STRUCTURAL AND BIOPHYSICAL CHARACTERIZATION OF VARIANTS OF THE MECHANOSENSITIVE CHANNEL OF LARGE CONDUCTANCE (MSCL)

Thesis by

Chinenye Abiodun Idigo

In Partial Fulfillment of the Requirements for the

degree of

Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2015

(Defended October 21st 2014)

ACKNOWLEDGEMENTS

I would like to acknowledge those who have been my support system on the journey to the completion of my graduate degree. When I arrived on Caltech's campus in the fall of 2008, I did not know what was ahead of me, but I knew that I wanted to join a prolific lab headed by a knowledgeable and involved professor. When I found the lab of **Professor Douglas Rees**, I knew that I had found my lab home. I would like to thank Doug for creating an environment where I had the opportunity to explore science with few limits. Doug's unique blend of optimistic realism gave me the freedom, not only to pursue membrane protein crystallography, but also to know when it was time to resort to plan B. Doug, your honest and wise advice helped me navigate the convoluted yet highly rewarding path of scientific discovery, and for that I am truly grateful. I would also like to thank the professors on my committee, **Henry Lester, Bil Clemons, David Chan,** and **Rob Phillips**, the members of the Dean's office, **Joe Shepherd, Felicia Hunt, Natalie Gilmore** and **Barbara Green**, the staff of the Caltech Center for Diversity, **Eva Graham, Portia Harris, Linda Webb** and **Taso Dimitriadis**, and the administrators, **Phoebe Ray, Alison Ross** and **Liz Ayala**. Thank you for the guidance and support over the years.

I would like to thank my colleagues in the Rees Lab for creating a lab environment that is supportive, collaborative and entertaining. I would like to thank **Zhenfeng Liu** for introducing me to the world of MscL and crystallography, **Chris Gandhi** and **Troy Walton** for their many guiding insights, and **Jens Kaiser** and **Julie Hoy** for their help with crystallography projects. I am grateful to my bay-mate and friend **Jeffrey Lai**, for the positive energy, life advice and for sharing your wealth of knowledge on

experimental procedures. Jeff, you served as a daily support system for me and I cannot thank you enough for that.

To the many great friends I made at Caltech; **Nadia Herrera, Alysia Ahmed, Qi Wen Li, Amy McCarthy, Matt Gethers, Peter Agbo, Schetema Nealy, Tara Gomez, Kaitlyn Lucey, Keeshia Wang,** and **Gabriele Meloni**, thank you for all the memories. To my best friends in the world, **Ijeoma Ohiaeri** and **Michelle Burke**, thank you for the endless encouragement and support.

Finally, I would like to thank my amazing family; words cannot express how much your love and support mean to me. I thank my parents, **Pius** and **Felicia Idigo**, for always believing in me and encouraging me to push past my perceived limitations. Daddy you taught me the importance of a good education and to always aim high. Mommy you taught me patience, quiet strength, and perseverance. To my siblings **Nnaemeka, Oliaku,** and **Obiora Idigo**: Nnaemeka, thank you for always helping me laugh through the tough times and listening to my frustrations. Oliaku, you are my best friend, my confidant and my biggest fan. You mean the world to me. Obiora, thank you for all the encouragement and for sending me care packages. I always enjoy our engaging in-depth philosophical conversations that challenge me to rethink the way I view life. My family, you are my heart and I dedicate this thesis to you all.

ABSTRACT

The ability to sense mechanical force is vital to all organisms to interact with and respond

to stimuli in their environment. Mechanosensation is critical to many physiological

functions such as the senses of hearing and touch in animals, gravitropism in plants and

osmoregulation in bacteria. Of these processes, the best understood at the molecular level

involve bacterial mechanosensitive channels. Under hypo-osmotic stress, bacteria are

able to alleviate turgor pressure through mechanosensitive channels that gate directly in

response to tension in the membrane lipid bilayer. A key participant in this response is

the mechanosensitive channel of large conductance (MscL), a non-selective channel with

a high conductance of ~3 nS that gates at tensions close to the membrane lytic tension.

It has been appreciated since the original discovery by C. Kung that the small subunit size

(~130 to 160 residues) and the high conductance necessitate that MscL forms a homo-

oligomeric channel. Over the past 20 years of study, the proposed oligomeric state of

MscL has ranged from monomer to hexamer. Oligomeric state has been shown to vary

between MscL homologues and is influenced by lipid/detergent environment. In this

thesis, we report the creation of a chimera library to systematically survey the correlation

between MscL sequence and oligomeric state to identify the sequence determinants of

oligomeric state. Our results demonstrate that although there is no combination of

sequences uniquely associated with a given oligomeric state (or mixture of oligomeric

states), there are significant correlations. In the quest to characterize the oligomeric state

of MscL, an exciting discovery was made about the dynamic nature of the MscL

complex. We found that in detergent solution, under mild heating conditions (37 °C – 60

°C), subunits of MscL can exchange between complexes, and the dynamics of this process are sensitive to the protein sequence.

Extensive efforts were made to produce high diffraction quality crystals of MscL for the determination of a high resolution X-ray crystal structure of a full length channel. The surface entropy reduction strategy was applied to the design of *S. aureus* MscL variants and while the strategy appears to have improved the crystallizability of *S. aureus* MscL, unfortunately the diffraction qualities of these crystals were not significantly improved. MscL chimeras were also screened for crystallization in various solubilization detergents, but also failed to yield high quality crystals.

MscL is a fascinating protein and continues to serve as a model system for the study of the structural and functional properties of mechanosensitive channels. Further characterization of the MscL chimera library will offer more insight into the characteristics of the channel. Of particular interest are the functional characterization of the chimeras and the exploration of the physiological relevance of intercomplex subunit exchange.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER I

INTRODUCTION

Portions of this chapter were adapted from "MscL: channeling membrane tension" T.A. Walton, C.A. Idigo, N. Herrera, D.C. Rees, Pflugers Arch - Eur J Physiol. DOI 10.1007/s00424-014-1535-x. Published online: 27 May 2014 [1]

1.1 Mechanosensitive Channel of Large Conductance (MscL)

Mechanosensors are a ubiquitous class of biomolecules that play key roles in transducing mechanical force into signals for regulation of proper development and survival of living organisms [2]. Mechanosensing channels are implicated in many functions such as the senses of hearing and touch in animals, gravitropism in plants, and osmoregulation in bacteria. As examples, the Piezo channel family is mechanically activated, cation selective, and exemplifies mechanotransduction in vertebrates and invertebrates [3, 4]. The MscS-like (MSL) family of proteins in the plant *Arabidopsis thaliana* helps maintain the correct shape and size of plastids, which are organelles involved in photosynthesis and gravity sensing [5]. Under hypo-osmotic stress, bacteria are able to alleviate turgor pressure through mechanosensitive channels that gate directly in response to tension in the membrane lipid bilayer [6, 7].

Stretch-activated bacterial ion channels were discovered by C. Kung's group using patch clamp experiments on *Escherichia coli* giant spheroplasts [8]. As bacteria contain a variety of different channels, the crucial breakthrough was the identification by Kung and

coworkers of channels that conducted only when tension was applied to the membrane by the application of suction to the patch pipette [8]. The first channel to be discovered, later identified as the mechanosensitive channel of small conductance (MscS) [9, 10], demonstrated a single channel conductance of ~1 nS and showed both pressure and voltage dependence as well as selectivity for anions. When subjected to even stronger suction, patches excised from *E. coli* giant spheroplasts showed activation of another ion channel with ~3 nS conductance [7]. This channel was identified as the mechanosensitive channel of large conductance (MscL). MscL shows no ion selectivity and is localized in the inner membrane [10-12]. Experiments on purified MscL reconstituted into azolectin liposomes indicate that MscL is gated solely by tension in the membrane lipid bilayer. The tension required to gate MscL is near the lytic limit of the membrane, ~10-12 mN m^{-1} [13]. In osmotic down shock assays, *mscS/mscL* double knockout cells have poor survivability [9]. These findings make it clear that a critical function of mechanosensitive channels is to serve as emergency release valves in bacteria. The discovery of such a simple yet robust system provides an ideal model to study the mechanism of mechanosensation and the structural dynamics of channel gating.

A key development in establishing the molecular basis of mechanosensation was the identification by Kung's group of the gene for *E. coli* MscL, which encoded a 136 amino acid polypeptide [10]. Given the small subunit size, the large conductance of MscL, and results of preliminary experiments on a size exclusion column where MscL elutes at a volume consistent with a molecular mass of roughly 70 kDa, it was postulated that the active form of MscL is a homo-oligomer [10]. MscL has since been identified in numerous

bacterial species [14, 15]. The best studied MscL orthologues are from *E. coli*, *Staphylococcus aureus*, and *Mycobacterium tuberculosis* (EcMscL, SaMscL, MtMscL, respectively); these channels contain 136, 120, and 151 residues, respectively, and the pairwise percent sequence identity between these proteins are EcMscL - MtMscL (37%), SaMscL - MtMscL (40%), and EcMscL - SaMscL (51%) [Fig. 1-1G].

The discovery that a relatively small membrane protein functioned as a physiologically gated channel directly motivated efforts in the Rees group to solve the structure of MscL in different conformational states [16]. To date there are three crystal structures of full-length or truncated versions of MscL; MtMscL [17, 18], a C-terminal truncation of SaMscL (SaMscL(CΔ26)) [19], and the EcMscL C-terminal domain (EcMscL-CTD) [20] [Fig. 1-1]. The structure of MtMscL determined at 3.5 Å resolution established the basic subunit architecture and revealed a pentameric channel with the subunits surrounding a central pore. In agreement with hydropathy predictions and other biochemical analyses [Fig 1-1 A & D] [10, 11], the MscL subunit was found to contain two transmembrane (TM) helices. The first 12 amino acids of the N-terminus are located on the cytoplasmic side of the membrane and form an amphipathic α-helix followed by the first transmembrane helix, (TM1), that crosses the membrane and lines the permeation pathway of the channel. Residues 48-68 are located on the periplasmic side of the membrane and form an extended loop with two antiparallel β-sheets. A second transmembrane α-helix (TM2) traverses the membrane again, from the periplasm towards the cytoplasm, and flanks the exterior of the channel. The C-terminal domain consists of a short loop (residues 102-105), followed by a left-handed coiled coil formed by α-helical residues 106-125. TM1 forms the core of a

complex network of interactions between subunits in the transmembrane region. The TM1 of one subunit interacts with the TM1s from two adjacent subunits as well as having an extensive antiparallel interface with TM2 from an adjacent subunit. In addition, the N-terminal α-helix of one subunit inserts between the TMs of an adjacent subunit at the cytoplasmic membrane interface. The narrowest part of the permeation pathway of the channel is largely hydrophobic, particularly at the constriction point, formed by the side chains of Ile 17 and Val 21, where the pore diameter is estimated to be between 2-3 Å. Theoretical models predict that a hydrophobic pore of ~9 Å diameter is required to conduct water and ~13 Å to conduct hydrated ions [21]. Therefore, the conformation of the MtMscL structure is designated as a closed or non-conducting state.

The SaMscL(CΔ26) crystal structure was reported at 3.8 Å resolution [Fig. 1-1 B & E] [19]. In contrast to the MtMscL structure, SaMscL(CΔ26) adopted a tetrameric oligomeric state with a shorter, widened conformation. In SaMscL(CΔ26), TM1 also lines the permeation pathway, but relative to MtMscL, it adopts a larger tilt angle with respect to the pore axis. As in the MtMscL structure, TM1 of one subunit interacts with TM1s from two adjacent subunits, but with an increased crossing angle. TM1 also exhibits a similar interaction with TM2 from an adjacent subunit. The pore diameter at the construction point, Val 21, is expanded to ~6 Å. This pore size is still not large enough to allow passage of water or hydrated ions; therefore, this structure is designated as a non-conducting expanded intermediate state.

The structure of EcMscL-CTD spanning residues 108-136 was determined at 1.45 Å resolution [Fig. 1-1C & F] [20]. It revealed a pentameric α-helical coiled coil (residues 116-136) and an irregular extended region (residues 108-115). The EcMscL-CTD has almost three heptad repeats and shows the characteristic knob into hole packing of hydrophobic residues at periodic intervals expected for a coiled coil [20]. This arrangement is similar to the MtMscL-CTD and to other coiled coils, such as the cartilage oligomeric matrix protein (COMP) [22].
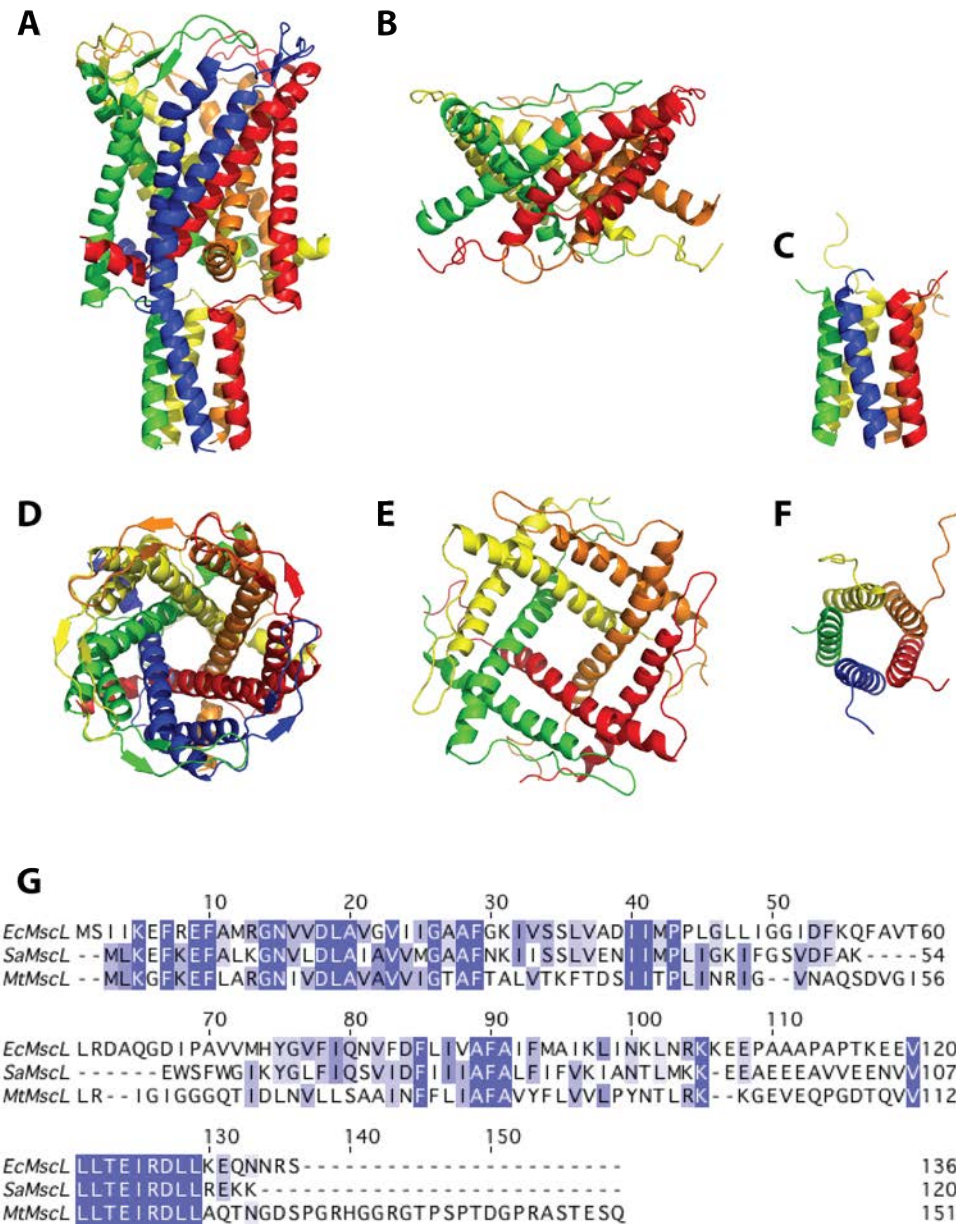
Figure 1-1. Crystal Structures of MtMscL, SaMscL(CΔ26), and EcMscL-CTD. (A, D) Side and top views of the crystal structure of MtMscL (PDBID: 2OAR). (B, E) Side and top views of the crystal structure of SaMscL(CΔ26) (PDBID: 3HZQ). (C, F) Side and top views of the crystal structure of EcMscL-CTD (PDBID: 4LKU). (G) Alignment of the amino acid sequence of EcMscL, SaMscL, and MtMscL. The pairwise percent sequence identity between these proteins are EcMscL - MtMscL (37%), SaMscL - MtMscL (40%), and EcMscL - SaMscL (51%).

1.2 Oligomeric State of MscL

To achieve nS conductance from a protein of only 136 residues, EcMscL was postulated to form a homo-oligomer when it was first identified [10]. Determination of the correct oligomeric state of MscL is critical to inform estimations for pore diameter and gating mechanism. The oligomeric state of MscL has been studied using various biochemical and biophysical techniques, which include covalent crosslinking [11, 12, 18, 23-28], tandem subunit fusions [11, 26, 29], 2-D electron crystallography [30], atomic force microscopy (AFM) [31], X-ray crystallography [17-20], analytical ultra centrifugation (AUC) [24, 26], size exclusion chromatography coupled to multiangle laser light scattering (SEC-MALS) [24, 32], and oligomer characterization by addition of mass (OCAM) [32]. Researchers have studied the oligomeric state of various wild type and mutant MscL constructs *in vivo* and *in vitro*. After nearly 20 years of study, the reported oligomeric state of MscL has ranged from monomer [12] to hexamer [11]. In this section, the results of these studies, including the strengths and weaknesses of different techniques used, will be discussed, along with the influence of lipid/detergent environment and sequence space on the variability of MscL oligomeric state.

Perhaps the most labor intensive method to determine the oligomeric state of MscL has been structural biology. X-ray crystallography, 2-D electron microscopy, and atomic force microscopy (AFM) have all been used to image MscL with various results [17-20, 30, 31]. The highest resolution structures can be resolved by X-ray crystallography, and to date there are three crystal structures of MscL homologs or subdomains [17-20]. X-ray crystallography is a low-throughput technique because the limiting factors in X-ray

structure determination of MscL are crystal formation and crystal quality. MscL is readily overexpressed in *E. coli* and purification gives relatively high protein yields. Nevertheless, as for most membrane proteins, MscL is recalcitrant to crystallization; even when crystals appear, they are typically of poor diffraction quality. Chang *et al.* screened 9 MscL homologs, 20 detergents, and over 24,000 crystal conditions before they were able to resolve the structure of MtMscL. Liu *et al.* determined a structure of truncated SaMscL after several years of work on full length EcMscL and SaMscL. Furthermore, Walton and Rees crystallized the soluble C-terminal domain of EcMscL to provide the first structural information for the *E. coli* homologue (while crystals of full length EcMscL have been obtained for nearly 20 years, their resolution has never exceeded ~8 Å). These structures revealed MtMscL as a pentamer, SaMscL(CΔ26) as a tetramer, and EcMscL-CTD as a pentamer. The origins of this variability in oligomeric state have not been established, and may reflect the intrinsic preferences of the protein sequence, as well as external factors such as the use of detergents, or the selective crystallization of a particular oligomeric state. It should be noted, however, that, the oligomeric state of detergent solubilized MtMscL and SaMscL(CΔ26) observed crystallographically have been confirmed with other methods [32].

