular loop (~52

residues), and a shorter second extracellular loop (16-33 residues) and a cytoplasmic C-terminal sequence that varies considerably in length between different isoforms (21-63 residues). The cytoplasmic tails are the most diverse in sequence. The first extracellular loop contains charged amino acids and a potential disulfide bridge for intramolecular stabilization or intercellular oligomerization, with some of the residues conserved in different claudin isoforms. There is considerable evidence that the first extracellular loop lines the paracellular pathway and determines the charge selectivity of paracellular transport [17].

The functions of the second extracellular loop are less well understood. The second extracellular loop may associate with itself, and possess a holding function, narrowing the paracellular cleft. Another theory is that the second extracellular loop forms trans-interactions with other claudins between cells. Studies by Piontek and colleagues on claudin-5-transfected HEK cells suggest a role in the formation of tight junction strands through trans-interactions between cells [18]. Piontek and colleagues hypothesize that the second extracellular loop is folded in a helix-turn-helix motif and forms dimers between claudins of opposing cell membranes through hydrophobic interaction of conserved aromatic residues. The second extracellular loop of claudin-3 and claudin-4 have also been shown to act as a receptor for *Clostridium perfringens* enterotoxin, an attribute that could be used in barrier modulation and drug delivery [19].

All claudins, with the exception of claudin-12, end in PDZ-binding motifs which interact with PDZ domains in the cytoplasmic scaffolding proteins ZO-1, -2,

and -3 [20]. Claudins lacking the last three amino acids or those in which the PDZ-binding sites are blocked by addition of an epitope-tag still localize to cell-cell contacts and form freeze-fracture strands, suggesting that they have an inherent ability to polymerize, independent of PDZ interactions [15]. However, the strands formed by PDZ-blocked claudins are poorly organized and not restricted to the apical pole [21].

Little is known about the function of the cytoplasmic domains in claudin proteins. The last three amino acids are critical for PDZ-dependent interactions and other regions are likely important for targeting or proteolytic degradation. The membrane-proximal region of the C-terminus is palmitoylated on conserved cysteines, and it is likely that most claudins can be phosphorylated on serines or threonines in the cytoplasmic tail [22]. Swapping of tails between different claudins altered their unique protein half-lives [23]. This appears to be a result of different interactions between the tails and their cytoplasmic binding partners. Tight junction associated MAGUKs, ZO-1, ZO-2, ZO-3, MUPP1, and PATJ directly link claudins to the actin cytoskeleton [24].

Claudin expression levels are variable among different cell types, thus drastically altering paracelllular permeability properties across tissue types. For example, some claudins, such as claudin-1, are ubiquitously expressed. Some cell types express only a single claudin, such as sertoli cells which express only claudin-11 [25]. Segment-specific expression of different claudins can occur along the nephron, whose paracellular properties are known to differ in each segment [26]. Claudins can thus confer gradients of resistance and charge selectivity through differential

expression in various tissue types. The physiological significance of differential paracellular permeability remains largely unexplored due to the unknown functions of individual claudin family members. Expression levels can differ by a thousand-fold within a defined segment; the emerging data indicate that there may be "housekeeping" claudins expressed widely and at high levels as well as specialized claudins which introduce specialized permselective properties to distinct tissue types [12].

The size and charge-selective properties of claudin pores remain ill-defined. Experiments involving tracer flux imply that the pore size diameter is within the range of 8-9 Å, which is much larger than transmembrane pores or channel [27]. It is possible that claudins create variable pore sizes. Vascular permeability studies in mice in which claudin-5 was deleted showed an incremental increase in the size of tracers allowed to exit from the vascular space in the brain. This result implies that the remaining claudins formed larger pores in the absence of claudin-5 [28]. It is still unknown whether claudins allow any significant passage of water. Kovbasnjuk and colleagues elegantly demonstrate that there is no detectable paracellular water movement, even in leaky epithelia [29]. Hence, how ionic solutes pass through tight junctions in the absence of a solvent remains controversial. Understanding water passage will require the experimental manipulation of claudins, biophysical studies of claudin structure, and assessment of transjunctional water transport.

In addition, claudins are largely responsible for the differences in electrical resistance across tight junctions. Differences in charges of the first extracellular loop

appear to discriminate between cation and anion selectivity [30]. Tissue and culture cell lines show variable charge selectivity as a result of differential claudin expression. Most epithelia in the body contain a mixture of different types and are cation selective. Charge-reversing mutagenesis of residues in the first extracellular loop shows that permeability for solute ions is increased by introducing opposite charges on the protein [31]. Some charged residues must line the pore, creating the selectivity filter that confers paracellular resistance. The number and combination of claudins, acting in cis or trans-oligomerization states, produces the permselective properties. Increasing evidence points to dynamic organization in the structure of claudin pores, with constant reassembly of pores influencing permselectivity [32]. The physiological consequences of dynamic assembly of claudin pores remain unexplored.

2. Claudins and their contribution to inherited human disease

The importance of claudins in early development and transepithelial barrier maintenance has been examined in mouse models and through categorization of human phenotypes. The effects of mutations in claudin-1 and claudin-11 will be reviewed here. Other claudin knockouts and gene depletion studies display a myriad of phenotypes, including skin barrier defects, size-selective blood-brain barrier defects, CNS myelin defects, loss of endocochlear potential, human deafness, renal failure, and defective skin or bile ducts and are reviewed elsewhere [8, 12].

Claudin-1 is a critical component of the epidermal barrier and an essential co-receptor for HCV entry. Claudin-1 KO mice die shortly after birth from transdermal

water [33]. Prior to this finding, it had been assumed that the skin barrier was formed by the intracellular accumulation of hydrophobic material and the cumulative layers of dead squamous cells. Examination of claudin-1 contacts in wild-type mice show continuous junctional contacts formed by claudin-1 and claudin-4. Furthermore, claudin-1 has been shown to be required for HCV infection of human hepatoma cell lines and is the first factor to confer susceptibility to HCV when expressed ectopically in non-hepatic cells. Residues within the first extracellular loop are critical for HCV entry; an antibody directed against the first extracellular loop was sufficient to block viral entry [9]. As the leading cause of cirrhosis and liver cancer worldwide, HCV and its interactions with claudin-1 represent a new target for antiviral drug development.

Claudin-11 is expressed in many epithelia in the body but is best known for the formation of continuous adhesive contacts around axons in oligodentrocytes. Claudin-11 is expressed normally in the epithelium lining of the endocochlear space and maintains an appropriate environment for the mechanotransduction by hair cells. When claudin-11 is impaired, a loss of electrical potential and absence of compensation by transcellular transport help contribute to deafness. Disruption of claudin-11 abrogates salutatory conduction between nodes. Homozygous null mice display severely delayed CNS conduction velocities and represent a specific defect in permselectivitys [25] Similar phenotypes are also observed in disruptions of claudin-14 [34].

3. **Claudins as potential chemotherapeutic targets**

Claudins represent a tantalizing target for drug development. Although most hydrophobic drugs cross through the transcellular pathway, hydrophilic drugs such as proteins and peptides utilize the paracellular route. Manipulation of tight junctions has long been considered a possible way to enhance transepithelial drug absorption. The large number of claudins, multiple roles in disease, and their restricted locations render them a particularly intriguing target for drug delivery.

In support of this hypothesis, studies have shown that the application of *Clostridia perfringins* enterotoxin, CPE, utilizes a limited number of receptors. CPE binds claudin-3 and claudin-4 in the intestine, inducing cytolysis and diarrhea [35]. Claudin-4 is dramatically upregulated in ovarian and pancreatic cancer, making the receptor a particularly attractive target as a selective chemotherapeutic agent [36]. Claudin-1 represents another attractive drug target in combating HCV due to its essential role as a co-receptor for late viral entry [9].