2-D electron crystallography was used to resolve an EcMscL structure to a resolution of 15 Å. This study suggested that EcMscL is a hexamer [30]. Detergent purified EcMscL was reconstituted into *E. coli* lipid liposomes at high protein to lipid ratios to form vesicular 2-D crystals. Negatively stained single layer 2-D crystals were imaged, and the 14 best images were processed. Two types of analysis, a filtered image without symmetry applied, and

filtered images with a six-fold symmetry plane group applied, produced projection maps with apparent six-fold symmetry. The EcMscL molecules were visualized as hexagonal particles with a central depression. This was interpreted to be six EcMscL subunits in a ring lining a pore. At the resolution of the study, however, it is difficult to see such features with high certainty. AFM studies were performed on *Salmonella typhimurium* MscL (StMscL) immobilized on organic monolayers that allowed for experimental control of protein tension and adhesion [31]. These studies show that under certain conditions, StMscL is present in the form of nanoscale entities with a diameter of 4-5 nm, interpreted as a pentameric structure. The dimensions of the AFM probe, and the influence of the probe on the lateral dimensions recorded for subject proteins, likely limit the resolution of these studies.

Historically, covalent cross-linking has been the most popular technique for determining the oligomeric state of MscL, as it is relatively simple, fast, and inexpensive. Cross-linking, however, has generated the most varied and ambiguous results, due to differences in cross-linker specificity, length, concentration, and reaction environment. Several studies have used non-specific cross-linking agents, such as disuccinimidyl suberate (DSS), which reacts with primary amine groups. Studies using DSS have given different results, even with comparable protocols. An early cross-linking study by Blount *et al.* used DSS to cross-link purified EcMscL solubilized in β-octylglucoside [11]. The products of the reactions were separated by SDS-PAGE and visualized via silver staining. Six distinct bands were observed, indicating EcMscL is hexameric [11]. Similar results were observed

when purified membrane fractions were cross-linked, and products were visualized via western blot.

Subsequently, Sukharev *et al.* performed cross-linking experiments with OG purified EcMscL, using variable concentrations of DSS [26]. At lower concentrations of DSS, there were over eight distinct bands on an SDS-PAGE gel, visualized with Coomassie Blue dye. At higher concentrations of DSS, both the higher and lower molecular weight bands disappeared, leaving three prominent bands with the most intense migrating at a molecular weight consistent with pentameric MscL. Based on these results and those of similar experiments using other crosslinking agents, EcMscL oligomeric state was interpreted as pentameric. Since DSS has a relatively long linker length (11 Å), the higher molecular weight bands were attributed to intercomplex crosslinking. It should be noted, however, that these studies used slightly different crosslinking protocols.

To address some of the issues of non-specific crosslinking agents and detergent effects, several studies have used a disulfide trapping strategy to crosslink MscL in membranes [23-25, 33]. Both wild-type SaMscL and EcMscL lack cysteine residues, making site-specific cysteine mutants relatively simple to generate. The double cysteine mutant, SaMscL L10C/M91C, has intracellular cysteine residues in the N-terminal and TM2 regions of SaMscL. Experiments conducted in whole cells require an oxidizing agent, such as copper phenanthroline to catalyze disulfide bond formation in the cytosol following osmotic shock to activate the channel. This process allows the copper phenanthroline complex to enter the cell directly through the MscL pore, since it will not otherwise diffuse

across the membrane. The products of the reaction are then detergent extracted in non-reducing SDS sample buffer, without quenching the reaction, and visualized by western blot. Based on such studies, Dorwart *et al.* suggest that SaMscL is pentameric when overexpressed in the *E. coli* membrane.

In general, crosslinking experiments are interpreted in a qualitative fashion, which can be problematic for the unambiguous determination of the oligomeric state of proteins. Reaction conditions that produce a ladder of bands for easy 'counting' of subunits may lead to incomplete reactions and possibly bands that represent concatamers of cross-linked subunits [34]. Furthermore, reaction conditions that produce fewer bands (or a single band) can lead to products that migrate at molecular weights inconsistent with the true molecular weight of the complex, due to concatamers as well as intramolecular cross-links. A particularly challenging situation occurs when mixtures of different oligomeric states are present; the series of bands for the smaller oligomer(s) will be a subset of the ladder for the largest oligomer. Due to these intrinsic difficulties with cross-linking, alternatives methods have been employed to determine the oligomeric state of MscL.

Genetic fusion of individual MscL subunits, followed by electrophysiology and crosslinking, also led to the conclusion that MscL is hexameric [11]. In these experiments, tandem fusions of MscL subunits generate functional channels in the bacterial membrane, as well as when reconstituted into liposomes. This observation of channels, presumably formed by an even number of subunits, combined with cross-linking experiments provided additional evidence for the hexameric form of EcMscL. A subsequent study using a similar

approach, but including triple subunit fusions, also produced functional channels [26]. In this study, the observation that double subunit fusion constructs form larger channels by SEC than monomeric constructs implied that the wild type monomeric protein assembles into a smaller pentameric channel [26]. Poolman and co-workers later repeated this genetic fusion approach with additional subunits fused together, forming up to six in a single protein chain. Interestingly, all constructs, from single to hexameric subunit fusions, produced functional channels [29]. Remarkably these fusions form functional channels with only one [11, 26] or two [29] amino acid containing linkers between the C-terminus of one subunit and the N-terminus of the next. This arrangement appears to require that the structure of either the N-terminal helix or the C-terminal bundle will be perturbed if the fusions assemble into the same channel. Given that the N-terminus is more sensitive to modification than the C-terminus [35], we would expect the perturbation to be at the C-terminal bundle in cases where these fusions assemble into functional channels.

Mass separation techniques such as SEC-MALS and AUC have been used to study the detergent solubilized molecular weight of several MscL homologs [24, 26, 32]. Sedimentation equilibrium experiments of detergent solubilized EcMscL were unable to observe a single species due to protein aggregation in the detergent/lipid mixture analyzed [26]. The data obtained showed the presence of two or more species and was fit to a two non-interacting species model. Calculating the molecular weight of the EcMscL complex from the experimental reduced molecular mass requires the estimation of detergent mass contribution. Since the amount of detergent bound to each MscL complex is unknown, it is difficult to estimate the respective mass contributions of protein and detergent. Sukharev *et*

al. concluded that MscL forms aggregates in OG and speculated that this may also be true in native membranes based on the observation that addition of lipids to the mixture did not ablate aggregation. Sedimentation equilibrium was also used to calculate the molecular mass of SaMscL [24]. The detergent C_8E_5 was chosen for solubilization, as it is neutrally buoyant in the experimental conditions. The data was fit to a single species model and the protein mass was calculated to be 71.2 kDa, which corresponds to a pentamer (72.2 kDa). Results from a sedimentation velocity experiment of SaMscL solubilized in LDAO led to a calculated protein mass of 62.5 kDa, which corresponds to a tetramer (57.8 kDa). This result is consistent with the oligomeric state observed in the SaMscL(CΔ26) structure. Detergent mass and buoyancy were accounted for with protein-free controls. SEC-MALS has the ability to separate aggregates from single channel protein detergent complexes, but uncertainty in determining the protein and detergent contributions to the scattering mass observed introduces error into the measurement. SEC-MALS has been used with limited success in measuring the molecular weight, and thus oligomeric state, of several MscL homologues. In our laboratory, we have observed good agreement between measured and theoretical pentameric molecular weight for most homologs tested; an exception is EcMscL, which is consistently measured at ~100 kDa, a mass more consistent with a hexameric oligomer [32]. If mixtures of oligomers exist in the detergent solubilized state, currently available SEC columns are not able to separate them.

In view of the challenges in determining the oligomeric state of membrane protein complexes, the Rees group developed a new method to the measure oligomeric state of multiprotein complexes. Oligomer characterization by addition of mass (OCAM) is capable

(in principle) of unambiguous measurement of the oligomeric state for any protein complex and also has the ability to detect mixtures of oligomers, if they exist [32]. OCAM counts protein subunits by selectively removing a mass tag fused to a protein subunit via a short peptide linker. Cleavage of each mass tag through a specific proteolytic site in the linker peptide reduces the total mass of the protein complex by an amount defined by the fused mass partner. Limited proteolysis and separation of the reaction products by size, via Blue Native PAGE, reveals a ladder of reaction products corresponding to the number of subunits present in the target complex plus one additional band for the completed reaction product. The pattern of bands may be used to distinguish the presence of a single homo-oligomer from a mixture of oligomeric states. Using this approach, we were able to determine the oligomeric state of several MscL homologs. Full length MtMscL and SaMscL are pentameric, while the CTD truncations are pentameric and pentamer/tetramer mixtures, respectively. EcMscL is a mixture of hexamer and pentamer forms by OCAM. Unfortunately OCAM, as described above, does not address the *in vivo* oligomeric state because it requires detergent solubilized protein. There are also compatibility issues between certain detergents and BN-PAGE that limit the range of conditions accessible by OCAM, and the possibility that some oligomers may partially dissociate under the conditions used to run BN-PAGE experiments. Nevertheless, the results of our OCAM studies are consistent with the X-ray crystallography for MtMscL and some of the cross linking results.

1.3 Conclusions and focus of thesis

As this overview has established, the oligomeric state (or states) of MscL remains a controversial topic. The multiple oligomeric states observed of MscL not only raise questions concerning the physiologically relevant oligomeric state of MscL, but they also generate uncertainty regarding the effects of modifications of the protein sequence and detergent solubilization (common practices for *in vitro* membrane protein experiments, especially to prepare samples for biophysical and structural studies). An understanding of these effects is crucial for addressing the function of MscL, since ultimately, the oligomeric state of MscL is an important aspect of the pore size (conductance) and the gating mechanism.

The objectives of this thesis are to address these issues through two complementary approaches:

1. determination of the high resolution crystal structure of full length SaMscL

2. preparation and characterization of a comprehensive library of chimeras between EcMscL and SaMscL to identify the sequence elements responsible for distinct structural and functional properties of these close homologues.

As described in the following chapters, despite considerable effort and preparation of numerous crystal forms, it was not possible to obtain high resolution crystals of SaMscL. The chimera library was successfully prepared and the characterization to date has provided exciting clues concerning the regions of MscL that influence oligomeric state and other properties of this fascinating channel.

Table 1-1. Summary of the Reported Oligomeric State of MscL Homologues Determined Using Various Techniques.

X-ray Crystallography

Construct	Detergent	Oligomer	
MtMscL	DDM	5	[17, 18]
SaMscL(CΔ26)	LDAO	4	[19]
EcMscL-CTD	n/a	5	[20]

Oligomer Characterization by Addition of Mass (OCAM)

Construct	Detergent	Oligomer	
MtMscL	DDM	5	[32]
MtMscL(CΔ29)	DDM	5	[32]
SaMscL	DDM	5	[32]
SaMscL(CΔ26)	DDM	5/4 mixture	[32]
EcMscL	DDM	6/5 mixture	[32]
	$C_{12}E_8$	5	[32]

Size Exclusion Chromatography - Multi-Angle Light Scattering (SEC-MALS)

Construct	Detergent	Protein Mass (kDa)	Modifier Mass (kDa)	Error	Oligomer	
MtMscL	DDM	87.4	89.4	5.2%	5	[32]
SaMscL	DDM	72.9	92.1	7.6%	5	[32]
	C_8E_5	72.8	31.3*	0.8kDa	5	[24]
	LDAO	60.0	36.8*	2.8kDa	4	[24]
SaMscL(CΔ26)	DDM	46.5	80.4	8.8%	4	[32]
EcMscL	DDM	103.3	91.5	0.02%	6	[32]

Analytical Ultracentrifugation (AUC)

Construct	Detergent	Method	Protein Mass (kDa)	Oligomer	
SaMscL	C_8E_5	Sedimentation Equilibrium	71.2	5	[24]
	LDAO	Sedimentation Velocity	62.5	4	[24]

Crosslinking

Construct	Detergent	Method	Oligomer	
MtMscL	DDM	Non-specific	5	[18]
	DDM	Carbodiimide chemistry	5	[36]
SaMscL	LDAO	Non specific	4	[19, 24]
	TritonX-100	Non specific	5	[24]
	In membrane	Disulfide trapping	5	[24, 25]
SaMscL(CΔ26)	LDAO	Non specific	4	[19]
	In membrane	Disulfide trapping	5	[25]
EcMscL	DDM	Non specific	5	[18]
	OG	Non specific	6	[11]
	OG	Non specific	5	[26, 28]
	TritonX-100	Non specific	5	[24]
	In membrane	Non specific	5	[26]
	In membrane	Disulfide trapping	5	[23, 27]

References

1.   Walton, T.A., et al., *MscL: channeling membrane tension.* Pflugers Arch, 2014.

2.   Kung, C., *A possible unifying principle for mechanosensation.* Nature, 2005. **436**(7051): p. 647-654.

3.   Coste, B., et al., *Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels.* Science, 2010. **330**(6000): p. 55-60.

4.   Coste, B., et al., *Piezo proteins are pore-forming subunits of mechanically activated channels.* Nature, 2012. **483**(7388): p. 176-81.

5.   Haswell, E.S. and E.M. Meyerowitz, *MscS-like proteins control plastid size and shape in Arabidopsis thaliana.* Curr Biol, 2006. **16**(1): p. 1-11.

6.   Booth, I.R. and P. Louis, *Managing hypoosmotic stress: aquaporins and mechanosensitive channels in Escherichia coli.* Curr Opin Microbiol, 1999. **2**(2): p. 166-9.

7.   Sukharev, S.I., et al., *Two types of mechanosensitive channels in the Escherichia coli cell envelope: solubilization and functional reconstitution.* Biophys J, 1993. **65**(1): p. 177-83.

8.   Martinac, B., et al., *Pressure-Sensitive Ion Channel in Escherichia-Coli.* Proceedings of the National Academy of Sciences of the United States of America, 1987. **84**(8): p. 2297-2301.

9.   Levina, N., et al., *Protection of Escherichia coli cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity.* EMBO J, 1999. **18**(7): p. 1730-7.

10.   Sukharev, S.I., et al., *A large-conductance mechanosensitive channel in E. coli encoded by mscL alone.* Nature, 1994. **368**(6468): p. 265-8.

11.   Blount, P., et al., *Membrane topology and multimeric structure of a mechanosensitive channel protein of Escherichia coli.* EMBO J, 1996. **15**(18): p. 4798-805.

12.   Hase, C.C., et al., *Cross-linking studies and membrane localization and assembly of radiolabelled large mechanosensitive ion channel (MscL) of Escherichia coli.* Biochem Biophys Res Commun, 1997. **232**(3): p. 777-82.

13.   Kung, C., B. Martinac, and S. Sukharev, *Mechanosensitive channels in microbes.* Annu Rev Microbiol, 2010. **64**: p. 313-29.

14.   Moe, P.C., P. Blount, and C. Kung, *Functional and structural conservation in the mechanosensitive channel MscL implicates elements crucial for mechanosensation.* Mol Microbiol, 1998. **28**(3): p. 583-92.

15.   Spencer, R.H., G. Chang, and D.C. Rees, *'Feeling the pressure': structural insights into a gated mechanosensitive channel.* Curr Opin Struct Biol, 1999. **9**(4): p. 448-54.

16.   Rees, D.C., G. Chang, and R.H. Spencer, *Crystallographic analyses of ion channels: lessons and challenges.* J Biol Chem, 2000. **275**(2): p. 713-6.

17.   Steinbacher, S., et al., *Structures of the prokaryotic mechanosensitive channels MscL and MscS.* Mechanosensitive Ion Channels, Part A, 2007. **58**: p. 1-24.

18.   Chang, G., et al., *Structure of the MscL homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel.* Science, 1998. **282**(5397): p. 2220-6.

19.   Liu, Z., C.S. Gandhi, and D.C. Rees, *Structure of a tetrameric MscL in an expanded intermediate state.* Nature, 2009. **461**(7260): p. 120-4.

20.   Walton, T.A. and D.C. Rees, *Structure and stability of the C-terminal helical bundle of the E. coli mechanosensitive channel of large conductance.* Protein Sci, 2013. **22**(11): p. 1592-601.

21.   Beckstein, O., K. Tai, and M.S. Sansom, *Not ions alone: barriers to ion permeation in nanopores and channels.* J Am Chem Soc, 2004. **126**(45): p. 14694-5.

22.   Malashkevich, V.N., et al., *The crystal structure of a five-stranded coiled coil in COMP: a prototype ion channel?* Science, 1996. **274**(5288): p. 761-5.

23. Anishkin, A., *On the Conformation of the COOH-terminal Domain of the Large Mechanosensitive Channel MscL.* The Journal of General Physiology, 2003. **121**(3): p. 227-244.

24. Dorwart, M.R., et al., *S. aureus MscL is a pentamer in vivo but of variable stoichiometries in vitro: implications for detergent-solubilized membrane proteins.* PLoS Biol, 2010. **8**(12): p. e1000555.

25. Iscla, I., R. Wray, and P. Blount, *The oligomeric state of the truncated mechanosensitive channel of large conductance shows no variance in vivo.* Protein Sci, 2011. **20**(9): p. 1638-42.

26. Sukharev, S.I., M.J. Schroeder, and D.R. McCaslin, *Stoichiometry of the large conductance bacterial mechanosensitive channel of E. coli. A biochemical study.* J Membr Biol, 1999. **171**(3): p. 183-93.

27. Yang, L.M., et al., *Three routes to modulate the pore size of the MscL channel/nanovalve.* ACS Nano, 2012. **6**(2): p. 1134-41.

28. Yoshimura, K., J. Usukura, and M. Sokabe, *Gating-associated conformational changes in the mechanosensitive channel MscL.* Proc Natl Acad Sci U S A, 2008. **105**(10): p. 4033-8.

29. Folgering, J.H., et al., *Lactococcus lactis uses MscL as its principal mechanosensitive channel.* J Biol Chem, 2005. **280**(10): p. 8784-92.

30. Saint, N., et al., *A hexameric transmembrane pore revealed by two-dimensional crystallization of the large mechanosensitive ion channel (MscL) of Escherichia coli.* J Biol Chem, 1998. **273**(24): p. 14667-70.

31.   Ornatska, M., et al., *Biomolecular stress-sensitive gauges: surface-mediated immobilization of mechanosensitive membrane protein.* J Am Chem Soc, 2003. **125**(42): p. 12722-3.

32.   Gandhi, C.S., T.A. Walton, and D.C. Rees, *OCAM: a new tool for studying the oligomeric diversity of MscL channels.* Protein Sci, 2011. **20**(2): p. 313-26.

33.   Sukharev, S., et al., *The gating mechanism of the large mechanosensitive channel MscL.* Nature, 2001. **409**(6821): p. 720-4.

34.   Azem, A., et al., *Cross-linking of porin with glutardialdehyde: a test for the adequacy of premises of cross-linking theory.* Biochim Biophys Acta, 1995. **1243**(2): p. 151-6.

35.   Blount, P., et al., *Single residue substitutions that change the gating properties of a mechanosensitive channel in Escherichia coli.* Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11652-7.

36.   Maurer, J.A., et al., *Comparing and contrasting Escherichia coli and Mycobacterium tuberculosis mechanosensitive channels (MscL). New gain of function mutations in the loop region.* J Biol Chem, 2000. **275**(29): p. 22238-44.

CHAPTER II

SAECO CHIMERAS

2.1 Background

Initial studies of EcMscL and SaMscL by OCAM prompted questions concerning the sequence determinants underlying the variation in oligomeric state [1]. EcMscL and SaMscL share 51% sequence identity [Fig 2-1B], yet when purified in the detergent DDM, SaMscL is stringently a pentamer, while EcMscL is a mixture of pentamers and hexamers. The purpose of the study detailed in this chapter is to use chimeras of SaMscL and EcMscL to probe the sequence level determinants of oligomeric state and other structural and biophysical properties.

Chimeric proteins have been widely used to investigate and manipulate the structural and functional behavior of proteins [2-5]. Chimeras can have properties within the range of the parent proteins' or can develop more exaggerated or novel properties not seen in any parent protein. A rich literature on chimeric proteins documents chimeras of altered enzymatic activity, substrate specificity, thermostability, and structural conformations. An important aspect of the design of chimeras is the assignment of the constituent elements. The SCHEMA method for generating chimeric constructs was developed here at Caltech by Professor Arnold and colleagues [6, 7]. With this algorithm, the segments in the chimeras are identified using a SCHEMA disruption energy which is proportional to the number of amino acid contacts in the parent structure that are predicted to be broken in the chimera. Chimeric proteins designed by SCHEMA and other methods have

been used to design thermostable cytochrome P450s [8], screen for thermostabilizing sequence blocks of cellobiohydrolase class I enzymes [9], improve the activity of beta-lactamases [10], and to alter the substrate profile of subtilisins [11].