Attempts have already been made to facilitate drug absorption through targeting occludin, the endogenous zonulin receptor system, or through manipulation of signaling pathways that influence the barrier [37-39]. However, the modulation of tight junctions by absorption enhancers for paracellular drug delivery has been hampered by a lack of the structure and function of tight junctions [40]. A crystal structure of a claudin tight junction assembly would contribute immensely to understanding the permselective properties of the paracellular pathway and for rational chemotherapeutic development.

4. Method for expression, purification, and crystallization of claudin homologues

X-ray crystallography represents the most successful method for determining the three-dimensional structure of membrane proteins. In order to obtain an atomic resolution structure of claudin contacts responsible for creating the paracellular permeability barrier, claudin proteins must be overexpressed in milligram quantities, solubilized in crystallization detergent, and purified to homogeneity prior to crystallization screens. The claudin family represents a particularly attractive candidate for attempts at crystallization due to the following qualities:

1) Relevance of tight junctions and paracellular permeability in medical research and drug delivery.
2) Large size of the claudin family, with 24 family members total, providing many potential crystallization targets for optimization.
3) Ability to express claudin proteins in high yields in both yeast and Sf9 baculoviral cell culture.
4) High solubility in favorable crystallization detergents, including dodecyl maltoside and octyl glucoside.

The tight junction transmembrane protein claudin-4 has been expressed, solubilized in detergents, and biochemically characterized previously [41]. Claudin-4 was expressed robustly in a baculoviral expression system. Immuno-EM examination of claudin-4 expressing cells demonstrated that the protein is concentrated in intracellular membranes, most likely those of the ER or Golgi, and does not appear on the plasma membrane. As a result, the infected insect cells did not display the adhesive qualities conferred by claudin overexpression. Western blot analysis of

hexahistidine tagged claudin-4 displayed a single band at 23 kDa, implying that the protein is purified as a monomer in most detergents.

Claudin-4 solubility was determined by Mitic and colleagues in 39 different detergents, with variable solubility in 33 detergents. In general, ionic and phospholipid-like detergents were most effective at liberating claudin-4 from the membrane. Of interest for structural studies, dodecyl maltoside and octyl glucoside were among the leading candidate. Dodecyl maltoside and octyl glucoside are empirically the two most effective crystallization detergents, as determined previously [42]. Solubilization in perfluoro-octanoic acid and purification by size-exclusion chromatography reveal claudin-4 as a hexamer in quarternary structure. Neither octyl glucoside nor dodecyl maltoside maintained the hexameric arranging, implying that oligomerization state of claudins is fairly dynamic. Claudin-4 thus represents a suitable target for further biophysical characterization.

The structure of claudin-based pores, barriers, and the role of homotypic and heterotypic interactions remain poorly defined. Understanding the biochemical and structural aspects of claudin function in maintaining paracellular barriers would greatly facilitate our understanding of diffusion across different transepithelial layers. The complex expression and regulation of tight junctions provide an intriguing target for specific therapies such as targeted drug delivery and prevention of microbial infections.

B. Results and Discussion

1. Overexpression of claudin homologues in *Escherichia coli*

Claudin family member protein sequence (CLDN1-CLDN25) was cloned from a hIMP library into a 6xhis tagged Mistic gateway expression vector or subcloned from the OriGene library into the same Mistic expression vector, pMis-4.0 [43]. The resulting Claudin clones contain an n-terminal 6xhis Mistic separated by a thrombin cleavage site for subsequent cleavage and purification of the Claudin protein. Claudin proteins were expressed in the specialized C43(DE3)PRIL strain [44, 45]. The use of the C43(DE3) strain proved critical in minimizing the cellular toxicity of human membrane protein expression in a T7 RNA polymerase expression system. Mistic moderately enhanced Claudin expression in some but not all cases.

The pMis-Claudin (ampR) constructs were transformed into C43(DE3)PRIL expression strain and grown to 0.5 OD600 at 37 ˚C in 1 L TB supplemented with 2 mM $MgSO_4$. C43(DE3) PRIL were then induced with 250 µm Isopropyl-beta-D-thiogalactoside (IPTG) and shifted to 20 ˚C for overnight expression. Cells were harvested after 16-20 hours of expression via centrifugation and stored at -80˚C in 50 ml buffer containing 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 40 mM Imidazole, 1 mM PMSF, and 5% Glycerol. The following day, cell pellet was dethawed, Roche protease inhibitor cocktail added, and cell resuspension lysed with 1 mg/ml lysozyme and sonication (5 min pulse on ice). Cell debris was removed by centrifugation at 5,000 x g for 10 min. The resulting membrane suspension was further centrifuged at 100,000 x g for 30 min and resuspended in 50 ml buffer containing 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 40 mM Imidazole, 1 mM PMSF, 5% Glycerol, 1 mM β-ME,

and 5 mM DDM (Buffer A). Detergent extraction proceeded overnight at 6˚C. The following day, the detergent resuspension was centrifuged at 100,000 x g for 30 min and the supernatant was applied to a ~6 ml Ni-NTA slurry equilibrated with 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 40 mM Imidazole, 1 mM PMSF, 5% Glycerol, 1 mM β-ME, and 0.5 mM DDM (Buffer B). Ni-NTA bound protein was washed with 10 column volumes Buffer B and eluted with Buffer B containing 400 mM Imidazole.

A 10 microliter sample was assayed via coomassie stained SDS-PAGE (reducing gel, 4-12% Bis-Tris). Typical Ni-NTA purified expression samples from Claudin-1, Claudin-10, and Claudin-16 are pictured in Figure 1, after DDM extraction. Approximately 2 mg of purified protein was obtained per liter of cell culture, measured by Bradford and OD280. We observed some degradation in some samples (Figure 1D, Figure 2), a result of proteolytic cleavage of the flexible linker to Mistic. An increase in protease inhibitor, changes in the tag construct, or further optimization in purification conditions helped significantly reduce observed proteolysis.

2. Claudin homologues are variably solubilized by multiple detergents

After verifying the successful expression in *E. coli* of Mistic-tagged Claudin family members, we proceeded to optimize the extraction protocol by sampling a series of different detergents. We focused principally on detergents with a significant track record of success in crystallization trials. Most detergents tested extracted Claudin family member proteins with some efficiency, although non-ionic detergents with longer alkyl chains, such as DDM, DM, or NG, tended to be more successful

during our tests (Figure 2). Interestingly, Claudin-10 and other Claudin family members displayed SDS-resistant oligomers, even under reducing conditions (Figure 2 and Figure 6B). These higher order oligomers correspond roughly to the expected molecular weight of a Claudin dimer. It is possible that the β-ME we used during our tests only weakly reduced the expected disulfide bridge that helps form claudin dimers, thus resulting in the higher molecular weight band.

During our trials, we also quantified detergent extraction using c-terminal GFP fusion tagged Claudin-10. GFP-tagging of IMP with cytoplasmic c-terminal tails, such as claudins, has proved useful in following the progression of membrane protein expression and purification trials [46-49]. Longer alkyl chain non-ionic detergents proved most successful in extracting claudin-10, relative to SDS detergent (Figure 3). Further purification by FPLC size-exclusion chromatography under reducing conditions revealed a single, unaggregated peak (Figure 4). Our results indicate that claudin family proteins can be expressed and purified in an *E. coli* Mistic-tagged expression system.

3. Claudin expression in a Cell-free expression system

E. coli-derived, cell-free expression from *in vitro* reconstituted ribosomal components and T7 RNA polymerase is an emerging method for the rapid screening, labeling, and structural characterization of IMP via solution NMR spectroscopy [43, 50-54]. Cell-free expression is especially effective in overcoming the limitations of toxic expression in a host cell. In the absence of a hydrophobic membrane layer, IMPs

produced in a cell-free expression system form precipitates that can be solubilized and purified in mild detergents. We transferred claudin family proteins from Mistic-tagged expression vectors to cell-free expression vectors and adopted a previously described work flow (Figure 5).