Chimeras have also been used to great effect in the study of channels. As an important example, the crystal structure of a chimeric voltage-dependent K^+ (Kv) channel was solved at 2.4 Å resolution [3]. The voltage sensor paddle domain of rat Kv1.2 was replaced by that of rat Kv2.1 and the crystal structure revealed a novel conformation of the channel that led to new mechanistic insights. Of great relevance to the present study, Yang *et al.* successfully used chimeras of EcMscL and SaMscL to probe the sequence basis of variations in gating behavior in MscL [12]. The study examined the effects of substitutions in a small region of the protein, ranging from the periplasmic end of TM1 through the loop to the periplasmic end of TM2, on channel open dwell times. Systematic swaps between SaMscL and EcMscL in this region led to the discovery that the EcMscL TM1/periplasmic loop interface (residues 44 – 49) and the SaMscL periplasmic loop (residues 48 - 62) confer long open dwell times, while the EcMscL periplasmic loop (residues 50 - 74) and the SaMscL TM1/periplasmic loop interface (residues 42 - 47) confer short openings. A mutant ESE50-77, which had a combination of the EcMscL TM1/periplasmic loop interface and the SaMscL periplasmic loop, had an extremely long open dwell time, longer than either parent. A single residue, I49 in EcMscL and its equivalent F47 in SaMscL, was identified for its importance in gating kinetics. In particular the mutant SaMscL F47L showed a delay in both channel opening and closing as well as severe hysteresis.

In the present work, we have exploited chimeras of SaMscL and EcMscL to systematically explore the relationship between oligomeric state and sequence space of the full length MscL protein. As discussed in Chapter 1, a single subunit of MscL can be divided into five distinct structural regions based on secondary structure; N-terminus (N-term), transmembrane helix 1 (TM1), periplasmic loop (Loop), transmembrane helix 2 (TM2) and the C-terminal domain (C-Term) [Fig. 2-1]. A full chimera set of all the combinations of these regions from SaMscL and EcMscL will contain $2^5 = 32$ constructs [Fig. 2-2]. Of these 32 chimeras, two are parent constructs, ten (2x5) are "single variants" containing 4/5 segments from one parent, and twenty (4x5) are "double variants" containing 3/5 segments from one parent. This library of chimeras, named SaEco, enables a coarse but complete scan of the MscL sequence. The working hypothesis is that there may be a sequence commonality between the SaEco chimeras that exhibit the same oligomeric state (or mixture of oligomeric states), i.e., MscL chimeras containing a certain structural region from specific parents (or combinations of such structural regions) may have the same oligomeric state. Of course, these chimeras can also be used as a tool to explore the sequence dependence of other properties of MscL such as protein expression, thermostability, crystallizability, gating behavior, and ability to protect bacteria from osmotic down-shock.
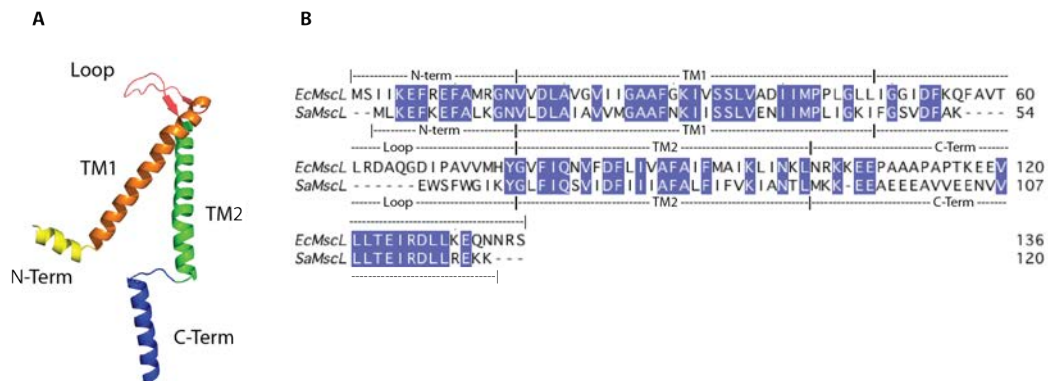
Figure 2-1. Crystal Structure of a Single Subunit of MtMscL and Sequence Alignment of EcMscL and SaMscL. (A) Crystal structure of a single subunit of MtMscL (PDB ID: 2OAR) with structural regions highlighted by color: N-term (yellow), TM1 (orange), Loop (red), TM2 (green), and C-term (blue). (B) Sequence alignment of SaMscL and EcMscL, which share a 51% sequence identity, with conserved residues highlighted in blue. The boundaries of the structural regions of EcMscL are labeled above its sequence and those of SaMscL are labeled below its sequence.

2.2 SaEco Chimera Library Design

The amino acid sequences of SaMscL and EcMscL were divided into five structural regions based on secondary structural elements. The crystal structure of SaMscL(CΔ26) was used to designate the regions of SaMscL. The corresponding elements of EcMscL were assigned based on homology to SaMscL and MtMscL [Fig 2-1B]. EcMscL was divided as follows: N-term (M1 - N15), TM1 (V16 - L48), Loop (I49 – Y75), TM2 (G76 – L102), and C-term (N103 – S136). SaMscL was divided as follows: N-term (M1 – N13), TM1 (V14 - I46), Loop (F47 – Y63), TM2 (G64 – L90), and C-term (M91 – K120). A library of 32 chimeras of all the combinations of the structural regions of EcMscL and SaMscL was constructed (SaEco27 – SaEco56) [Fig 2-2], as well as a library of sGFP fusions (SaEco27-sGFP – SaEco56-sGFP). A fluorescent protein fusion partner was chosen as the mass tag for OCAM measurements because of the ease of tracking the protein during expression and purification as well as its possible utility in

25

other assays. The superfolder GFP variant (sGFP) was chosen because wild-type GFP forms weak dimers at high concentrations and the A206V mutation disrupts the dimerization interface, creating a homogeneous monomeric population [13]. For the purposes of characterizing the oligomeric state of a multimeric protein, it is important that the fusion partner does not introduce any intersubunit interactions that could bias or occlude the protein's native state. The protein sequence of all the chimeras can be found in the Appendix section.

The chimeras in this study were numbered in the order they were designed in the cloning scheme. For example, SaEco27 – 34 was the first set of chimeras to be cloned and had two gene fragments, one from each parent, stitched together at a single seam point. The other chimeras, SaEco35 – 56, have multiple gene fragments from each parent distributed in an alternating fashion with multiple seam points. The numbering scheme of the SaEco chimeras starts at 27, because the first set of gene fragments amplified from the parent sequences were numbered 1 - 26, and to avoid confusion while cloning, the SaEco constructs were numbered starting at 27. Although the numbering scheme is odd, it introduces an element of randomization that reduces bias when analyzing data, because the name of the chimera is not easily translated to its sequence.

The SaEco chimera library was successfully cloned using homologous recombination and site-directed mutagenesis methods and placed in a pET15b vector for expression (see Methods and Materials for more detail). Each construct was cloned both with and without an sGFP fused at the C-terminal end via a short linker with a TEV recognition site.
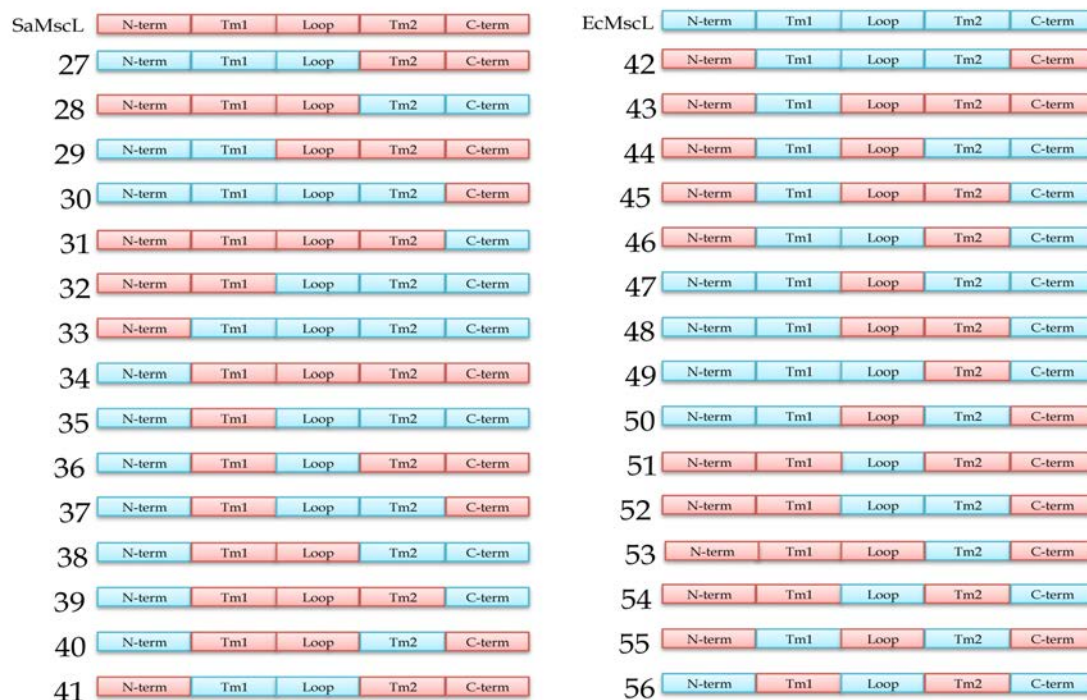
Figure 2-2. Schematic Illustration of the SaEco Chimera Library. A schematic illustration of all 32 SaEco chimeras (SaMscL, EcMscL, and SaEco27 – SaEco56) with structural regions delineated as labeled boxes. Blue boxes represent EcMscL parent origin and red boxes represent SaMscL parent origin. The protein sequence of all the chimeras can be found in the Appendix section.

2.3 Expression and Purification of SaEco Chimeras

The first step after cloning the SaEco chimeras was to test their expression levels. Whenever the sequence of a protein is altered, there is always the possibility of altered (decreased or increased) expression. All constructs were expressed in *BL21Δmscl E. coli* strain, cultured in auto-induction media, solubilized using the detergent DDM, and purified via metal affinity chromatography. Fluorescence images of cell pellets expressing SaEco-sGFP constructs clearly reveal variations in expression levels [Fig 2-3A]. Notably, SaEco 27-sGFP, 41-sGFP, 46-sGFP, 51s-GFP, and 54-sGFP have

27

significantly lower expression levels than the rest of the library. The low fluorescence is correlated with poor yields of the protein after purification [Fig. 2-3B], precluding the possibility that GFP is removed or doesn't mature in these constructs. As discussed later, there are multiple possible mechanisms for the low expression levels of these constructs; however, the presence of a cell pellet of comparable size to their counterparts indicates it is not due to a gain of function phenotype that leads to cell lysis. All constructs are cultured under the same conditions, which implies that differences are not due to variations in the cell growth protocols. Therefore it is presumed that the low expression level in these five constructs is due to a sequence level variation that leads to one or more problems in the DNA transcription, protein translation, folding, assembly, or degradation targeting pathways. The common sequence elements between SaEco 27, 41, 46, 51, and 54 include the EcMscL Loop and SaMscL TM2. This trend is not completely deterministic, however, since SaEco 36, 49 and 56 also have the Loop of EcMscL and the TM2 of SaMscL, yet express comparatively better. Larger scale cultures of SaEco 27, 41/-sGFP, 46/-sGFP, 51/-sGFP, and 54/-sGFP were able to yield enough protein for analysis. SaEco27-sGFP never yielded enough protein for analysis.
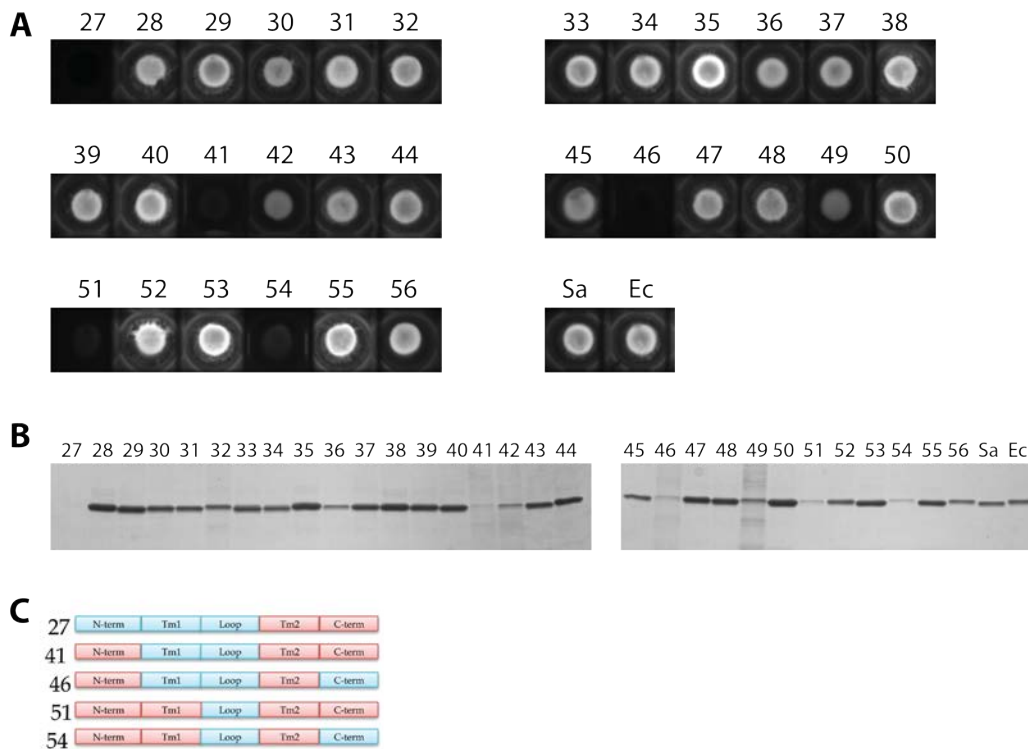
Figure 2-3. Test Expression of SaEco-sGFP Fusion Proteins Identifies Five Poorly Expressing Chimeras. (A) Fluorescence images of cell pellets expressing SaEco-sGFP constructs clearly reveal variations in expression levels. Cell pellets containing SaEco 27-sGFP, 41-sGFP, 46-sGFP, 51s-GFP, and 54-sGFP have little to no detectable fluorescence. (B) SDS-PAGE gel of purified SaEco-sGFP proteins further demonstrates that those five constructs have a significantly lower protein yield than the rest of the chimera library. (C) Schematic illustration of poorly expressing SaEco chimeras, with blue boxes representing EcMscL parent origin and red boxes representing SaMscL parent origin. The common sequence elements between SaEco 27, 41, 46, 51, and 54 include the EcMscL Loop and SaMscL TM2.

2.4 OCAM of SaEco Chimeras

All SaEco-sGFP constructs were analyzed by the OCAM technique as described in Chapter 1 [1]. Each protein was incubated with TEV protease at 34°C (the optimal temperature of TEV protease activity) and sampled at 15, 60, and 240 minute time points. Negative controls were also run without added TEV and sampled at the 0 minutes time point and after 240 minutes incubation at 34°C. The products of these reactions were analyzed by Blue Native PAGE (BN-PAGE), and the number of bands in the lanes were

quantified by densitometry and compared to that of wild-type EcMscL and SaMscL [Fig. 2-4]. SaMscL has six distinct bands, excluding the highest molecular weight band, which is believed to be a dimer of pentamers, and the lowest molecular weight band, which is sGFP liberated from the fusion protein upon TEV protease cleavage [1]. Six bands correspond to the expected proteolytic products from a pentameric species as it loses sGFP tags in a stepwise fashion (an n-mer will generate n+1 distinct species containing 0, 1, … n sGFPs; this assumes that all species with the same number of sGFPs will migrate as a single band on BN-PAGE). It should be noted that at the 0 time point, even before TEV is added, there is a second faint band below the major band. The major band corresponds to the mass of the entire fusion protein complex, and the second lower band corresponds to the mass of the fusion proton complex minus a single sGFP tag. This loss of the sGFP mass partner is presumed to be due to proteolytic activity in the expression host cell [1].

The banding pattern for EcMscL is more complicated than that of SaMscL. With EcMscL, eight distinct bands (excluding the dimer and s-GFP bands) can be resolved. At first glance, one might assume this indicates a heptameric complex (n+1, with n=7). However, based on work done in the original OCAM paper to estimate the mass of the bands based on migration distance, it was determined that this banding pattern is consistent with a mixture of hexamers and pentamers [1]. The principles of OCAM dictate that a mixture of hexamers and pentamers would have 13 distinct bands $((n_a+1)+(n_b+1)$, with $n_a=6$ and $n_b=5)$. However, it is not possible to resolve all 13 bands with the current BN-PAGE protocol because some of the products of the reaction are

similar in mass (within 8.6 kDa) and the estimated resolution limit of the gel is 13 kDa [1]. Therefore, multiple bands can overlap each other. Extensive efforts were made to improve the separation of bands on the gel by trying different percentage polyacrylamide non-gradient and gradient gels and extending electrophoresis times. These efforts were not successful in resolving additional bands, however, which ultimately will likely require construction of new fusions with a mass increment that would be more cleanly resolvable.

As proof of principle that it is possible for a mixture of oligomeric states to be fully resolved by OCAM, we can look to the OCAM analysis of SaMscL(CΔ26) [1]. SaMscL(CΔ26) is a mixture of pentamers and tetramers, and when analyzed by OCAM the products of a proteolytic reaction ran as eleven distinct bands by BN-PAGE ((n_a+1)+(n_b+1), with n_a=5 and n_b=4). The products for this construct have masses that differ by more than 13 kDa, which are resolvable by this gel system. For the purposes of this study, we will designate any SaEco chimera that runs as six distinct bands as a pentamer and those that run as eight distinct bands as a pentamer/hexamer mixture.