Claudin family proteins expressed robustly in an *E. coli*-derived cell-free expression system, with over 16 targets expressing at quantities sufficient for characterization by solution NMR spectroscopy (Figure 6). In contrast to the *E. coli* Mistic expression system, most of the claudin family proteins were only solubilized in the detergent LMPG (Figure 6D). It is possible that expression in the native membrane state provides a more suitable foundation for extraction and solubilization by mild non-ionic detergents such as DDM and OG (Figure 3). It remains unclear how detergent solubilizes precipitated IMP in the cell-free expression system without the creation of partially misfolded aggregates.

Here we present evidence for the expression and purification of claudin family proteins in both an *E. coli* mistic-tagged expression system and an *E. coli*-derived cell-free expression system. Our results indicate that these systems are suitable for obtaining quantities of functional protein sufficient for structural studies. Determining the structure of a claudin pore is important in understanding epithelial barrier function. These methods should prove useful in guiding future structural studies of claudins and the paracellular pathway.

C. References

1. Loll, P.J., *Membrane protein structural biology: the high throughput challenge.* J Struct Biol, 2003. **142**(1): p. 144-53.

2. Laage, R. and D. Langosch, *Strategies for prokaryotic expression of eukaryotic membrane proteins.* Traffic, 2001. **2**(2): p. 99-104.

3. Granseth, E., et al., *Membrane protein structural biology--how far can the bugs take us?* Mol Membr Biol, 2007. **24**(5-6): p. 329-32.

4. Muller, G., *Towards 3D structures of G protein-coupled receptors: a multidisciplinary approach.* Curr Med Chem, 2000. **7**(9): p. 861-88.

5. Ahmad, A., et al., *Identification and design of antimicrobial peptides for therapeutic applications.* Curr Protein Pept Sci, 2012. **13**(3): p. 211-23.

6. Nijmeijer, S., et al., *Molecular pharmacology of histamine H4 receptors.* Front Biosci, 2012. **17**: p. 2089-106.

7. Hou, J., et al., *Study of claudin function by RNA interference.* J Biol Chem, 2006. **281**(47): p. 36117-23.

8. Angelow, S., R. Ahlstrom, and A.S. Yu, *Biology of claudins.* Am J Physiol Renal Physiol, 2008. **295**(4): p. F867-76.

9. Evans, M.J., et al., *Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry.* Nature, 2007. **446**(7137): p. 801-5.

10. Saadat, I., et al., *Helicobacter pylori CagA targets PAR1/MARK kinase to disrupt epithelial cell polarity.* Nature, 2007. **447**(7142): p. 330-3.

11. Van Itallie, C.M., A.S. Fanning, and J.M. Anderson, *Reversal of charge selectivity in cation or anion-selective epithelial lines by expression of different claudins.* Am J Physiol Renal Physiol, 2003. **285**(6): p. F1078-84.

12. Van Itallie, C.M. and J.M. Anderson, *Claudins and epithelial paracellular transport.* Annu Rev Physiol, 2006. **68**: p. 403-29.

13. Morita, K., et al., *Claudin multigene family encoding four-transmembrane domain protein components of tight junction strands.* Proc Natl Acad Sci U S A, 1999. **96**(2): p. 511-6.

14. Furuse, M., et al., *A single gene product, claudin-1 or -2, reconstitutes tight junction strands and recruits occludin in fibroblasts.* J Cell Biol, 1998. **143**(2): p. 391-401.

15. Furuse, M., et al., *Claudin-1 and -2: novel integral membrane proteins localizing at tight junctions with no sequence similarity to occludin.* J Cell Biol, 1998. **141**(7): p. 1539-50.

16. Kubota, K., et al., *Ca(2+)-independent cell-adhesion activity of claudins, a family of integral membrane proteins localized at tight junctions.* Curr Biol, 1999. **9**(18): p. 1035-8.

17. Colegio, O.R., et al., *Claudin extracellular domains determine paracellular charge selectivity and resistance but not tight junction fibril architecture.* Am J Physiol Cell Physiol, 2003. **284**(6): p. C1346-54.

18. Piontek, J., et al., *Formation of tight junction: determinants of homophilic interaction between classic claudins.* FASEB J, 2008. **22**(1): p. 146-58.

19. Van Itallie, C.M., et al., *Structure of the claudin-binding domain of Clostridium perfringens enterotoxin.* J Biol Chem, 2008. **283**(1): p. 268-74.

20. Umeda, K., et al., *ZO-1 and ZO-2 independently determine where claudins are polymerized in tight-junction strand formation.* Cell, 2006. **126**(4): p. 741-54.

21. McCarthy, K.M., et al., *Inducible expression of claudin-1-myc but not occludin-VSV-G results in aberrant tight junction strand formation in MDCK cells.* J Cell Sci, 2000. **113 Pt 19**: p. 3387-98.

22. Van Itallie, C.M., et al., *Palmitoylation of claudins is required for efficient tight-junction localization.* J Cell Sci, 2005. **118**(Pt 7): p. 1427-36.

23. Van Itallie, C.M., O.R. Colegio, and J.M. Anderson, *The cytoplasmic tails of claudins can influence tight junction barrier properties through effects on protein stability.* J Membr Biol, 2004. **199**(1): p. 29-38.

24. Roh, M.H., et al., *The carboxyl terminus of zona occludens-3 binds and recruits a mammalian homologue of discs lost to tight junctions.* J Biol Chem, 2002. **277**(30): p. 27501-9.

25. Gow, A., et al., *CNS myelin and sertoli cell tight junction strands are absent in Osp/claudin-11 null mice.* Cell, 1999. **99**(6): p. 649-59.

26. Kiuchi-Saishin, Y., et al., *Differential expression patterns of claudins, tight junction membrane proteins, in mouse nephron segments.* J Am Soc Nephrol, 2002. **13**(4): p. 875-86.

27. Watson, C.J., M. Rowland, and G. Warhurst, *Functional modeling of tight junctions in intestinal cell monolayers using polyethylene glycol oligomers.* Am J Physiol Cell Physiol, 2001. **281**(2): p. C388-97.

28. Nitta, T., et al., *Size-selective loosening of the blood-brain barrier in claudin-5-deficient mice.* J Cell Biol, 2003. **161**(3): p. 653-60.

29. Kovbasnjuk, O., et al., *Water does not flow across the tight junctions of MDCK cell epithelium.* Proc Natl Acad Sci U S A, 1998. **95**(11): p. 6526-30.

30. Yu, A.S., et al., *Claudin-8 expression in Madin-Darby canine kidney cells augments the paracellular barrier to cation permeation.* J Biol Chem, 2003. **278**(19): p. 17350-9.

31. Colegio, O.R., et al., *Claudins create charge-selective channels in the paracellular pathway between epithelial cells.* Am J Physiol Cell Physiol, 2002. **283**(1): p. C142-7.

32. Sasaki, H., et al., *Dynamic behavior of paired claudin strands within apposing plasma membranes.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3971-6.

33. Furuse, M., et al., *Claudin-based tight junctions are crucial for the mammalian epidermal barrier: a lesson from claudin-1-deficient mice.* J Cell Biol, 2002. **156**(6): p. 1099-111.

34. Ben-Yosef, T., et al., *Claudin 14 knockout mice, a model for autosomal recessive deafness DFNB29, are deaf due to cochlear hair cell degeneration.* Hum Mol Genet, 2003. **12**(16): p. 2049-61.

35. McClane, B.A., *The complex interactions between Clostridium perfringens enterotoxin and epithelial tight junctions.* Toxicon, 2001. **39**(11): p. 1781-91.

36. Michl, P., et al., *Claudin-4: a new target for pancreatic cancer treatment using Clostridium perfringens enterotoxin.* Gastroenterology, 2001. **121**(3): p. 678-84.

37. Wong, V. and B.M. Gumbiner, *A synthetic peptide corresponding to the extracellular domain of occludin perturbs the tight junction permeability barrier.* J Cell Biol, 1997. **136**(2): p. 399-409.

38. Wang, W., et al., *Human zonulin, a potential modulator of intestinal tight junctions.* J Cell Sci, 2000. **113 Pt 24**: p. 4435-40.

39. Clarke, H., et al., *Modification of tight junction function by protein kinase C isoforms.* Adv Drug Deliv Rev, 2000. **41**(3): p. 283-301.

40. Salama, N.N., N.D. Eddington, and A. Fasano, *Tight junction modulation and its relationship to drug delivery.* Adv Drug Deliv Rev, 2006. **58**(1): p. 15-28.

41. Mitic, L.L., V.M. Unger, and J.M. Anderson, *Expression, solubilization, and biochemical characterization of the tight junction transmembrane protein claudin-4.* Protein Sci, 2003. **12**(2): p. 218-27.

42. Newstead, S., S. Ferrandon, and S. Iwata, *Rationalizing alpha-helical membrane protein crystallization.* Protein Sci, 2008. **17**(3): p. 466-72.

43. Klammt, C., et al., *Facile backbone structure determination of human membrane proteins by NMR spectroscopy.* Nat Methods, 2012. **9**(8): p. 834-9.

44. Miroux, B. and J.E. Walker, *Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels.* J Mol Biol, 1996. **260**(3): p. 289-98.

45. Arechaga, I., et al., *Over-expression of Escherichia coli F1F(o)-ATPase subunit a is inhibited by instability of the uncB gene transcript.* FEBS Lett, 2003. **547**(1-3): p. 97-100.

46. Drew, D., et al., *GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in Saccharomyces cerevisiae.* Nat Protoc, 2008. **3**(5): p. 784-98.

47. Hu, N.J., et al., *Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT.* Nature, 2011. **478**(7369): p. 408-11.

48. Newstead, S., et al., *Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2.* EMBO J, 2011. **30**(2): p. 417-26.

49. Newstead, S., et al., *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. 13936-41.

50. Maslennikov, I., et al., *Membrane domain structures of three classes of histidine kinase receptors by cell-free expression and rapid NMR analysis.* Proc Natl Acad Sci U S A, 2010. **107**(24): p. 10902-7.

51. Klammt, C., et al., *Cell-free production of G protein-coupled receptors for functional and structural studies.* J Struct Biol, 2007. **158**(3): p. 482-93.

52. Klammt, C., et al., *Cell-free production of integral membrane proteins on a preparative scale.* Methods Mol Biol, 2007. **375**: p. 57-78.

53. Klammt, C., et al., *Polymer-based cell-free expression of ligand-binding family B G-protein coupled receptors without detergents.* Protein Sci, 2011. **20**(6): p. 1030-41.

54. Kalmbach, R., et al., *Functional cell-free synthesis of a seven helix membrane protein: in situ insertion of bacteriorhodopsin into liposomes.* J Mol Biol, 2007. **371**(3): p. 639-48.

FIGURES

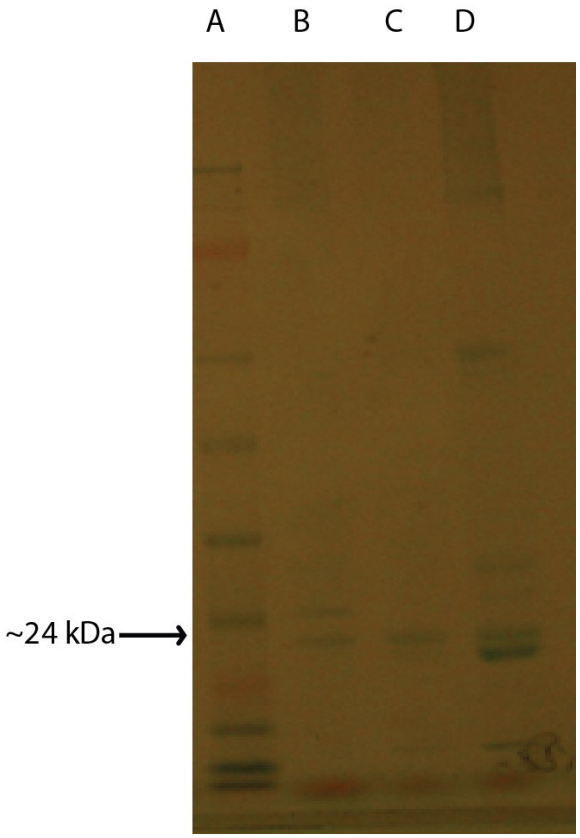


Figure 4.1 Expression and Ni-NTA purification of b) Claudin-1, c) Claudin-10, and d) Claudin-16 in Buffer B and DDM detergent. Protein standard loaded in lane a.

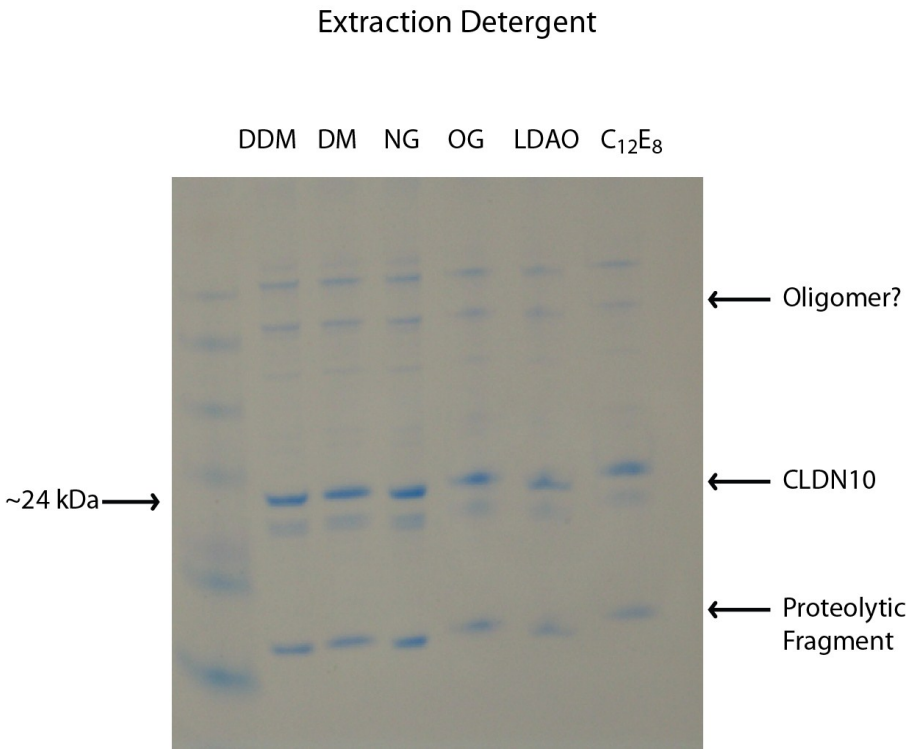


Figure 4.2 Detergent extraction test of Claudin-10 with DDM, DM, NG, OG, LDAO, and $C_{12}E_8$. All proteins were extracted at 10xCMC in Buffer B.