All SaEco chimeras (except the poorly expressing SaEco27-sGFP) were analyzed by OCAM. A majority of them ran as pentamers, like SaMscl, while most of the remaining ran as either a hexamer/pentamer mixture like EcMscL or a tetramer/pentamer mixture. The chimeras that strongly displayed the characteristic eight-band pattern of EcMscL are SaEco 30-sGFP, 37-sGFP, 40-sGFP, and 50-sGFP [Fig. 2-5]. SaEco 30, 37, 40, and 50 have in common the sequence elements of the N-term and TM2 of EcMscL and the C-

Term of SaMscL [Fig. 2-5F]. Please refer to the Appendix section for gels of the OCAM results of all the chimeras.
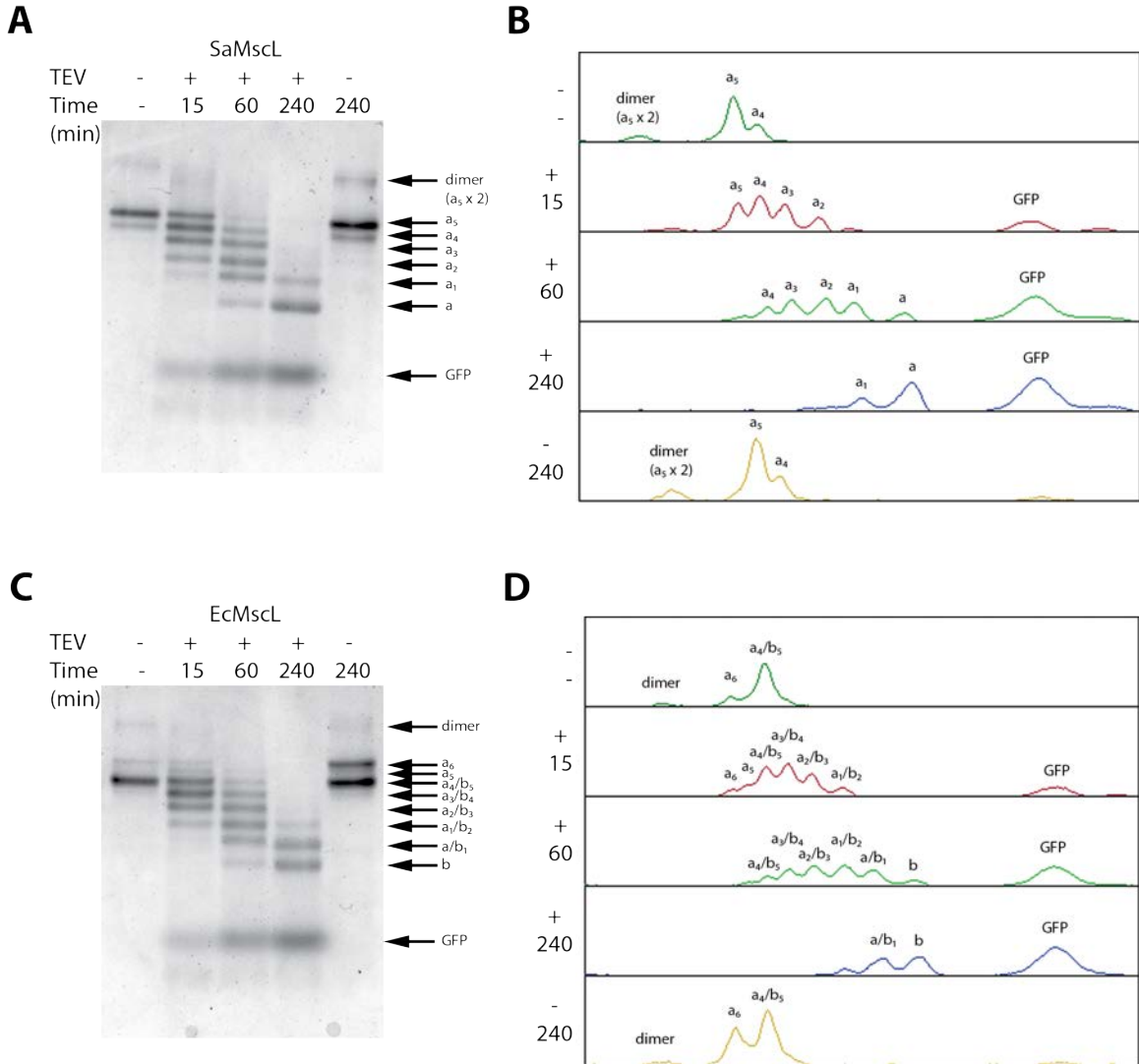
Figure 2-4. OCAM Measurements Reveal that SaMscL is a Pentamer and EcMscL is a Mixture of Pentamers and Hexamers. (A) BN-PAGE separation of SaMscL-sGFP OCAM reaction products and controls. TEV protease reactions were quenched at 15, 60, and 240 minute time points. Controls were run without added TEV and sampled at the 0 minutes time point and after 240 minutes incubation at 34°C. A total of six reaction products are observed corresponding to the loss of 0-5 sGFP mass tags of a pentameric species. The cleaved sGFP is observed at a lower gel migration distance. (B) Densitometry traces of (A) with each peak labeled to show the expected OCAM reaction product. (C) BN-PAGE separation of EcMscL-sGFP OCAM reaction products and controls. A total of eight reaction products are observed in a banding pattern consistent with a mixture of hexameric and pentameric species. (D) Densitometry traces of (C) with each peak labeled to show the expected OCAM reaction product.
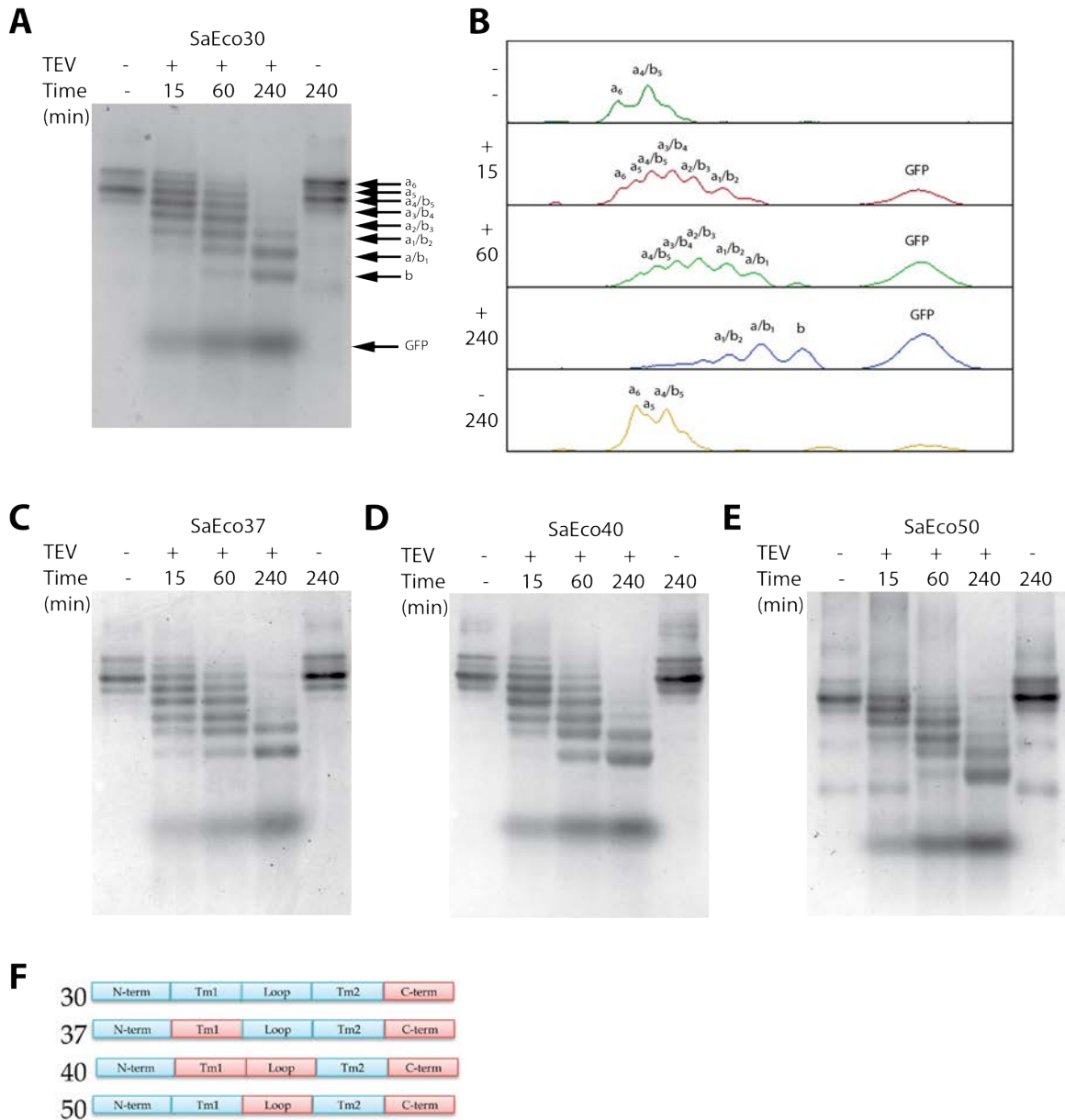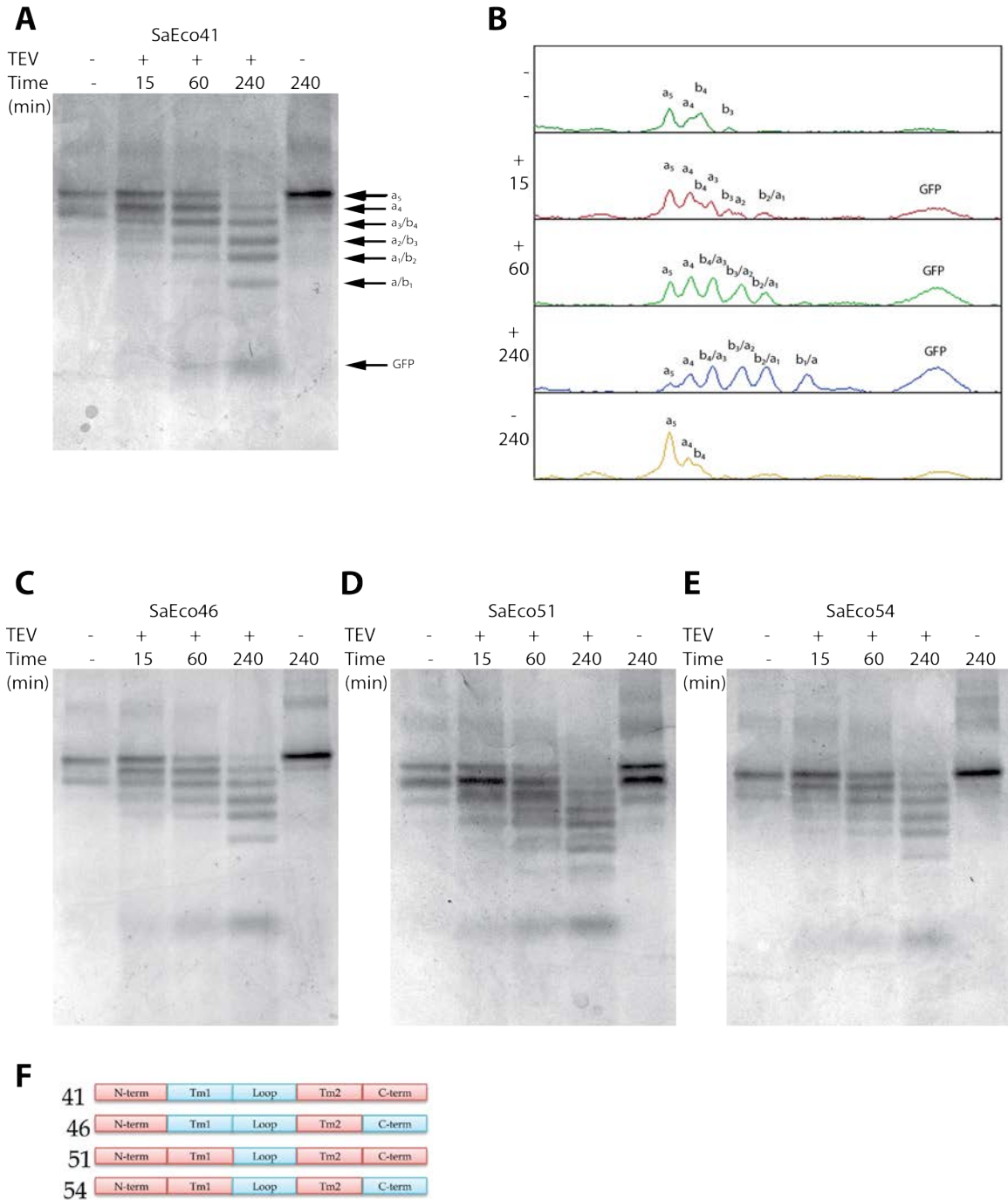
Figure 2-5. OCAM Measurements Reveal Four Chimeras that are a Mixture of Pentamers and Hexamers. (A) BN-PAGE separation of SaEco30-sGFP OCAM reaction products and controls. A total of eight reaction products are observed in a banding pattern consistent with a mixture of hexameric and pentameric species. (B) Densitometry traces of (A) with each peak labeled to show the expected OCAM reaction product. (C - E) BN-PAGE separation of OCAM reaction products and controls of SaEco 37-sGFP, 40-sGFP, and 50-sGFP, respectively. A total of eight reaction products are observed in a banding pattern consistent with a mixture of hexameric and pentameric species. (F) Schematic illustration of SaEco chimeras that are a mixture of pentamers and hexamers, with blue boxes representing EcMscL parent origin and red boxes representing SaMscL parent origin. The common sequence elements between SaEco 30, 37, 40, and 50 include the EcMscL N-term and TM2 and the SaMscL C-Term.

33

SaEco 41-sGFP, 46-sGFP, 51-sGFP, and 54-sGFP, which coincidentally are poorly expressing chimeras, have a mixed oligomeric stage, postulated to be a mixture of tetramers and pentamers [Fig. 2-6]. In the densitometry plot of SaEco41-sGFP [Fig. 2-6B], at certain time points, additional bands can be seen between the proteolytic products of a pentameric species. There is a band between the a_3 and a_4 bands (mw=168 kDa and 195 kDa, respectively) and another one between the a_2 and a_3 bands (mw = 141 kDa and 168 kDa, respectively). Based on the relative migration distance of the additional bands to that of the pentamer proteolytic product bands, a rough estimation of the molecular weights of the additional bands correlate with that of a tetrameric species (b_4, mw = 177 kDa and b_3, mw = 151 kDa).

Figure 2-6. OCAM Measurements Reveal Four Chimeras that are a Mixture of Tetramers and Pentamers. (A) BN-PAGE separation of SaEco41-sGFP OCAM reaction products and controls. Reaction products are observed in a banding pattern consistent with a mixture of tetrameric and pentameric species. (B) Densitometry traces of (A) with each peak labeled to show the expected OCAM reaction product. (C - E) BN-PAGE separation of OCAM reaction products and controls of SaEco 46-sGFP, 51-sGFP, and 54-sGFP, respectively. Reaction products are observed in a banding pattern consistent with a mixture of tetrameric and pentameric species. (F) Schematic illustration of SaEco chimeras that are a mixture of tetramers and pentamers, with blue boxes representing EcMscL parent origin and red boxes representing SaMscL parent

origin. The common sequence elements between SaEco 41, 46, 51, and 54 include the EcMscL Loop and the SaMscL N-Term and TM2.

Intriguingly, there were several chimeras that exhibited banding patterns that appear to represent unique oligomeric states different from those observed in either parent construct or in the literature. SaEco 30-sGFP, 50-sGFP, and 55-sGFP have two additional lower molecular weight protein bands that correspond to oligomeric states smaller than tetramer [Fig. 2-7]. In lanes 1, 2, and 5 of the gel of SaEco55-sGFP [Fig. 2-7A], there is one band below the "a" band (mw = 84 kDa) and another one between the a_1 and a_2 bands (mw = 110 kDa and 137 kDa, respectively). Based on the position of the proteolysis bands, a rough estimation of the molecular weights of the additional bands correlate with that of a trimer (c_3, mw = 131 kDa) and a dimer (d_1, mw = 60 kDa). Precise molecular weights of protein bands on a gradient BN-PAGE gel cannot be easily derived, because the relationship between protein molecular weight and migration distance can only be empirically established through the use of appropriate standards, especially with a gradient gel. Nevertheless, we can be certain that the bands represent oligomers of MscL because there are no other protein components in the 1st and 5th lane of each gel. SDS-PAGE gels of the purified proteins show that the preparations are >99% pure. It should also be noted that there are no detectable proteolytic products of TEV digestion of the "c" and "d" oligomers. The bands appear to fade over the course of the TEV reaction.
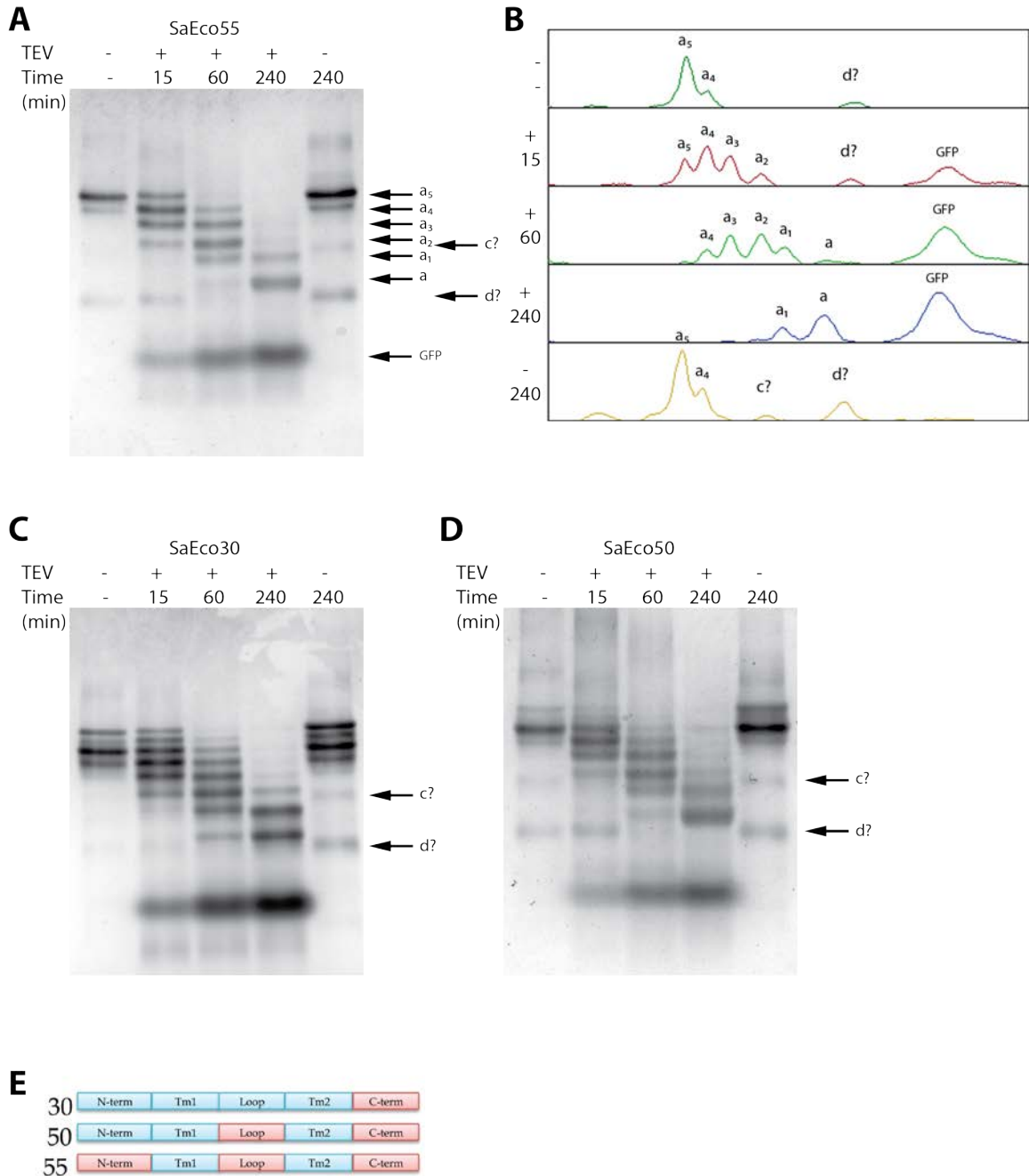
Figure 2-7. SaEco30-sGFP, SaEco50-sGFP, and SaEco55s-GFP have Lower Molecular Weight Oligomers Postulated to be Dimers and Trimers. (A) BN-PAGE separation of SaEco55-sGFP OCAM reaction products and controls. Two additional oligomer bands are observed at migration distances that likely correspond to dimers and trimers. (B) Densitometry traces of (A) with each peak labeled to show the expected OCAM reaction product. (C - D) BN-PAGE separation of OCAM reaction products and controls of SaEco30-sGFP and SaEco50-sGFP, respectively. Two additional oligomer bands are observed at migration distances that likely correspond to dimers and trimers. (E) Schematic illustration of SaEco chimeras that are a mixture of tetramers and pentamers, with blue boxes representing EcMscL parent origin and red boxes representing SaMscL parent origin. The common sequence elements between SaEco30, 50 and 55 include the EcMscL TM1 and TM2 and the SaMscL C-Term.

2.5 Heat Sensitivity of SaEco Chimeras

Using OCAM data, pentamer/hexamer mixed oligomeric state chimeras were unambiguously identifiable. Somewhat surprisingly, however, was the initial difficulty in interpreting the BN-PAGE of non-fusion SaEco chimeras. Non-fusion SaEco chimeras are anticipated to give one and two bands on BN-PAGE for single and mixed oligomeric states, respectively. With the protocols initially used, however, even wild-type EcMscL did not display the characteristic double band of mixed oligomeric state constructs [Fig. 2-8]. To troubleshoot the discrepancy between the interpretations of the OCAM data as compared to the SaEco non-fusion data, differences in protocols between these two approaches were systematically tested. These tests included assessing the influence of the following conditions on the non-fusion SaEco chimeras: addition of TEV buffer (20 mM Tris pH 7.5, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT), addition of TEV protease, addition of iodoacetamide (a TEV quencher) and incubating samples at 34°C for 240 minutes, before running BN-PAGE gels of the non-fusion samples. These initial results revealed that temperature had a significant effect on the oligomeric state of certain SaEco chimeras [Fig. 2-8]. Detailed examination of the OCAM data also confirmed this phenomena; certain hexamer/pentamer mixed oligomeric state constructs, including EcMscL-sGFP and SaEco50-sGFP, showed a marked increase in the population of hexamers versus pentamers after incubation at 34°C for 240 minutes in the absence of TEV [Fig. 2-4C and Fig. 2-5E].

To probe the effects of temperature on oligomeric state, pentamer/hexamer mixed oligomeric state SaEco chimeras were exposed to mild heating at 34 °C for 240 minutes,

and then visualized on BN-PAGE gels. The results showed that EcMscL, SaEco30, and SaEco50 are heat sensitive, while SaEco37 and SaEco40 do not appear to be [Fig. 2-8]. SaEco30 had the most marked shift in the percentage of the population of hexamers after heating, with the hexameric population increasing from 32% to 57% of the total protein. Before heating, EcMscl and SaEco50 exhibited a poorly defined hexamer band, which appeared to be a shoulder on the pentamer band that could not be integrated by gel analysis software. After heating, however, a clear hexamer band appeared corresponding to 28% and 42% for EcMscL and SaEco50, respectively. In contrast, SaEco37 and SaEco40 showed insignificant changes in hexamer population upon heating (18% to 20% and 22% to 17%, respectively). The sequence commonality between SaEco30 and SaEco50 are the N-term, TM1, and TM2 of EcMscL and the C-term of SaMscL. Again, these motifs are not strictly deterministic, as EcMscL does not have the C-term of SaMscL, but it still shows heat sensitivity.
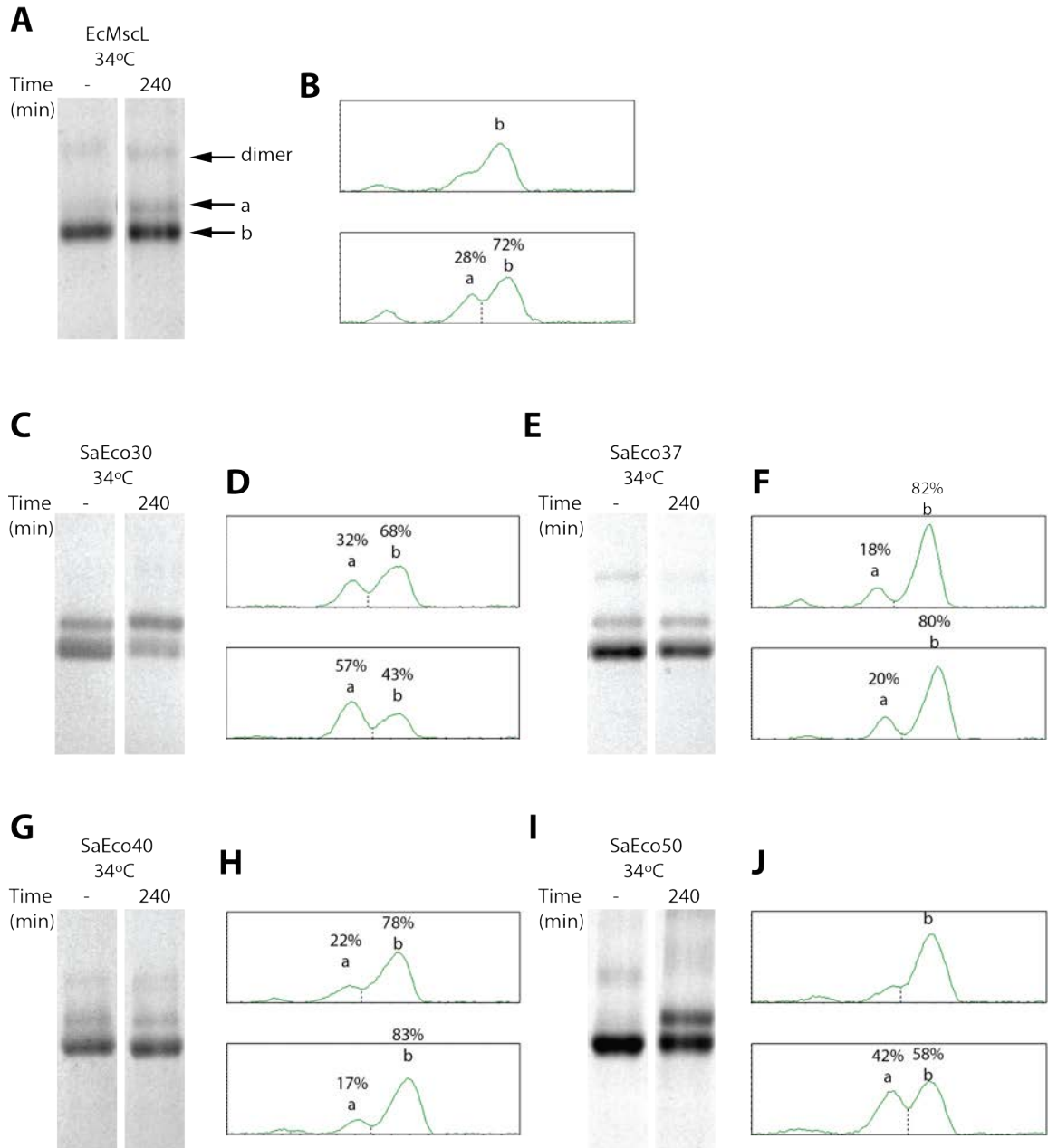
Figure 2-8. Temperature Effects on Hexamer/Pentamer Mixed Oligomeric State SaEco Chimeras. (A, C, E, G, and 1) BN-PAGE separation of EcMscL, SaEco30, 37, 40, and 50, respectively, before and after incubation at 34 °C for 240 minutes. EcMscL, SaEco30, and 50 show an increase in the percentage of the population of hexamers after heating. (B, D, F, H, and J) Densitometry traces of (A, C, E, G, and 1, respectively) with each peak labeled to show the percentage of each oligomer.

To further explore the time dependence of this shift in population of oligomeric states for these constructs, a time course experiment was conducted. EcMscL, SaEco30, and SaEco50 were heated at 34°C with time points sampled at 30, 60, 120, and 240 minutes

and then analyzed by BN-PAGE [Fig. 2-9]. For EcMscL, the population of hexamers increased during this time period from 20% ➔ 23% ➔ 24% ➔ 30% ➔ 36%. For SaEco30, the population of hexamers increased for the first 120 minutes, but seemed to have plateaued by the 240 minute mark: 35% ➔ 43% ➔ 51% ➔ 57% ➔ 59%. SaEco50 showed a sharp increase in the population of hexamers from an indistinct small shoulder to 37% after only 30 minutes of incubation after which the population seems to have plateaued: ~0% ➔ 37% ➔ 41% ➔ 42% ➔ 43%.
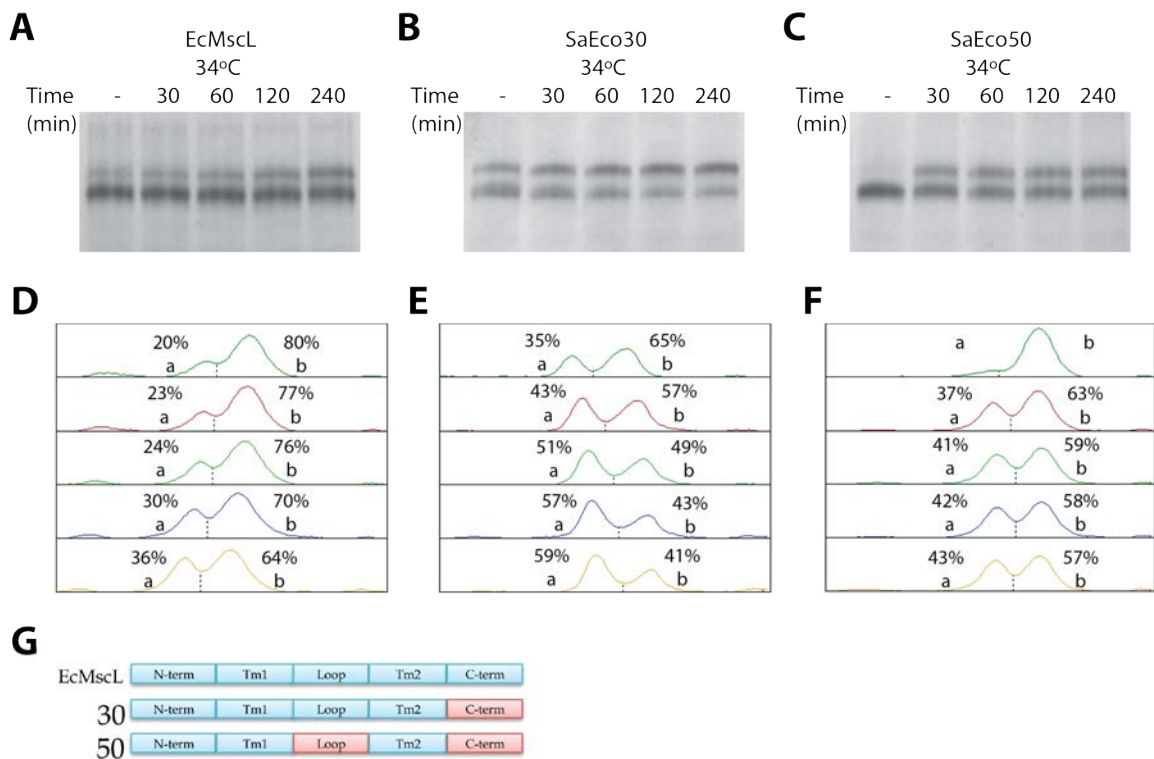
Figure 2-9. Time Course Temperature Experiments Track the Shift in Population of Oligomers Over Time (A, B, and C) BN-PAGE separation of time course temperature experiments of EcmscL, SaEco30, and SaEco50, respectively. Oligomeric state is monitored after incubation at 34°C for 30, 60, 120, and 240 minutes. (D, E, and F) Densitometry traces of (A, B, and C, respectively) with each peak labeled to show the percentage of each oligomer. (G) Schematic illustration of EcMscL, SaEco30, and SaEco50, with blue boxes representing EcMscL parent origin and red boxes representing SaMscL parent origin. The common sequence elements between EcMscL SaEco30, and SaEco50 include the EcMscL N-term, TM1, and TM2.

2.6 Subunit Exchange of SaEco Chimeras

Evidence that the oligomeric state of some MscL constructs is dynamic motivated experiments to explore the underlying mechanism of this phenomenon. If in a closed system (such as a fixed sample volume under mild heating conditions) the equilibrium between pentameric and hexameric complexes shifts, there should be evidence of subunit exchange between complexes. An experiment was accordingly designed to test for subunit exchange in MscL constructs. Purified SaEco-sGFP constructs and their corresponding SaEco non-fusion constructs would be mixed in a tube, incubated at various temperatures and visualized by BN-PAGE. If there is indeed subunit exchange between sGFP fusion and non-fusion complexes, one would expect to see a ladder of intermediate species similar to that of an OCAM experiment. SaEco30 was chosen for the preliminary experiments, because it showed the most marked change in oligomeric state after incubation at 34 °C. SaEco30 and SaEco30-sGFP were incubated both separately and together at 4 °C and 34 °C for four hours [Fig. 2-10A]. In the control experiment when SaEco30 and SaEco30-sGFP were mixed together and incubated at 4 °C, there are no intermediate species. However, when incubated together at 34 °C, as predicted, we see a ladder of heterogeneous complexes similar to that of an OCAM experiment. A second time course experiment was conducted to monitor the evolution of these intermediate species. SaEco30 and SaEco30-sGFP were incubated at 34 °C and sampled at 30, 60, 120, and 240 minutes [Fig. 2-10B & C]. As early as 30 minutes incubation, intermediate species begin to appear and continue to increase in abundance over the course of the experiment. The results of these preliminary experiments confirmed that subunit

exchange can occur, and therefore subunit exchange experiments were conducted for the rest of the SaEco chimeras.
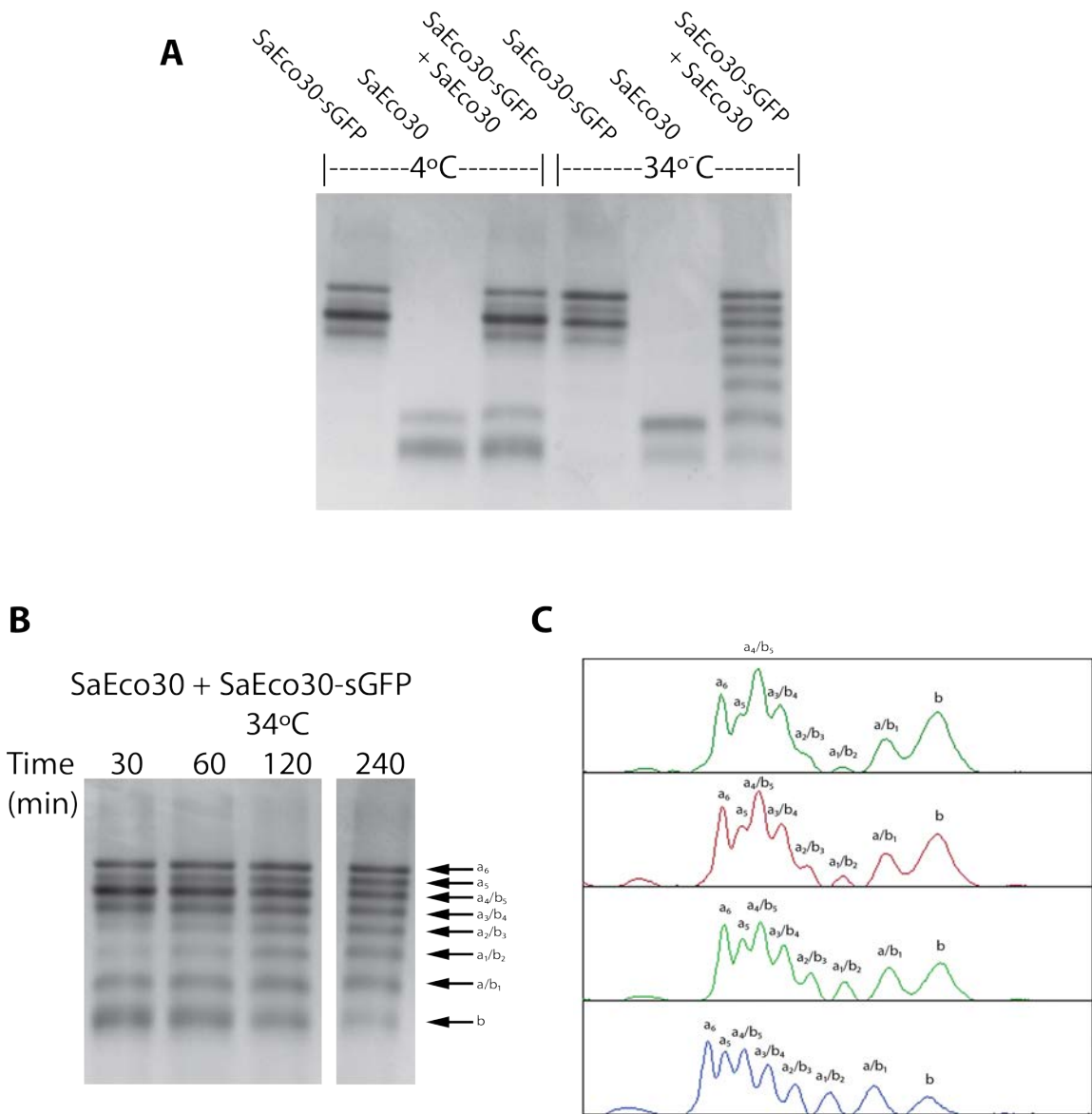
Figure 2-10. Preliminary Experiments Reveal that Subunit Exchange Can Occur Between SaEco30-sGFP and SaEco30 Complexes. (A) BN-PAGE separation of SaEco30 subunit exchange experiment products and controls. SaEco30 and SaEco30-sGFP were incubated both separately and together at 4 °C and 34 °C for four hours. (B) BN-PAGE separation of SaEco30 subunit exchange time course experiment products. SaEco30 and SaEco30-sGFP were incubated at 34 °C and sampled at 30, 60, 120, and 240 minutes (C) Densitometry traces of (B) with each peak labeled to show the expected species.

To test for subunit exchange in all SaEco chimeras, purified SaEco-sGFP constructs and their corresponding SaEco non-fusion constructs were mixed in a tube and incubated at 4 °C, 37 °C, 50 °C, and 60 °C for four hours, and then visualized by BN-PAGE. All SaEco chimeras (except SaEco27) were tested and the results from these experiments showed that most constructs exhibit a temperature dependent subunit exchange [Fig. 2-11] (see Appendix for gel images of all the SaEco subunit exchange experiments). Within the temperature range tested, all constructs except SaEco39 show a ladder of intermediate species. For many constructs, the final distribution of bands is peaked at intermediate values, as anticipated for a randomized distribution of subunits. Given the number of constructs and the limited range of temperatures tested, it was not possible to evaluate the kinetics of subunit exchange, and instead the extent of exchange after four hours incubation at different temperatures was used to estimate the temperature dependence of subunit exchange. The temperature of exchange (T_E) for each construct was estimated as the temperature at which exchange would be ~50% complete after four hours. For example, T_E for SaMscL = 55 °C, EcMscL = 44 °C, SaEco38 = 60 °C, SaEco53 = 50 °C, and SaEco55 = 37 °C [Fig. 2-11] (see Table 2-2 at the end of the chapter for a list of estimates of T_E of all the SaEco chimeras). Given the relatively large intervals in the temperature screen, evaluation of T_E is necessarily imprecise, but the subsequent analyses based on these values were not particularly sensitive to the precise numerical values.