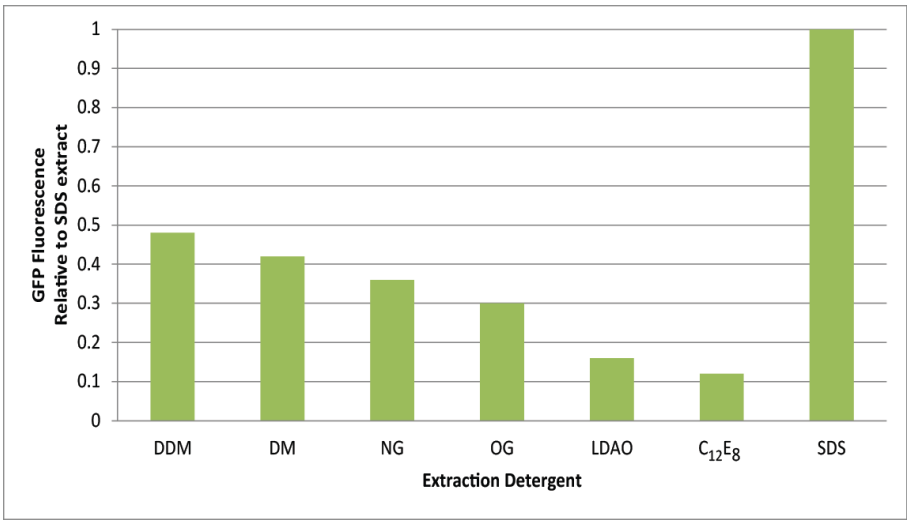


Figure 4.3 Claudin-10 detergent extraction efficiency measured by GFP fluorescence.

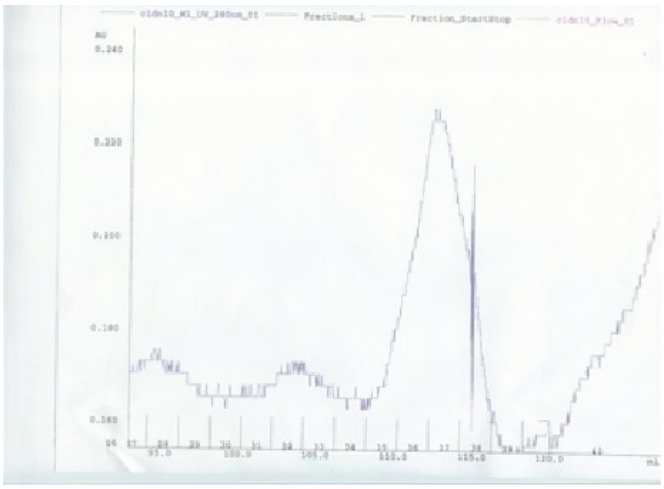


Figure 4.4 Size-exclusion chromatogram of Claudin-10 purified in Buffer B and 0.5 mM DDM (reducing conditions).

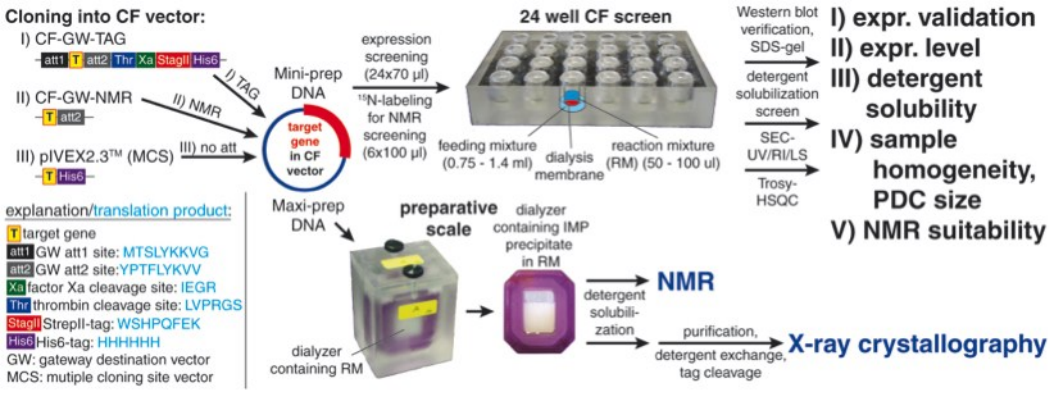


Figure 4.5 Workflow for cloning, cell-free expression, and screening of Claudin proteins for structural studies.

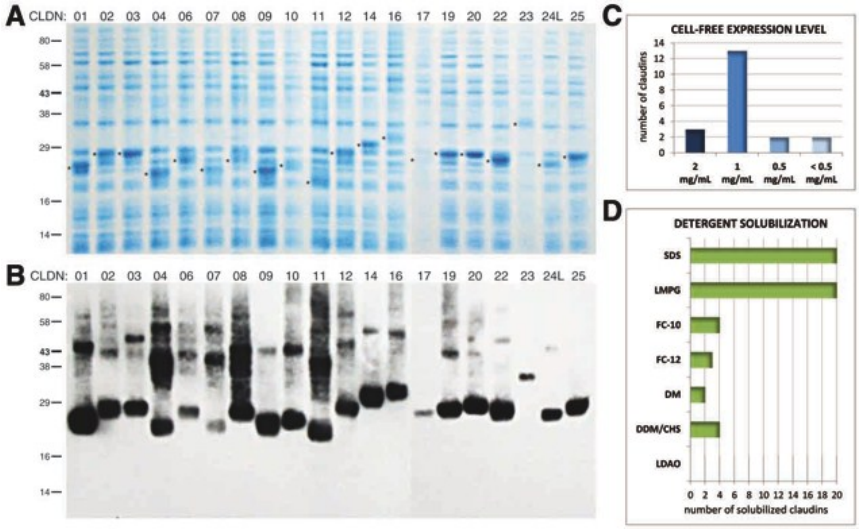


Figure 4.6 Cell-free expression and purification of Claudin family members. SDS-PAGE resolved Claudin cell-free expression extracts, visualized by a) Coomassie stain and b) Western Blot for α-His. Also pictured, c) estimated expression levels and d) solubility in different detergents.

ACKNOWLEDGEMENTS

Chapter IV, in part, was a collaboration with Christian Klammt involving cell-free expression

CHAPTER V

Conclusions and Perspective

In summary, my studies on the biochemical properties of MISTIC identify a potential mechanism for oligomerization that contributes to the membrane protein chaperoning properties of MISTIC. Functional analysis of MISTIC in *Bacillus subtilis* also identified novel roles for MstX and the downstream putative potassium channel, YugO, in the regulation of biofilm formation. I showed that the *mstX* gene is regulated by the master regulator for biofilm formation, SinR. Functional expression of both MstX and YugO were required to restore biofilm formation to *B. subtilis*. The activity of MstX and YugO depend upon the presence of the potassium-sensing kinase, KinC. I propose that MstX and YugO participate in a positive autoregulatory loop involving potassium leakage, KinC activation, and SinR derepression of genes involved in biofilm assembly. Furthermore, I demonstrate the ability of the MISTIC protein to assist in the expression of Claudin family proteins, a group of human integral membrane proteins that are of significant research interest. Thus, MISTIC proves to be a versatile platform for the expression of integral membrane proteins in multiple biological contexts.

A. Future Directions

1. Directed Evolution of Mistic for enhanced expression of integral membrane constructs

The ability of Mistic to support high levels of membrane protein expression present a novel opportunity to tackle previously intractable problems in biology. Due to its ability to be used for high-throughput protein overexpression in the familiar

expression species *E. coli*, Mistic provides an ideal platform for structural genomic studies of membrane proteins such as mouse GPCR's (MepNET), *Mycobacterium tuberculosis* membrane proteins (TBSGC), and *Pseudomonas aeruginosa* membrane proteins (Harvard Institute of Proteomics) [1]. However, it is unlikely that this heterologous *B. subtilis* protein is optimal for membrane protein expression in *E. coli* so it is likely that Mistic could be improved through genetic optimization and careful screening for improved membrane protein expression via a folding reporter fluorescent assay [2-4]. Such an approach would necessitate random mutagenesis and familial recombination directed toward the isolation of Mistic variants with improved expression capabilities [5-7]. Such Directed evolution approaches have proved instrumental in developing enhanced characteristics in many different proteins, including fluorescent proteins, tobacco etch virus protease, unnatural amino acid protein synthesis, thermostable DNA polymerases, GroEL, cytochrome P450, homing endonucleases, and a myriad of other industrial and biomedically relevant molecules [8-14]. Obtaining an improved isolate of Mistic would provide an immense resource to the biophysical community and biologists as a whole, enabling the structural study of a wide variety of previously unattainable yet immensely important proteins, including CCR5 (essential HIV-1 integrase co-receptor), presenilin-1 (alzheimer's disease), EGFR (cancer), CTFR (cystic fibrosis), beta-2 adrenergic receptor (lung and heart disease), and CRFR2-beta (anxiety and depression) [15-19].

Researchers can borrow Nature's algorithm for problem solving by selecting for beneficial mutations in MISTIC that improve the folding dynamics of membrane

proteins expressed *in vivo* in *E. coli*. One option is to assemble large, randomly mutagenized libraries of Mistic created by error-prone PCR and staggered extension process PCR with Mistic family members present in *Bacillus licheniformis*, *Bacillus atrophaeaus*, and *Bacillus mojavensis* [20-22]. Mutagenized libraries could be efficiently cloned into an arabinose-inducible pBAD33 vector by MEGAWHOP, minimizing the emergence of unmutagenized MISTIC among a mutagenized population [23, 24]. Researchers could use a positive selection marker with the expression of a boron efflux transporter (BOR1) fused to the N-terminus of GFP [25]. The successful folding of the chromophore and correlative fluorescence of green fluorescent protein serves as an excellent indicator of successful folding of upstream domains [26]. Integral membrane proteins that do not fold will form non-fluorescent inclusion bodies that can be selected against through diminished fluorescence [4]. Furthermore, functional expression of boron efflux proteins confers resistance to heavy metals, thus integrating a primary selection strategy that reduces the fluorescent screening workload [27]. By employing this rapid qualitative and quantitative screening with a direct selection method, Mistic variants that facilitate the integration of membrane proteins with a cytoplasmic tail can be efficiently selected for with the assistance of an Illumatool, microplate reader, GFP in-gel fluorescence, and resistance to boron.

Directed evolution and screening can proceed rapidly through the following work-flow. Researchers would isolate co-transformants containing mutagenized pBAD33-Mistic and pET-BOR1 on plates with arabinose, IPTG, and a wide range of

boron concentrations. BOR1 expresses only weakly with the aid of Mistic, does not express at all in the absence of Mistic, contains a prominent C-terminal tail (unpublished results), and represents a significant target for structural characterization. After a day of growth, colonies can be selected from boron plates containing at least ten-times more colonies than a no vector control. Colonies would then be transferred from these selection plates to replica plates will be grown at 30˚C for three hours, after which the colonies will be examined for fluorescence using an Illumatool (Lightools Research, Encinitis, CA). Screening 20,000 – 30,000 colonies per PCR reaction, as previously described [28], researchers could select for the top 400 "improved" expressers. Researchers could prepare replicates and examine their fluorescence quantitatively through the use of a microplate reader, using only colonies with significantly different and reproducible fluorescence for additional rounds of shuffling by staggered extension process and error-prone PCR. Following several repetitions of evolution through the application of this method, we will compare sequences of high expression mutants and identify critical sites of mutagenesis shared among them. These sites can be further manipulated through designed, PCR-based mutagenesis approaches in an attempt to isolate "ideal" mutants [29]. If these screening approaches prove insufficient for achieving dramatic improvements in membrane protein expression, researchers could apply quantitative FACS for screening millions of colonies with high efficiency [30].

Structural and biochemical characterization of membrane protein interactions requires considerable quantities of purified product for analysis. Evolutionary

optimization of Mistic for expression of eukaryotic membrane proteins could prove helpful in surmounting traditional constraints that impair the study of biological systems.

2. Potassium and Osmotic Pressure in the regulation of *B. subtilis* biofilm formation

Bacteria lack the ability to actively transport water. Therefore, osmotic processes determine the cellular water content and help buffer changes in external salinity or osmolality that would influence the hydration status of the cytoplasm and the magnitude of turgor [31]. Osmotically induced water fluxes that threaten collapse of turgor under hyperosmotic conditions or cell rupture under hypoosmotic conditions make it imperative for bacteria to actively modulate their intracellular solute pools [32]. Bacteria do this by amassing cytoplasmic and periplasmic ions and organic osmolytes external osmolarity is high and by expelling these compounds through mechanosensitive channels when external osmolarity is low [33, 34].

B. subtilis and *E. coli* respond to a sudden osmotic increase in their environment with a two-phase adaptation reaction [34-36]. In *B. subtilis*, the basal concentration of potassium is estimated at 350 mM, with osmolarities exceeding 650 mM creating severe osmotic challenges [35, 36]. Large amounts of potassium are imported through specific transport systems in order to compensate for the efflux of water associated with an increase in osmolarity. However, high intracellular concentrations of potassium have negative effects on protein function, DNA-protein interactions, and protein synthesis [37-39]. Thus, the initial response is only

satisfactory for dealing with temporary high osmolality. The second phase of osmoadaptation entails the synthesis of compatible solutes and the efflux of potassium [40, 41]. This second phase enables the long-term survival of bacteria exposed to extreme osmotic stress.

Potassium ions play an essential role in bacterial physiology through osmoregulation, maintenance of cellular pH, and also serve specific roles in the structure and function of many enzymes [42]. The *B. subtilis* genome encodes a variety of potential potassium transporters and channels. Two potassium transporters are encoded by KtrAB (higher-affinity) and KtrCD (lower-affinity). These two transporters play an important role in osmoregulation in response to high osmolarity [34], likely functioning in potassium uptake during the initial phase of adaptation. KtrAB and KtrCD have also been shown to facilitate colony spreading in *B. subtilis* by supporting sufficient levels of intracellular potassium ions for sliding motility [43].

More recently, researchers have identified potassium efflux and osmotic pressure as potential regulators of biofilm formation in *B. subtilis* [44, 45]. High concentrations of exopolysaccharide polymers have been shown to increase osmotic pressure and to coordinate gene expression in different cell types [44, 46]. You need to explain the connection with these results and potassium efflux. My studies on Mistic in *B. subtilis* imply that the putative potassium channel, YugO, might play a role complementary to KtrAB and KtrCD in regulating intracellular concentrations of potassium for coping with osmotic stress and for inducing the signal transduction pathways necessary to support biofilm formation.

During biofilm growth, *B. subtilis* is confronted with adverse conditions associated with nutrient depletion and desiccation. An intriguing possibility is that the potassium transporters and potassium channels might genetically coordinate the transition to multicellular growth through their osmoregulatory roles. Previous research groups have already demonstrated that surfactin and potassium transporters are essential for the *B. subtilis* multicellular behavior known as swarming motility [34, 43, 47, 48]. I also demonstrated that the *yugO* null mutant displayed increased colony spreading during growth in non-swarming conditions (unpublished observation), indicating that even at low basal level expression, YugO plays a significant role in mediating physiological processes in addition to biofilm formation. Additional studies are required to verify the role of YugO as a potassium efflux channel, to identify other genetic regulatory components involved in potassium sensing, and to identify the variety of roles potassium plays in multicellular growth in *B. subtilis*.

A. References

1. Terwilliger, T.C., et al., *The TB structural genomics consortium: a resource for Mycobacterium tuberculosis biology.* Tuberculosis (Edinb), 2003. **83**(4): p. 223-49.

2. Waldo, G.S., et al., *Rapid protein-folding assay using green fluorescent protein.* Nat Biotechnol, 1999. **17**(7): p. 691-5.

3. Cabantous, S. and G.S. Waldo, *In vivo and in vitro protein solubility assays using split GFP.* Nat Methods, 2006. **3**(10): p. 845-54.

4. Waldo, G.S., *Genetic screens and directed evolution for protein solubility.* Curr Opin Chem Biol, 2003. **7**(1): p. 33-8.