To test for a correlation between T_E and the structural regions of the chimeras, a principal component analysis was performed using a Mathematica script. The chimeras were represented numerically by the parent origin of each structural domain (N-term, TM1,

44

loop, TM2, C-term) where -1 represented SaMscL parent origin and +1 represented EcMscL parent origin. By this convention, SaMscL = (-1,-1,-1,-1,-1), EcMscL = (1,1,1,1,1), SaEco30 = (1,1,1,1,-1), etc. The estimated T_Es for the chimeras were scaled so that the average value equaled 0 and variance equaled 1. SaEco31 and SaEco39 were excluded from the analysis because their $T_E > 60$ °C and therefore could not be estimated and SaEco27 was also excluded due to insufficient protein to study experimentally. The principal component analysis yielded coefficients for the components (0.432, -0.765, -0.175, -0.170, 0.378) associated with the temperature dependence of subunit exchange. The correlation between the experimental values and predicted values of T_E is 0.773. The negative values for TM1, loop, and TM2 mean that SaMscL parent origin is associated with higher T_E, while the positive values for N-term and C-term mean that EcMscL parent origin is associated with a higher T_E. The analysis accurately predicted that the chimera with the highest T_E would be (1,-1,-1,-1,1), which represents the chimera SaEco39. SaEco39 was the only chimera that showed no evidence of subunit exchange even after four hours of heating at 60 °C (and was therefore not included in the principal components analysis). Conversely, the analysis predicts that the chimera with the lowest T_E would be (-1, 1, 1, 1, -1), which is SaEco42, and SaEco42 has a T_E of 37 °C.
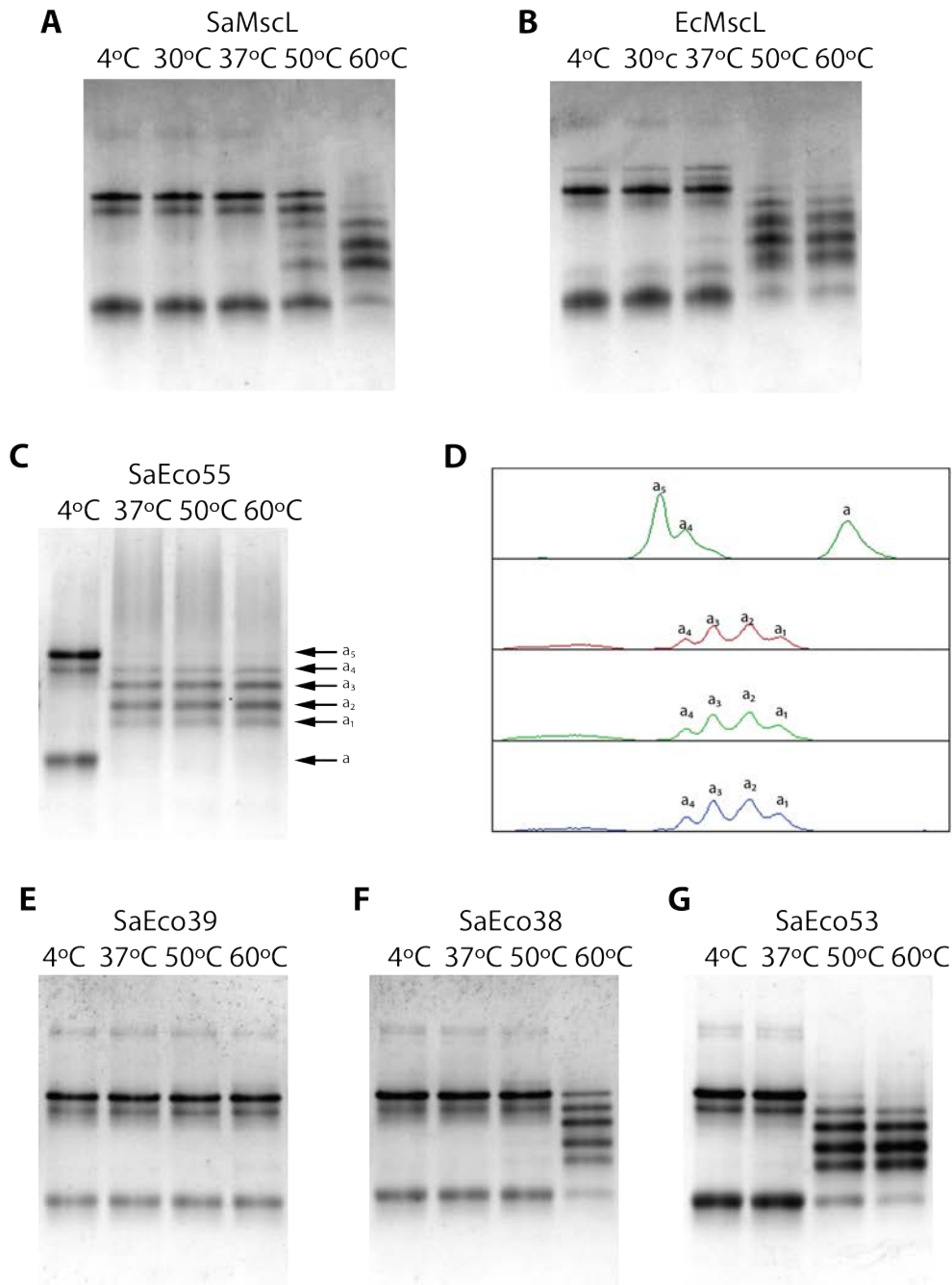
Figure 2-11. Experiments Reveal that Most MscL Constructs Exhibit Temperature Dependent Subunit Exchange. (A, B, C, E, F, G) BN-PAGE separation of subunit exchange experiment intermediates for SaMscL, EcMscL, SaEco55, SaEco39, SaEco38, and SaEco53, respectively. Purified SaEco-sGFP constructs and their corresponding SaEco non-fusion constructs were mixed in a tube and incubated at 4 °C, 37 °C, 50 °C, and 60 °C for four hours. The temperature of exchange (T_E) for each construct was estimated as the temperature at which exchange would be ~50% complete after four hours. T_E for SaMscL = 55 °C, EcMscL = 44 °C, SaEco38 = 60 °C, SaEco53 = 50 °C, and SaEco55 = 37 °C. (D) Densitometry traces of (C) with each peak labeled to show the expected species.

46

Examination of subunit exchange of the poorly expressing chimeras more clearly indicates the presence of a mixture of tetrameric and pentameric complexes [Fig. 2-12]. In particular, gels of SaEco41 and SaEco51 show a distinct, lower molecular weight band that corresponds to a tetrameric non fusion (oligomer "b"), and in the 60 °C lane of SaEco41 the b_1 band is visible [Fig. 2-12A & D].
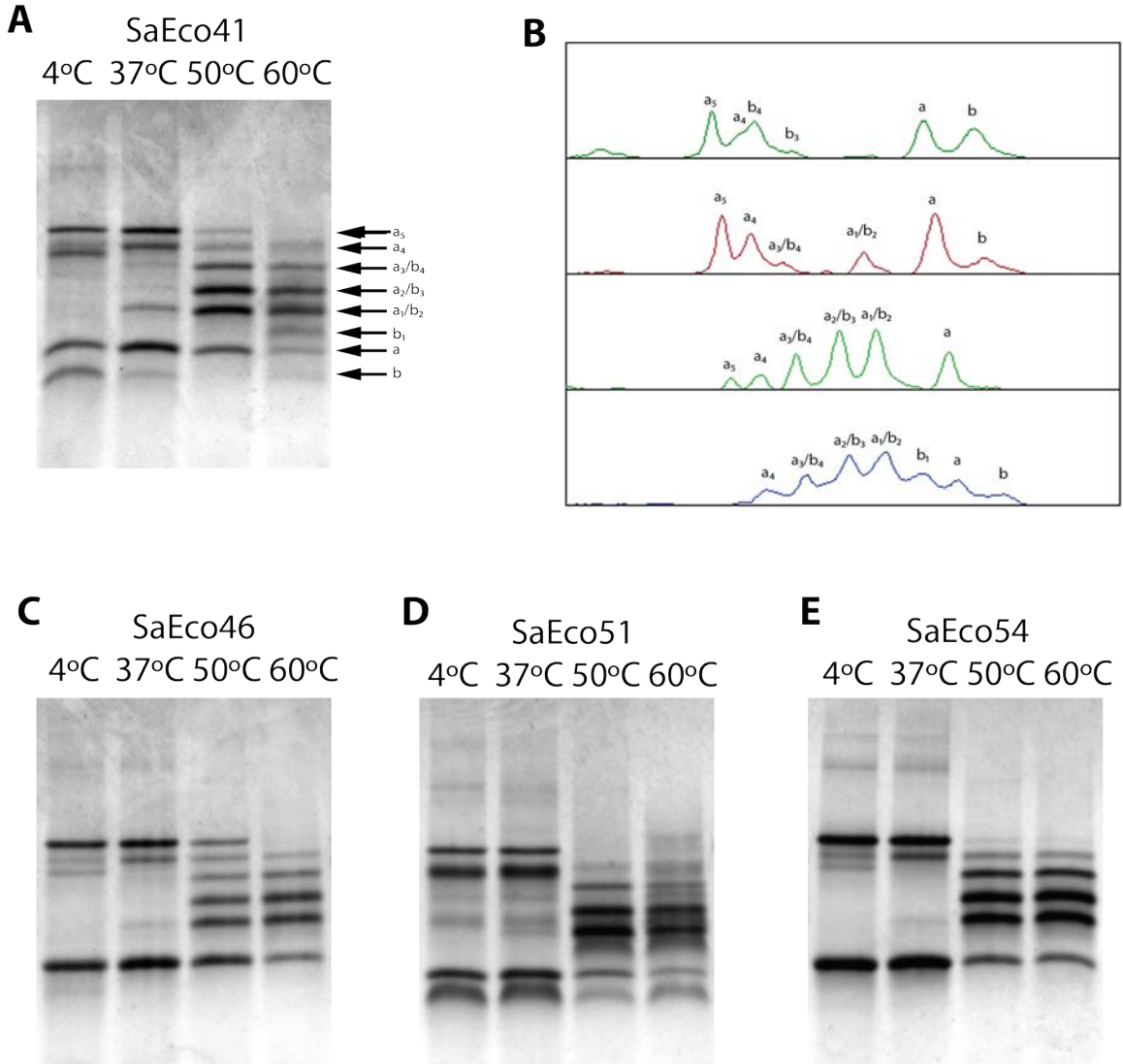
Figure 2-12. Subunit Exchange of the Poorly Expressing Chimeras. (A, C, D, E) BN-PAGE separation of subunit exchange experiment intermediates for SaEco41, SaEco46, SaEco51, and SaEco54, respectively. Purified SaEco-sGFP constructs and their corresponding SaEco non-fusion constructs were mixed in a tube and incubated at 4 °C, 37 °C, 50 °C, and 60 °C for four hours. The temperature of exchange (T_E) for each construct was estimated as the temperature at which exchange would be ~50% complete after four hours. T_E for SaEco41 = 44 °C, SaEco46 = 44 °C, and SaEco51 = 44 °C, and SaEco54 = 44 °C. (B) Densitometry traces of (A) with each peak labeled to show the expected species.

2.7 Crystallography of SaEco Chimeras

Four constructs were chosen as candidates for crystallization based on their homogeneity (single oligomeric state) and high expression levels. SaEco 28, 31, 32, and 52 were expressed in the *BL21Δmscl E. coli* strain, cultured in TB-AMP media, solubilized using DDM, and purified via metal affinity chromatography and size exclusions chromatography. For each construct, the sample was split into two equal parts, one of which had the 6His tag removed by thrombin cleavage and underwent a second size exclusion column. All eight samples were screened for crystallization using the 96-condition commercial screens MemGold™ and MemGold2™ (Molecular Dimensions). All trays were vapor diffusion sitting drops set up by the Crystal Gryphon™ (Art Robbins Instruments) at room temperature and incubated at 4°C for crystal growth. These screens were chosen because they are specifically designed for screening membrane proteins and in the past have produced crystal hits for MscL homologues and their mutants. Also, previous crystallization studies have shown that MscL almost exclusively crystallizes at 4°C.

Four of the samples, SaEco28 + His tag, SaEco28 - His tag, SaEco31 + His tag, and SaEco31 - His tag, gave over 45 initial crystal hits. Some of the more promising hits are shown below in Figure 2-13. Most of the crystal conditions were similar to conditions previously used to crystallize MscL; low molecular weight PEGs (e.g., PEG 400, PEG 550MME, etc), low to neutral pH (4.6 – 7.5), and divalent cation salts (e.g $CaCl_2$, Mg acetate, etc). The constructs crystallized in many crystal forms including needles, rods, plates, and cubic forms. Crystals large enough to harvest were looped, frozen under liquid

nitrogen, and shipped to the Stanford Synchrotron Radiation Lightsource (SSRL) for X-ray diffraction screening at Beamline 12-2. The best crystals diffracted out to ~12 Å resolution in the best direction and had reasonably well defined spots

Several of the conditions that produced the most promising crystal hits were optimized in hopes of improving the crystal quality. Custom crystallization trials were set up with varied precipitant concentration, pHs, buffers, salts, and additives. Several cryo-protectants were also screened to help preserve crystal integrity during harvesting. Unfortunately after extensive effort, crystal quality could not be improved.
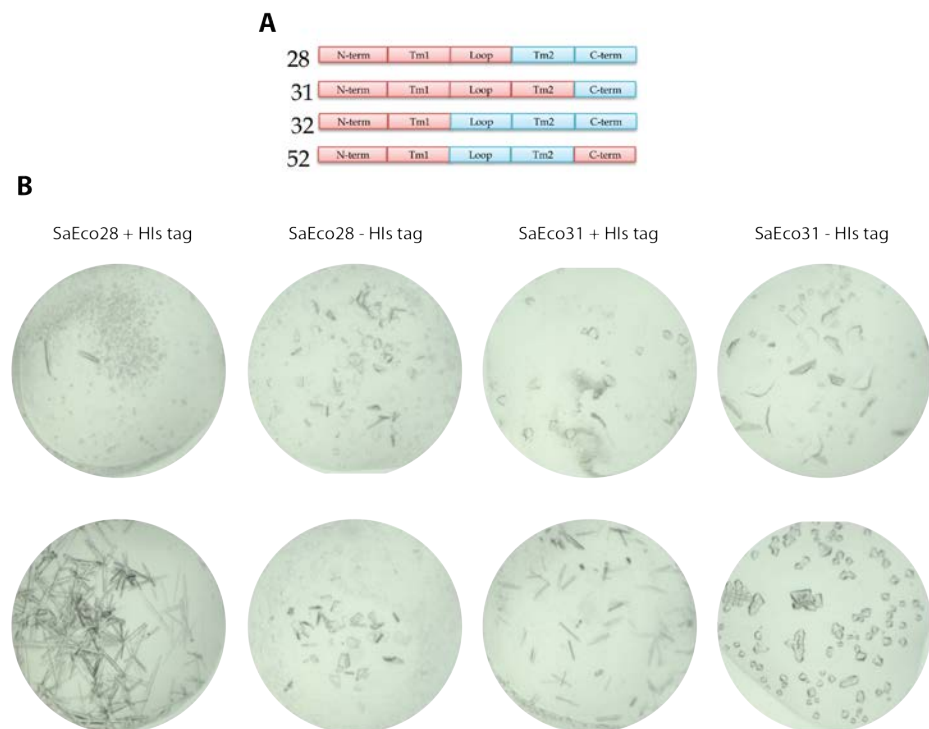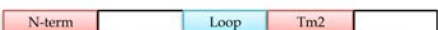
Figure 2-13. SaEco28 and SaEco31 Readily Form Crystals. (A) SaEco 28, 31, 32, and 52 were chosen as candidates for crystallization based on their homogeneity (single oligomeric state) and high expression levels. (B) SaEco28 + His tag, SaEco28 - His tag, SaEco31 + His tag, and SaEco31 - His tag gave over 45 initial crystal hits, but even after optimization the best crystals diffracted out to ~12 Å resolution in the best direction.

2.8 Discussion

In this study, we created a chimera library to systematically survey the correlation between MscL sequence and phenomena like expression level, mixed oligomeric state, heat sensitivity, and subunit exchange. Our results demonstrate that although there is no combination of sequences uniquely associated with each property, there are correlations. Low expression is correlated to EcMscL loop and SaMscL TM2. Pentamer/Hexamer mixed oligomeric state is correlated with the EcMscL N-term and TM2 and SaMscL C-term. Tetramer/pentamer mixed oligomeric state is correlated to EcMscL loop and SaMscL N-term and TM2. High temperature of subunit exchange is linked to EcMscL N-term and C-term and SaMscL TM1, Loop and TM2, with SaMscL TM1 providing the dominant contribution.

Table 2-1. Summary Table Highlighting the Sequence Elements Correlated to Low Expression, Mixed Oligomeric State and High Temperature of Subunit Exchange.

Protein Expression

Our understanding of overexpression of recombinant membrane proteins in hosts like *E. coli* is still very imprecise, but there is an ever-growing literature on the subject [14-17]. There are various stages in the protein expression pathway that if disturbed could lead to poor expression levels. First, it is important to choose an appropriate expression strain, plasmid vector, and promoter for expression. Upon translation, oversaturation of the translocation system can lead to improper membrane protein targeting, insertion, and folding. There have been a few proposed ways to circumvent this issue, including slowing down the rate of transcription using tunable *E. coli* expression strains like Lemo21(DE3) or co-expressing proteins with Sec translocon components [18-20]. Also, at the translocation stage the transmembrane segments of the protein have to be recognized and correctly inserted into the membrane. Sequence determinants of proper insertion include but are not limited to hydrophobicity, segment length, and the position of polar, aromatic, and charged residues [16, 21]. After insertion into the membrane, proper folding and assembly into tertiary and quaternary structures is critical and in some cases is facilitated by chaperones and other proteins. As lipid interactions are essential for proper folding and function of membrane proteins, host lipid composition can affect assembly. It is not uncommon for membrane proteins to misfold in a host with significantly different lipid composition than its native environment [16]. Even after proper assembly of the protein, it is possible that the protein (or large quantities of the protein) is toxic to the cell [17]. The protein might also have a short half-life and be targeted for degradation.

In general, SaEco constructs expressed relatively well, with the exceptions of SaEco-27, SaEco41/-sGFP, SaEco46/-sGFP, SaEco51/-sGFP, and SaEco54/-sGFP, which had significantly reduced expression levels, and SaEco27-sGFP, which could not be detected at all. The common sequence elements between these poorly expressing constructs are the Loop of EcMscL and the TM2 of SaMscL. As detailed above, there are multiple possible causes for the lower expression levels of these constructs and more extensive studies are needed to determine the exact cause(s), but from the current results some inferences can be made. In cultures of the low expressing SaEco chimeras, the presence of a cell pellet of comparable size to their counterparts indicates they do not cause a gain of function phenotype that leads to cell lysis. All constructs were cultured under the same conditions, which implies it is not an issue with the cell growth protocol. Therefore, it can be speculated that the low expression level in these five constructs is due to a sequence level variation that leads to problems in the transcription, protein translation, folding, assembly, or degradation targeting pathways. There have been documented cases of closely related proteins (like SaEco chimeras) with vastly different expression levels [14, 22]. The slightest change in sequence can affect protein expression.

Oligomeric state

Over many years of study, the oligomeric state of MscL has been proposed to vary from monomer to hexamers [23, 24]. From the literature, we can infer that the oligomeric state of MscL is governed by a complex set of factors, including environment (lipid, detergent, etc) and protein sequence (homologues, mutations, etc). Several studies show that oligomeric state is undoubtedly controlled on the sequence level. When translated *in vitro*

by total chemical synthesis, EcMscL and MtMscL were functional when reconstituted into lipid vesicles [25]. Even when the EcMscL protein sequence was expressed in two halves (N-half (M1-D67) and C-half (I68 – S136)), expressed and purified separately, and then reconstituted into lipid vesicles together, they formed functional channels [26]. MscL has the ability to self-assemble into functional oligomers and this intrinsic property is encoded in the amino acid sequence. Here we examine the sequence level determinants of oligomeric state, while keeping the environment constant, using SaEco chimeras.

In the OCAM analysis, SaEco30-sGFP, 37-sGFP, 40-sGFP, and 50-sGFP run as a mixture of pentamers and hexamers. These constructs have in common the N-term and TM2 of EcMscL and the C-Term of SaMscL. SaEco41-sGFP, 46-sGFP, 51-sGFP, and 54-sGFP appear to run as mixture of pentamers and tetramers. The common sequence elements between these constructs are the Loop of EcMscL and the N-term and TM2 of SaMscL. SaEco30-sGFP, 50-sGFP, and 55-sGFP also appear to have lower oligomeric states, which are postulated to be dimers and trimers. These constructs have in common the sequence elements of the TM1 and TM2 of EcMscL and the C-term of SaMscL.

There are several interfaces in the MscL channel where there are complex inter-subunit interactions that could influence oligomeric state. Based on the crystal structure of MtMscL [27, 28], TM1 from each subunit interacts with two TM1s from adjacent subunits. TM1 also interacts with two TM2s, one from the same subunit and another from an adjacent subunit. The N-term of one subunit is inserted between the TM2s of two adjacent subunits. There are also interactions between subunits in the C-terminal helical

coiled coil and between adjacent loops [29, 30]. Mutations that affect any of these interactions could lead to altered oligomeric behavior.