5. Arnold, F.H., *Fancy footwork in the sequence space shuffle.* Nat Biotechnol, 2006. **24**(3): p. 328-30.

6. Santoro, S.W., et al., *An efficient system for the evolution of aminoacyl-tRNA synthetase specificity.* Nat Biotechnol, 2002. **20**(10): p. 1044-8.

7. Collins, C.H., J.R. Leadbetter, and F.H. Arnold, *Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR.* Nat Biotechnol, 2006. **24**(6): p. 708-12.

8. Doyon, J.B., et al., *Directed evolution and substrate specificity profile of homing endonuclease I-SceI.* J Am Chem Soc, 2006. **128**(7): p. 2477-84.

9. van den Berg, S., et al., *Improved solubility of TEV protease by directed evolution.* J Biotechnol, 2006. **121**(3): p. 291-8.

10. Kermekchiev, M.B., A. Tzekov, and W.M. Barnes, *Cold-sensitive mutants of Taq DNA polymerase provide a hot start for PCR.* Nucleic Acids Res, 2003. **31**(21): p. 6139-47.

11. Wang, J.D., et al., *Directed evolution of substrate-optimized GroEL/S chaperonins.* Cell, 2002. **111**(7): p. 1027-39.

12. Wang, L., et al., *Expanding the genetic code of Escherichia coli.* Science, 2001. **292**(5516): p. 498-500.

13. Shaner, N.C., et al., *Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein.* Nat Biotechnol, 2004. **22**(12): p. 1567-72.

14. Otey, C.R., et al., *Functional evolution and structural conservation in chimeric cytochromes p450: calibrating a structure-guided approach.* Chem Biol, 2004. **11**(3): p. 309-18.

15. Balistreri, C.R., et al., *CCR5 receptor: biologic and genetic implications in age-related diseases.* Ann N Y Acad Sci, 2007. **1100**: p. 162-72.

16. Chai, C.K., *The genetics of Alzheimer's disease.* Am J Alzheimers Dis Other Demen, 2007. **22**(1): p. 37-41.

17. Zhang, X., et al., *An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor.* Cell, 2006. **125**(6): p. 1137-49.

18. Carlin, K.M., W.W. Vale, and T.L. Bale, *Vital functions of corticotropin-releasing factor (CRF) pathways in maintenance and regulation of energy homeostasis.* Proc Natl Acad Sci U S A, 2006. **103**(9): p. 3462-7.

19. Whalen, E.J., et al., *Regulation of beta-adrenergic receptor signaling by S-nitrosylation of G-protein-coupled receptor kinase 2.* Cell, 2007. **129**(3): p. 511-22.

20. Roosild, T.P., et al., *NMR structure of Mistic, a membrane-integrating protein for membrane protein expression.* Science, 2005. **307**(5713): p. 1317-21.

21. Roosild, T.P., et al., *Characterization of the family of Mistic homologues.* BMC Struct Biol, 2006. **6**: p. 10.

22. Zhao, H., et al., *Molecular evolution by staggered extension process (StEP) in vitro recombination.* Nat Biotechnol, 1998. **16**(3): p. 258-61.

23. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.* J Bacteriol, 1995. **177**(14): p. 4121-30.

24. Miyazaki, K. and M. Takenouchi, *Creating random mutagenesis libraries using megaprimer PCR of whole plasmid.* Biotechniques, 2002. **33**(5): p. 1033-4, 1036-8.

25. Dumon-Seignovert, L., G. Cariot, and L. Vuillard, *The toxicity of recombinant proteins in Escherichia coli: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3).* Protein Expr Purif, 2004. **37**(1): p. 203-6.

26. Drew, D., et al., *Optimization of membrane protein overexpression and purification using GFP fusions.* Nat Methods, 2006. **3**(4): p. 303-13.

27. Sutton, T., et al., *Boron-toxicity tolerance in barley arising from efflux transporter amplification.* Science, 2007. **318**(5855): p. 1446-9.

28. Drummond, D.A., et al., *Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins.* J Mol Biol, 2005. **350**(4): p. 806-16.

29. Sawano, A. and A. Miyawaki, *Directed evolution of green fluorescent protein by a new versatile PCR strategy for site-directed and semi-random mutagenesis.* Nucleic Acids Res, 2000. **28**(16): p. E78.

30. Nguyen, A.W. and P.S. Daugherty, *Evolutionary optimization of fluorescent proteins for intracellular FRET.* Nat Biotechnol, 2005. **23**(3): p. 355-60.

31. Wood, J.M., *Osmosensing by bacteria: signals and membrane-based sensors.* Microbiol Mol Biol Rev, 1999. **63**(1): p. 230-62.

32. Poolman, B. and E. Glaasker, *Regulation of compatible solute accumulation in bacteria.* Mol Microbiol, 1998. **29**(2): p. 397-407.

33. Booth, I.R. and C.F. Higgins, *Enteric bacteria and osmotic stress: intracellular potassium glutamate as a secondary signal of osmotic stress?* FEMS Microbiol Rev, 1990. **6**(2-3): p. 239-46.

34. Holtmann, G., et al., *KtrAB and KtrCD: two K+ uptake systems in Bacillus subtilis and their role in adaptation to hypertonicity.* J Bacteriol, 2003. **185**(4): p. 1289-98.

35. Whatmore, A.M., J.A. Chudek, and R.H. Reed, *The effects of osmotic upshock on the intracellular solute pools of Bacillus subtilis.* J Gen Microbiol, 1990. **136**(12): p. 2527-35.

36. Whatmore, A.M. and R.H. Reed, *Determination of turgor pressure in Bacillus subtilis: a possible role for K+ in turgor regulation.* J Gen Microbiol, 1990. **136**(12): p. 2521-6.

37. Record, M.T., Jr., W. Zhang, and C.F. Anderson, *Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts.* Adv Protein Chem, 1998. **51**: p. 281-353.

38. Record, M.T., Jr., et al., *Biophysical compensation mechanisms buffering E. coli protein-nucleic acid interactions against changing environments.* Trends Biochem Sci, 1998. **23**(5): p. 190-4.

39. Record, M.T., Jr., et al., *Responses of E. coli to osmotic stress: large changes in amounts of cytoplasmic solutes and water.* Trends Biochem Sci, 1998. **23**(4): p. 143-8.

40. Brown, A.D., *Microbial water stress.* Bacteriol Rev, 1976. **40**(4): p. 803-46.

41. Yancey, P.H., et al., *Living with water stress: evolution of osmolyte systems.* Science, 1982. **217**(4566): p. 1214-22.

42. Epstein, W., *The roles and regulation of potassium in bacteria.* Prog Nucleic Acid Res Mol Biol, 2003. **75**: p. 293-320.

43. Kinsinger, R.F., et al., *Genetic requirements for potassium ion-dependent colony spreading in Bacillus subtilis.* J Bacteriol, 2005. **187**(24): p. 8462-9.

44. Rubinstein, S.M., et al., *Osmotic pressure can regulate matrix gene expression in Bacillus subtilis.* Mol Microbiol, 2012. **86**(2): p. 426-36.

45. Lopez, D., et al., *Structurally diverse natural products that cause potassium leakage trigger multicellularity in Bacillus subtilis.* Proc Natl Acad Sci U S A, 2009. **106**(1): p. 280-5.

46. Russo, F.D. and T.J. Silhavy, *EnvZ controls the concentration of phosphorylated OmpR to mediate osmoregulation of the porin genes.* J Mol Biol, 1991. **222**(3): p. 567-80.

47. Angelini, T.E., et al., *Bacillus subtilis spreads by surfing on waves of surfactant.* Proc Natl Acad Sci U S A, 2009. **106**(43): p. 18109-13.

48. Seminara, A., et al., *Osmotic spreading of Bacillus subtilis biofilms driven by an extracellular matrix.* Proc Natl Acad Sci U S A, 2012. **109**(4): p. 1116-21.