Heat Sensitivity

Our experiments on EcMscL, SaEco30, and SaEco50 show that oligomeric state can be dynamic in solution. Under mild heating conditions, we see that the relative ratios of hexamers to pentamers can shift in these chimeras. Initial gels of EcMscL showed a very minor hexamer band, compared to the original OCAM paper where EcMscL exhibited a more nearly 50:50 distribution of pentamers and hexamers. The purification scheme used in the original OCAM paper is significantly different from the purification scheme used in the studies on SaEco chimeras, however. The original OCAM paper protocol cultured cells in TB media, induced expression by addition of IPTG at $OD_{600} = 2$, lysed the cells using a microfluidizer and ran two columns, a metal affinity column at 4 °C and a SEC column at room temperature. The protocol used to study SaEco chimeras cultured cells overnight in ZYM-5052 auto-induction media, lysed cells by freeze-thaw cycles, and ran only a metal affinity column at 4 °C (no SEC column) followed by dialysis to remove imidazole. Based on the results of heat sensitivity experiments, we believe the most impactful difference between the two protocols was the extended time samples spent at room temperature while on the SEC column.

Temperature effects on MscL stability have been tested using circular dichroism. Temperature melts of EcMscL and MtMscL were monitored at 220 nm at pH 7.5 and showed mostly alpha helical CD spectra, with irreversible thermal unfolding [33].

EcMscL and MtMscL have two transitions, one at 65 °C and another at 85 °C, where there is a significant change in ellipticity. Another CD study of MtMscL shows a similar transition point at 60 °C [30]. The heat-induced shift in oligomeric states we observe at 34°C takes place at a temperature significantly lower than those associated with tertiary structure unfolding for EcMscL and MtMscL (~60 - 65 °C) and closer to physiological temperature.

Subunit Exchange

The results of the subunit exchange experiments show definitive evidence that in detergent solution, subunits of MscL can exchange between complexes. In the literature, there is only one prior study that documented a dynamic change in oligomeric state of MscL in solution [31]. In that study, SaMscL solubilized in detergent LDAO was analyzed by SEC-MALS and yielded an estimated protein mass ~61 kDa (tetramer). The sample was then exchanged into detergent C_8E_5, and when re-analyzed by SEC-MALS yielded a higher estimated protein mass of ~74 kDa (pentamer). This process was reversible, because when exchanged back into LDAO, the protein mass decreased back to ~60 kDa. The authors dismissed this as a detergent artifact, and looked no further into the phenomenon. The results of the subunit exchange experiments detailed above gives us more insight into the mechanism of the movement of subunits between complexes.

Presumably in order for subunit exchange to occur there has to be some combination of channel fusion and fission, which mandates an intermediate species of lower oligomeric state (tetramer, trimer, dimer, or monomer). There is evidence of lower oligomeric states

in some chimeras, but not all. It is therefore particularly interesting that SaEco30, SaEco50, and SaEco55, three chimeras that have what are postulated to be trimer and dimer bands, have low estimated T_E of 37°C. It is possible that lower oligomeric state species exist in all chimeras that show subunit exchange, but these species are so transient or are in such low abundance that they cannot be visualized on the BN-PAGE gels.

Even though subunit exchange in detergent solution has now been proven, it is still unclear whether this is a physiologically relevant phenomenon. Several studies have co-expressed two MscL mutants and seen a population of hetero-oligomers that contain subunits of both mutants. Mika *et al.* co-expressed EcMscL and EcMscL(G22C) from the same plasmid under control of different promoters and found they had a heterogeneous population of channels [32]. Chromatofocusing was used to separate these complexes by pI. Although this experiment, and others like it, indicate that protein subunits translated from multiple mRNAs can form hetero-oligomers, it cannot be determined whether this occurred during channel assembly or through subunit exchange of homogeneous channels after assembly. Further experiments will have to be designed to probe *in vivo* subunit exchange.

Crystallography

Crystallization trials of SaEco 28, 31, 32, and 52 yielded no crystal structures. As it often is with MscL, the barrier for structure determination was not failure to obtain crystals, but rather that the crystals diffracted poorly. It was promising that there were over 45 initial

crystal hits, but several rounds of optimization did not improve diffraction quality. It was interesting to note, however, that SaEco28 (T_E = 60 °C) and SaEco31 (T_E > 60 °C) crystallized while SaEco32 (T_E = 44 °C) and SaEco52 (T_E =37 °C) did not produce crystals. It should also be noted that MscL constructs (including SaEco chimeras) rarely crystallize at 20 °C and almost exclusively crystallize at 4 °C. This begins to show a potential connection between thermostability (in terms of subunit exchange) and crystallizability. It might be that subunit exchange introduces heterogeneity that impedes crystal growth and quality and this is exacerbated at higher crystallization temperatures. Screening MscL constructs for high T_E may be a possible way to select ideal candidates for crystallography.

Table 2-2. List of the Oligomeric State and Temperature of Exchange (T_E) of the SaEco Chimeras

Chimera	Oligomeric State	T_E (°C)	Chimera	Oligomeric State	T_E (°C)
SaMscL	5	55	41	4/5	44
EcMscL	5/6	44	42	5	37
27	n/a	n/a	43	5	37
28	5	60	44	5	44
29	5	37	45	5	44
30	5/6	37	46	4/5	44
31	5	> 60	47	5	44
32	5	44	48	5	50
33	5	44	49	5	55
34	5	55	50	5/6	37
35	5	44	51	4/5	44
36	5	50	52	5	37
37	5/6	44	53	5	50
38	5	60	54	4/5	44
39	5	> 60	55	5	37
40	5/6	55	56	5	55

References

1.	Gandhi, C.S., T.A. Walton, and D.C. Rees, *OCAM: a new tool for studying the oligomeric diversity of MscL channels.* Protein Sci, 2011. **20**(2): p. 313-26.

2.	Trudeau, D.L., M.A. Smith, and F.H. Arnold, *Innovation by homologous recombination.* Curr Opin Chem Biol, 2013. **17**(6): p. 902-9.

3.	Long, S.B., et al., *Atomic structure of a voltage-dependent K+ channel in a lipid membrane-like environment.* Nature, 2007. **450**(7168): p. 376-82.

4.	Carbone, M.N. and F.H. Arnold, *Engineering by homologous recombination: exploring sequence and function within a conserved fold.* Curr Opin Struct Biol, 2007. **17**(4): p. 454-9.

5.	Rico, J.A. and B. Hocker, *Design of chimeric proteins by combination of subdomain-sized fragments.* Methods Enzymol, 2013. **523**: p. 389-405.

6.	Voigt, C.A., et al., *Protein building blocks preserved by recombination.* Nat Struct Biol, 2002. **9**(7): p. 553-8.

7.	Meyer, M.M., et al., *Library analysis of SCHEMA-guided protein recombination.* Protein Sci, 2003. **12**(8): p. 1686-93.

8.	Li, Y., et al., *A diverse family of thermostable cytochrome P450s created by recombination of stabilizing fragments.* Nat Biotechnol, 2007. **25**(9): p. 1051-6.

9.	Heinzelman, P., et al., *Efficient screening of fungal cellobiohydrolase class I enzymes for thermostabilizing sequence blocks by SCHEMA structure-guided recombination.* Protein Eng Des Sel, 2010. **23**(11): p. 871-80.

10.   Clouthier, C.M., et al., *Chimeric beta-lactamases: global conservation of parental function and fast time-scale dynamics with increased slow motions.* PLoS One, 2012. **7**(12): p. e52283.

11.   Jones, D.D., *Recombining low homology, functionally rich regions of bacterial subtilisins by combinatorial fragment exchange.* PLoS One, 2011. **6**(9): p. e24319.

12.   Yang, L.M., D. Zhong, and P. Blount, *Chimeras reveal a single lipid-interface residue that controls MscL channel kinetics as well as mechanosensitivity.* Cell Rep, 2013. **3**(2): p. 520-7.

13.   Pedelacq, J.D., et al., *Engineering and characterization of a superfolder green fluorescent protein.* Nat Biotechnol, 2006. **24**(1): p. 79-88.

14.   Grisshammer, R., *Understanding recombinant expression of membrane proteins.* Curr Opin Biotechnol, 2006. **17**(4): p. 337-40.

15.   Wagner, S., et al., *Rationalizing membrane protein overexpression.* Trends Biotechnol, 2006. **24**(8): p. 364-71.

16.   Bowie, J.U., *Solving the membrane protein folding problem.* Nature, 2005. **438**(7068): p. 581-9.

17.   Luirink, J., et al., *Biogenesis of inner membrane proteins in Escherichia coli.* Biochim Biophys Acta, 2012. **1817**(6): p. 965-76.

18.   Wagner, S., et al., *Tuning Escherichia coli for membrane protein overexpression.* Proc Natl Acad Sci U S A, 2008. **105**(38): p. 14371-6.

19.   Schlegel, S., et al., *Optimizing membrane protein overexpression in the Escherichia coli strain Lemo21(DE3).* J Mol Biol, 2012. **423**(4): p. 648-59.

20. Wagner, S., et al., *Consequences of Membrane Protein Overexpression in Escherichia coli.* Molecular & Cellular Proteomics, 2007. **6**(9): p. 1527-1550.

21. Hessa, T., et al., *Recognition of transmembrane helices by the endoplasmic reticulum translocon.* Nature, 2005. **433**(7024): p. 377-81.

22. Akermoun, M., et al., *Characterization of 16 human G protein-coupled receptors expressed in baculovirus-infected insect cells.* Protein Expr Purif, 2005. **44**(1): p. 65-74.

23. Blount, P., et al., *Membrane topology and multimeric structure of a mechanosensitive channel protein of Escherichia coli.* EMBO J, 1996. **15**(18): p. 4798-805.

24. Hase, C.C., A.C. Le Dain, and B. Martinac, *Molecular dissection of the large mechanosensitive ion channel (MscL) of E. coli: mutants with altered channel gating and pressure sensitivity.* J Membr Biol, 1997. **157**(1): p. 17-25.

25. Clayton, D., et al., *Total chemical synthesis and electrophysiological characterization of mechanosensitive channels from Escherichia coli and Mycobacterium tuberculosis.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(14): p. 4764-4769.

26. Park, K.H., et al., *Purification and functional reconstitution of N- and C-halves of the MscL channel.* Biophys J, 2004. **86**(4): p. 2129-36.

27. Chang, G., et al., *Structure of the MscL homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel.* Science, 1998. **282**(5397): p. 2220-6.

28.   Steinbacher, S., et al., *Structures of the prokaryotic mechanosensitive channels MscL and MscS.* Mechanosensitive Ion Channels, Part A, 2007. **58**: p. 1-24.

29.   Maurer, J.A., et al., *Comparing and contrasting Escherichia coli and Mycobacterium tuberculosis mechanosensitive channels (MscL). New gain of function mutations in the loop region.* J Biol Chem, 2000. **275**(29): p. 22238-44.

30.   Maurer, J.A., et al., *Confirming the Revised C-Terminal Domain of the MscL Crystal Structure.* Biophysical Journal, 2008. **94**(12): p. 4662-4667.

31.   Dorwart, M.R., et al., *S. aureus MscL is a pentamer in vivo but of variable stoichiometries in vitro: implications for detergent-solubilized membrane proteins.* PLoS Biol, 2010. **8**(12): p. e1000555.

32.   Mika, J.T., et al., *On the role of individual subunits in MscL gating: "all for one, one for all?".* FASEB J, 2013. **27**(3): p. 882-92.

33.   Strop, P. (2002). *Characterization of the mechanosensitive channel of large conductance.* (Doctoral dissertation), California Institute of Technology. http://resolver.caltech.edu/CaltechETD:etd-05092002-155216 I

CHAPTER III

LT CHIMERAS

3.1 Design of LT Chimeras

Chimeric mutants of MtMscL, EcMscL, and SaMscL were designed to probe the role of the periplasmic loop and C-terminal domain in function, oligomeric state, and crystallizability. These loop and 'tail" (LT) chimeras are of particular interest because previous studies have highlighted the importance of these regions. The C-terminus has been of significance in all three MscL crystal structures solved to date. A 26 amino acid C-terminal truncation to SaMscL improved the quality of crystals and led to the resolution of its crystal structure (PDB ID: 3HZQ) [1]. It is also hypothesized that this truncation may have caused the switch in oligomeric state to a tetramer, instead of the expected pentamer like MtMscL. In the crystal structure of MtMscL (PDB ID: 2OAR), the last twenty residues of the C-terminus were not resolvable from the diffraction data, presumably because they were disordered in the crystal [2, 3]. A high resolution crystal structure of the isolated C-terminal domain of EcMscL was solved at 1.45 Å resolution (PDB ID: 4LKU) [4]. The MtMscL and SaMscl(CΔ26) crystal structures had crystal contacts in the loop and C-terminus.

Targeted mutations to residues in the C-terminus of MtMscL decreased its thermostability [5], shortening the linker between TM2 and the α-helical coiled coil, decreased the conductance of the channel [6], and deleting the entire C-terminus increased the open dwell time of the channel [1]. The periplasmic loop was also selected

as a substitution site, because it is the least conserved region amongst the homologues. Mutagenesis studies on MtMscL and EcMscL have identified mutations in the loop that cause both gain of function and loss of function phenotypes [7-9]. Two studies on EcMscL, one that expressed the protein in two halves divided at the loop and another that proteolytically cleaved the protein at the loop, showed that the channel was still functional, but had increased mechanosensitivity [10, 11]. These studies have led to the hypothesis that the loop acts as a spring that resists channel openings and assists channel closings.

As with the SaEco chimeras, the structural regions were defined based on secondary structure elements and were designated based on the crystal structures of MtMscL and SaMscL(CΔ26) [Fig 2-1]. In the current study, four LT chimeras were designed by simple "cut and paste" substitutions [Fig 3-1]. The protein sequence of all the chimeras can be found in the Appendix section. LT0 was previously designed by Dr. Zhenfeng Liu, and LT1, LT2, and LT3 are novel designs. The codon optimized genes of LT1, LT2, and LT3 were ordered from Genscript™ and were inserted into pET15b vectors using NdeI and BamHI restriction sites. LT-sGFP fusions were also cloned using homologous recombination. Plasmids were transformed into *BL21Gold E. coli* strain for crystallization screening and *BL21Δmscl E. coli* strain for OCAM and subunit exchange studies.
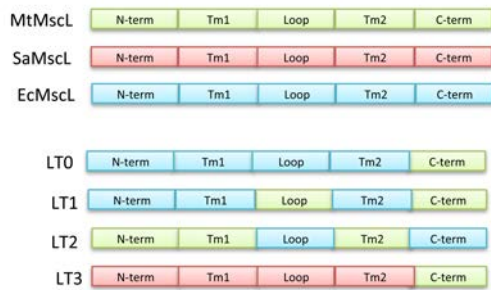
Figure 3-1. Schematic Illustration of the LT Chimera Library. A schematic illustration of the wild-type and chimeric constructs; MtMscL, SaMscL, EcMscL, LT0, LT1, LT2, and LT3. Structural regions are delineated as labeled boxes, green boxes, represent MtMscL parent origin, blue boxes represent EcMscL parent origin, and red boxes represent SaMscL parent origin. The protein sequence of all the chimeras can be found in the Appendix section.

3.2 OCAM

When solubilized in the detergent DDM, OCAM analysis demonstrated that SaMscL and MtMscL are pentamers, while EcMscL is a mixture of pentamers and hexamers [12]. Using the purification and OCAM protocols described in Chapter 2, all LT-sGFP constructs were analyzed and compared to wild-type and SaEco constructs [Table 3-1]. LT0 ran as a hexamer/pentamer mixture, LT1 ran as a tetramer/pentamer mixture, and LT2 and LT3 ran as pentamers [Fig. 3-2].
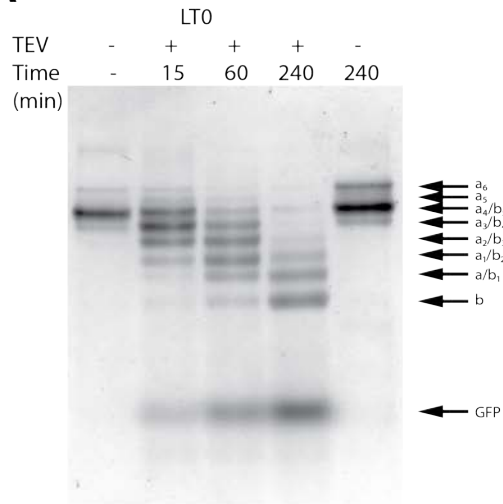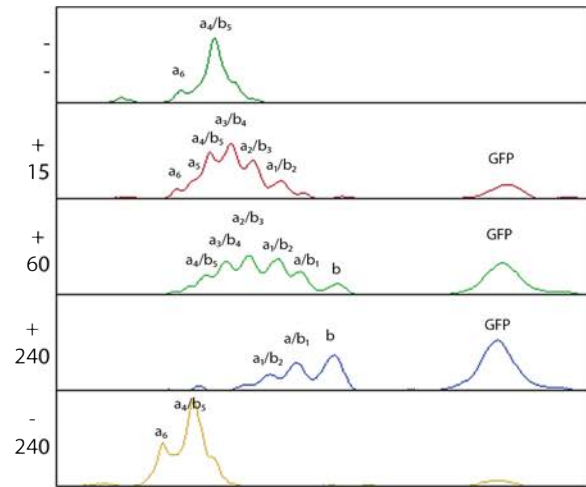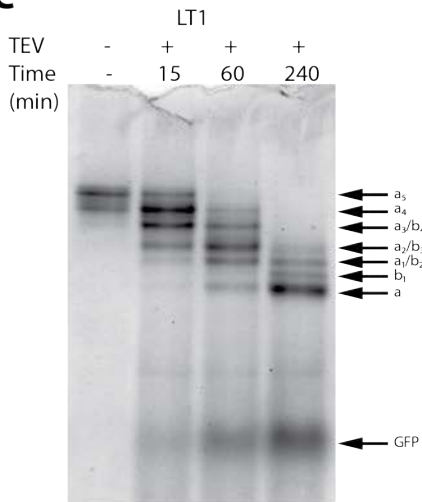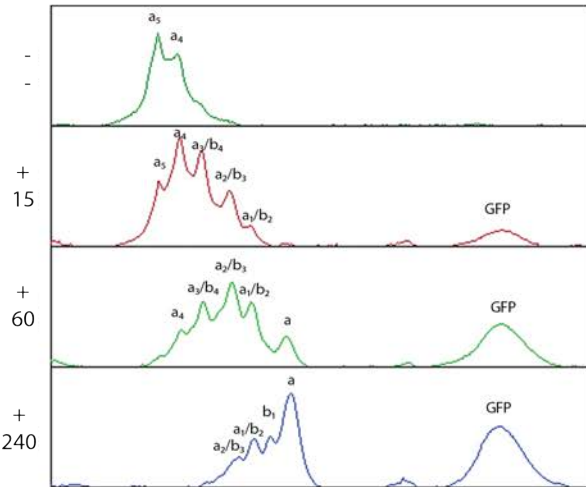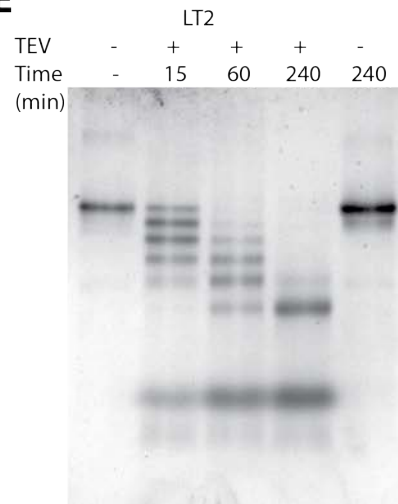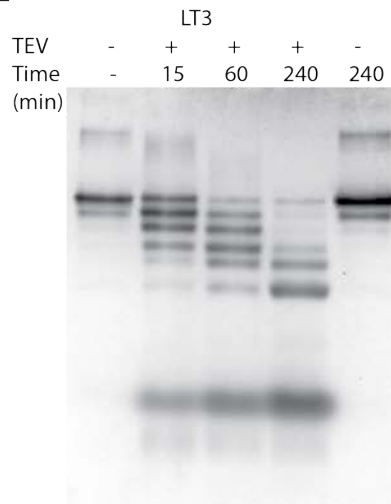
Figure 3-2. OCAM Measurements Reveal that LT Chimeras Have Varied Oligomeric States. (A) BN-PAGE separation of LT0-sGFP OCAM reaction products and controls. TEV protease reactions were quenched at 15, 60, and 240 minute time points. Controls were run without added TEV and sampled at the 0 minutes time point and after 240 minutes incubation at 34°C. A total of eight reaction products are observed in a banding pattern consistent with a mixture of hexameric and pentameric species. (B) Densitometry traces of (A) with each peak labeled to show the expected OCAM reaction product. (C) BN-PAGE separation of LT1-sGFP OCAM reaction products and controls. Reaction products are observed in a banding pattern consistent with a mixture of tetrameric and pentameric species. (D) Densitometry traces of (C) with each peak labeled to show the expected OCAM reaction product. (E - F) BN-PAGE separation of OCAM reaction products and controls of LT2 and LT3, respectively. Reaction products are observed in a banding pattern consistent with pentameric species.