49. Nevo-Dinur, K., et al., *Translation-independent localization of mRNA in E. coli.* Science, 2011. **331**(6020): p. 1081-4.

APPENDIX

Intracellular localization and solubility of a Mistic-GFP fusion construct

A. RESULTS

Mistic is a highly hydrophilic and negatively charged protein with no identifiable secretory signals or transmembrane helices. In spite of these characteristics, Mistic has been reported to autonomously associate with the membrane, thereby facilitating the targeting of fused integral membrane constructs to the membrane. In order to ascertain the mechanism of action and to verify the localization of Mistic in living cells, I carried out fluorescent microscopy studies in *Escherichia coli* and *Bacillus subtilis* using n-terminal and c-terminal GFP fusions. Mistic-GFP fusions were expressed under the control of an arabinose or xylose-inducible promoter. Observation of membrane localized GFP fluorescence would indicate that a Mistic fusion is sufficient for targeting GFP constructs to the membrane, whereas disperse cytoplasmic fluorescence would indicate that Mistic requires other factors in order to localize to the membrane. Briefly, in *Escherichia coli*, Mistic was cloned into a pBAD33 vector containing a c-terminal GFP with a c-terminal 6xhis tag for purification. The pBAD33 vector is under control of an arabinose-inducible promoter. E. coli DH10B containing pBAD33-Mistic-GFP-6xhis construct were induced at 0.3 OD600 with 0.2% Arabinose in LB media and observed by fluorescent microscopy after an hour of growth. In *Bacillus subtilis*, Mistic was cloned into pSG1154 (c-terminal GFP tag) or pSG1729 (n-terminal GFP tag) and ectopically integrated at the *amyE* locus. Expression of the resulting fusion was induced at 0.3 OD600 with the addition of 0.5% xylose in LB media and observed after an hour of growth.

Results from *Escherichia coli* indicate that c-terminal Mistic-GFP-6xhis constructs were predominantly soluble and present throughout the cytoplasm (Figure 1B). The results of the Mistic-GFP-6xhis construct were similar to GFP-6xhis expressed alone, with the exception of slightly lower signal intensity (Figure 1A). Interestingly, I was unable to observe any fluorescence in both n-terminal and c-terminal Mistic-GFP fusion constructs in *Bacillus subtilis*, implying that the construct failed to express or fold properly. It is thus unclear if Mistic localizes *in vivo* to the cytoplasm or the membrane fraction of *B. subtilis*.

The c-terminal GFP fusion moderately influenced the *in vitro* solubility of Mistic in *E. coli* (Figure 2). *E. coli* DH10B cells containing pBAD33-Mistic-GFP-6xhis construct were induced at 0.3 OD600 with 0.2% Arabinose in 200 ml LB media and grown for three hours prior to collection. Cells were pelleted by centrifugation at 5,000 x g for ten minutes and lysed by a fluidizer. The lysate was cleared by centrifugation at 80,000 x g for one hour. The resulting supernatant was loaded onto a Ni-NTA column and eluted with 250 mM Imidazole after washing with five column volumes of buffer. The washed and eluted protein was resolved on SDS-PAGE. More than half of the Mistic-GFP-6xhis fusion partitioned to the soluble fraction after high-speed centrifugation (Figure 1G, expected size of fusion Mistic-GFP-6xhis fusion ~35 kDa). The pelleted membrane fraction from the previous step was resuspended and solubilized in buffer containing 20 mM Lauryldimethylamine Oxide (LDAO) for one hour. The solution was cleared by centrifugation at 80,000 x g for one hour. The resulting supernatant was loaded onto a Ni-NTA column and eluted

with 250 mM Imidazole after washing with five column volumes of buffer containing 4 mM LDAO. A smaller amount of Mistic-GFP-6xhis fusion was observed in the membrane fraction (Figure 2D) compared to the purified, soluble fraction (Figure 2G).

These results suggest that Mistic is not sufficient to target fused cargo proteins to the membrane in *E. coli*. The addition of a 250 amino acid GFP-6xhis tag to the c-terminus of Mistic significantly increased the solubility of the resulting fusion protein (Figure 2). Moreover, fluorescent microscopy revealed that a Mistic-GFP-6xhis fusion predominantly occupies the cytoplasm as opposed to the membrane layer (Figure 1B). It is more probable that Mistic facilitates the expression of associated integral membrane proteins through enhanced folding or improved pre-translational mRNA localization to the membrane than through direct membrane association or targeting [49]. Additional experiments are required to further validate and distinguish between these hypotheses.

FIGURES

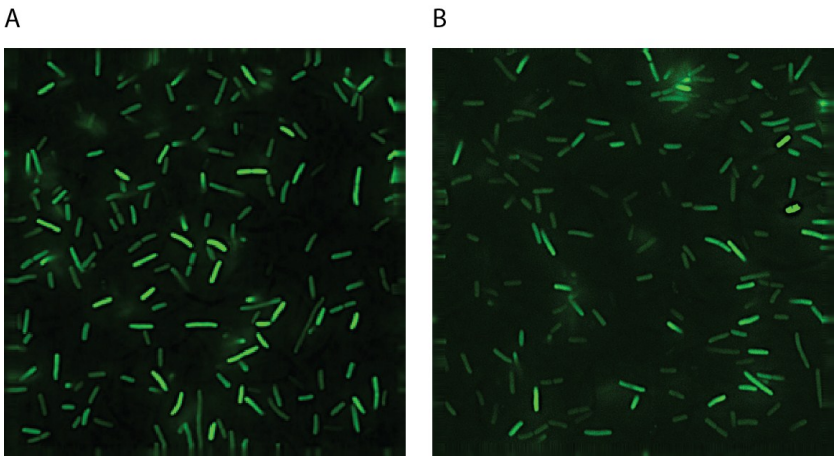


Figure A.1 Mistic compartmentalization assay using a c-terminal GFP tag. Fluorescent microscopy of a DH10B strain expressing Mistic-GFP-6xhis from a pBAD33 expression vector. The Mistic-GFP-6xhis protein was induced with the addition of 0.2% Arabinose at 0.3 OD600 in LB media. Cells were imaged after one hour of growth.

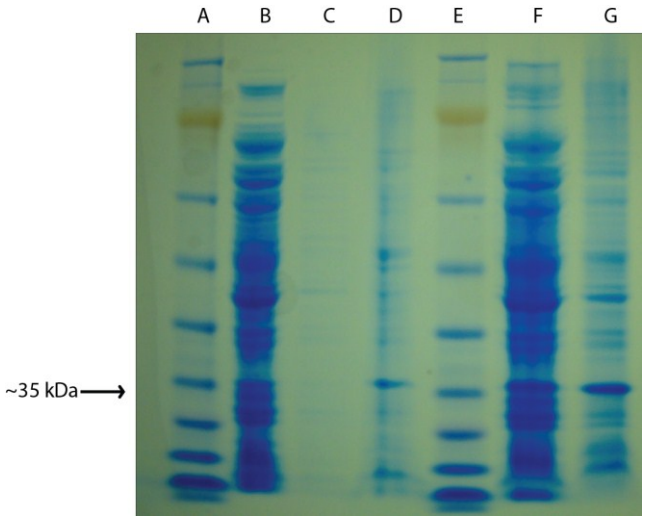


Figure A.2 SDS-PAGE and Coomassie stain of unpurified and purified extracts of a strain expressing Mistic-GFP-6xhis. DH10B strain expressing Mistic-GFP-6xhis from a pBAD33 expression vector. Membrane and soluble fractions were isolated and purified as described. Lane (a) protein ladder standard, (b) membrane fraction flow-through, (c) membrane column wash flow-through, (d) imidazole eluted membrane fraction, (e) protein ladder standard, (f) soluble fraction flow-through, (g) imidazole eluted soluble fraction.