3.3 Subunit Exchange

The LT chimeras were tested for subunit exchange using the protocol described in Chapter 2. All four constructs showed evidence of subunit exchange with temperature of exchange (T_E) for LT0 = 37 °C, LT1 = 37 °C, LT2 = 37 °C, and LT3 > 60 °C [Fig. 3-3]. These values are compared to those of wild-type and SaEco constructs in Table 3-1.
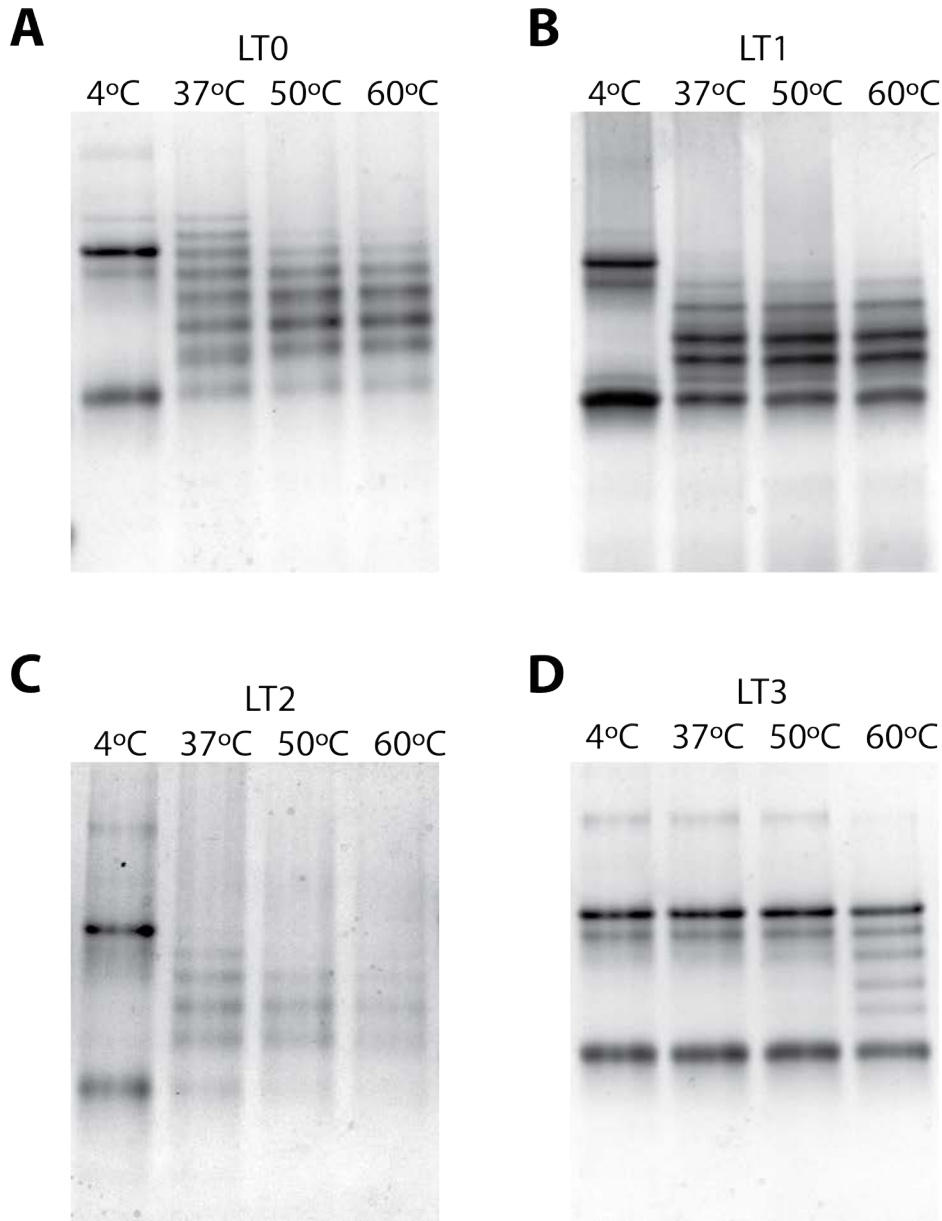
Figure 3-3. Experiments Reveal that LT Chimeras Exhibit Temperature Dependent Subunit Exchange. (A, B, C, D) BN-PAGE separation of subunit exchange experiment intermediates for LT0, LT1, LT2, and LT3, respectively. Purified LT-sGFP constructs and their corresponding LT non-fusion constructs were mixed in a tube and incubated at 4 °C, 37 °C, 50 °C, and 60 °C for four hours. The temperature of exchange (T_E) for each construct was estimated as the temperature at which exchange would be ~50% complete after four hours. T_E for LT0 = 37 °C, LT1 = 37 °C, LT2 = 37 °C, and LT3 > 60 °C.

3.4 Crystallization

LT0, LT1, LT2, and LT3 were expressed in the *BL21Gold E. coli* strain, cultured in TB-AMP media, solubilized using detergent DDM, and purified via metal affinity chromatography and size exclusion chromatography. All samples were screened for crystallization using the 96-condition commercial screens MemGold™, MemGold2™, MemSys™, and MemStart™ (Molecular Dimensions). All trays were vapor diffusion sitting drops set up by hand or using the Crystal Gryphon™ (Art Robbins Instruments) at room temperature and incubated at 4°C for crystal growth. Only two LT chimeras produced crystal hits, LT0 and LT2, and they had less than 10 combined initial crystal hits. Crystal conditions were optimized and crystals were screened for X-ray diffraction at Beamline 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL), but they all diffracted poorly.

Solubilization Detergent Screen

* Names and abbreviations of detergents mentioned in this section: n-Octyl-β-D-Glucopyranoside (OG), n-Decyl-β-D-Maltopyranoside (DM), n-Dodecyl-β-D-Maltopyranoside (DDM), 5-Cyclohexyl-1-pentyl-β-D-maltopyranoside (CYMAL-5), 6-Cyclohexyl-1-hexyl-β-D-maltopyranoside (CYMAL-6), 7-Cyclohexyl-1-heptyl-β-D-maltopyranoside (CYMAL-7), Lauryldimethylamine-oxide (LDAO), n-Dodecylphosphocholine (FOS-CHOLINE-12), n-Tetradecylphosphocholine (FOS-CHOLINE-14), 7-Cyclohexyl-1-heptylphosphocholine (CYCLOFOS-7), and Zwittergent 3-14.

According to the Protein Data Bank of Transmembrane Proteins (PDBTM), as of 08-22-2014 there are 2,272 reported structures of transmembrane proteins (http://pdbtm.enzim.hu) [13]. These proteins are solubilized in a wide variety of detergents and the choice of solubilization detergent can greatly affect crystallizability and crystal quality. There are an overwhelming number of commercially available detergents and selecting targets for a detergent screen can be difficult. During the first round of screening it is common to use detergents that have previously been successfully used to solubilize proteins for X-ray crystal structure determinations. The Membrane Protein Data Bank has easily searchable statistics and can be used to identify the most commonly used detergents (http://www.mpdb.tcd.ie/index.asp). As of the last entry in January 2011, 159 structures of proteins solubilized in OG had been solved, 152 in DDM, 126 in LDAO, 84 in DM, 48 in FOS-CHOLINE-12, etc. All these top targets except OG were included in the LT0 detergent screen. In particular, DDM and LDAO were of interest because MtMscL and SaMscL(CΔ26), respectively, solubilized in those two detergents yielded crystal structures [1-3]. OG was excluded from these trials because previous studies show that MscL aggregates and is not stable in OG [14, 15].

Other considerations to be made during detergent selection are the physical properties of the detergent [16]. These include detergent solubility, the critical micellar concentration (CMC) at which the detergent self-assembles into micelles, and aggregation number, which is the number of detergent monomers per micelle and is related to micelle size by the detergent monomer molecular weight. Detergent molecules are made up of a hydrophilic head group and a hydrophobic alkyl chain. The size and charge of the head

group and the length of the alkyl chain affect the physical properties of the detergent. The length of the alkyl chain affects the aggregation number, CMC, and solubility of the detergent. Short chains have higher CMC and lower aggregation number and longer alkyl chains have poor solubility. Detergent head group properties influence the harshness and solubility of detergents. Small ionic head groups tend to have better solubility, but are harsher (e.g., sulfates). Larger neutral head groups tend to be gentler, but less soluble (e.g., maltopyranosides).

The size and dynamic nature of the detergent belt around the transmembrane protein can impede formation of an ordered crystal lattice. Membrane protein crystals tend to have a high solvent content, due to the accommodation of the detergent belt and the limited exposed protein surface for crystal contact formation. The ideal detergent for a membrane protein is one that is able to extract the protein from the cell membrane and is gentle enough not to perturb the native structure of the protein. A small micelle size is also favorable because more protein surface is exposed that can potentially form crystal contacts, and the amount of space occupied by detergent in the crystal lattice is minimal. However, the micelle size has to be adequately large to protect the hydrophobic region of the protein and prevent aggregation in aqueous solution. A common strategy used to improve crystal quality after an initial crystal hit is the "nearest neighbor" strategy. For example, if a protein forms poorly diffracting crystals in DDM, the protein can then be screened in DM which has a smaller micelle size, but the same head group.

For the first detergent screen, LT0 was extracted in DDM, detergent exchanged into DM, LDAO, CYMAL-7, CYCLOFOS-7, and FOS-CHOLINE-12 and screened for crystallization using MemGold™ screen and incubated at 4 °C. LT0 did not crystallize in any of the other detergents. Figure 3-4 shows a comparison of a crystallization condition (15% w/v PEG 2000, 0.02 M Bis-Tris pH 7), which gave a crystal hit in DDM, but not in the other detergents. LT0 in DDM formed small needles, but in other detergents the drop stayed clear or formed precipitate except CYMAL-7, which seemed to form what looks like amorphous aggregate.
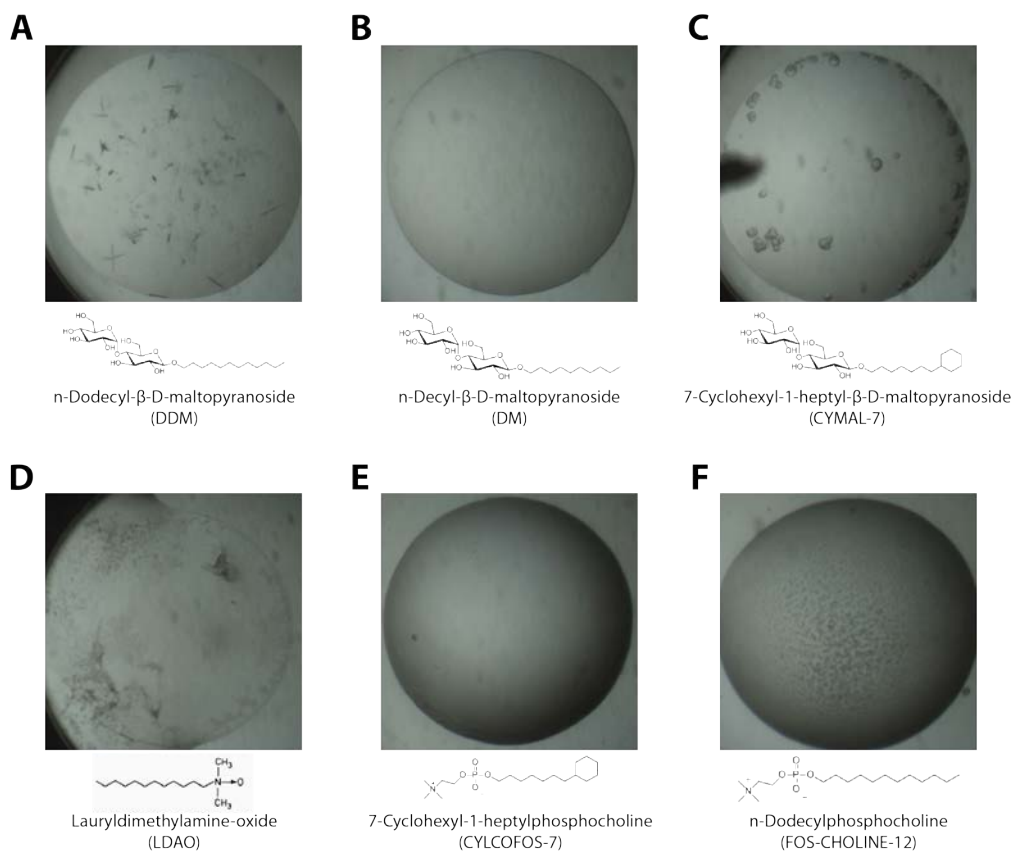
Figure 3-4. Detergent Screen Shows LT0 Only Crystallizes When Solubilized in DDM. Images of crystallizing drops of LT0 solubilized in various detergents in crystallization condition; 15% w/v PEG 2000, 0.02 M bis-tris pH 7.0. (A) LT0 in DDM forms small needles, (B) LT0 in DM, (C) lT0 in CYMAL-7, (D) LT0 in LDAO, (E) LT0 in CYCLOFOS-7, and (F) LT0 in FOS-CHOLINE-12

Extensive effort was made to optimize the LT0 in DDM crystal hit, including screening around the initial crystallization condition and an additive screen, but the crystals could not be optimized past unharvestable needles. Focus was then turned to LT2, which was the only other LT chimera that crystallized. An initial screen of LT2 in DDM gave a few crystal hits, including the thin plates shown in Figure 3-5A. Initial optimizations screened around the initial crystallization condition (22% v/v PEG 400, 0.05 M Na Citrate pH 5.4, 0.07 M NaCl) as well as an additive screen. The additive screen showed improved or altered crystal forms with additive detergents CYMAL-6, CYMAL-5 and Zwittergent 3-14 [Fig. 3-5B, C, and D, respectively]. Therefore CYMAL-5, CYMAL-6, and Zwittergent 3-14 were chosen as targets for a solubilization detergent screen of LT2. CYAML-5 and CYMAL-6 are ideal targets, because they also satisfy the "nearest neighbor" strategy. Like DDM, they both have gentle maltopyranoside head groups, but their hydrophobic chains include a cyclohexyl ring [Fig. 3-5B & C]. Detergents with cyclohexyl rings in their alkyl chains were developed because they offer the benefits of both short and long chain detergents in that they usually have good solubility and a lower aggregation number. For example, compared to DDM, which has an aggregation number of 78-149 and mw = 510.6 Da, CYMAL-5 and CYMAL-6 have aggregation numbers of 47 and 91, respectively, and mw= 494.5 Da and 508.5 Da, respectively. Therefore the CYMAL-5 and CYMAL-6 are predicted to have smaller micelle size and therefore potentially allow for the formation of better ordered crystal lattices.

LT2 was extracted and purified in CYMAL-5, CYMAL-6, Zwittergent 3-14, and FOS-CHOLINE-14, screened for crystallization using MemGold™ screen, and incubated at 4

°C. LT2 crystallized in CYMAL-6, but not in the other detergents [Fig. 3-5E, F & G]. These crystal hits were optimized and screened for X-ray diffraction, but did not yield high quality diffraction data.
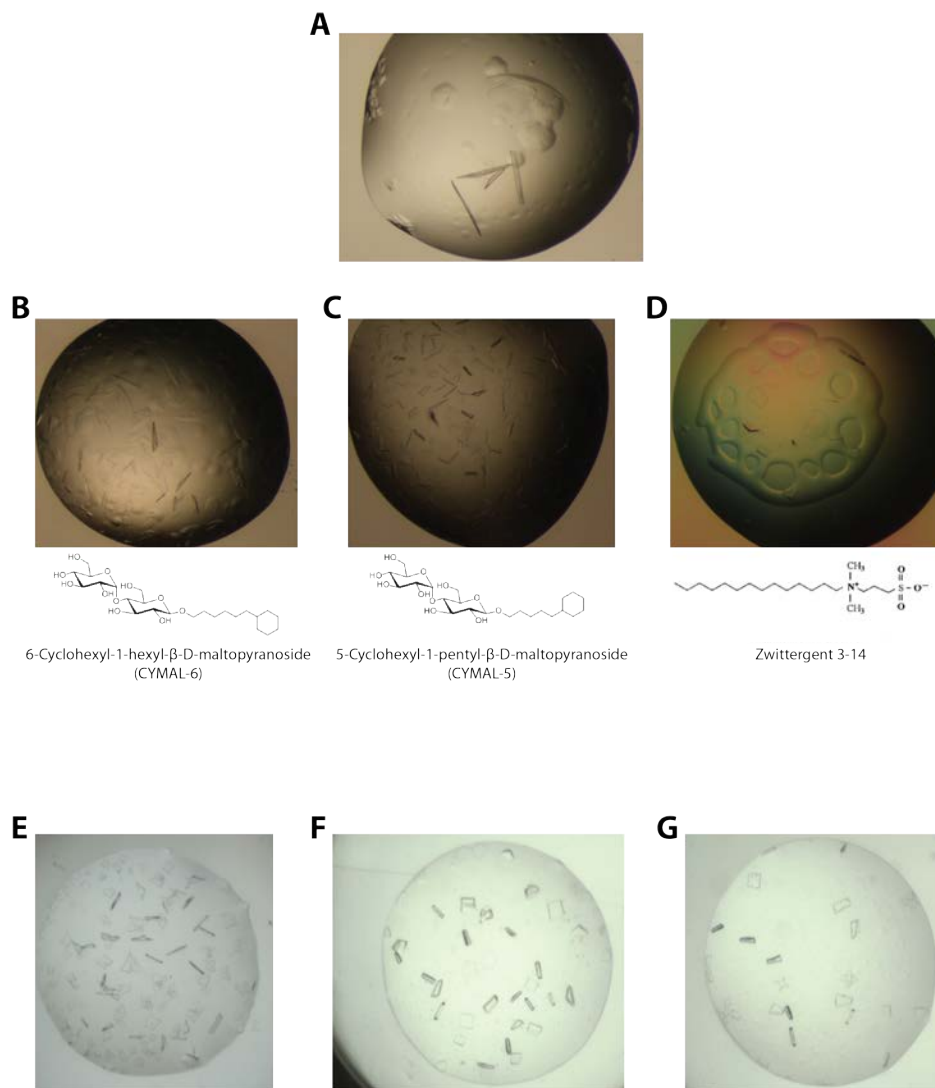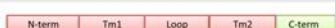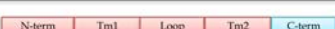
Figure 3-5. Detergent Screen Shows LT2 Crystallizes When Solubilized in Both DDM and CYMAL-6. (A) LT2 solubilized in DDM crystallized in condition 22% v/v PEG 400, 0.05 M sodium citrate pH 5.4, 0.07 M NaCl. (B - D) Crystal condition (A) with additive CYMAL-6, CYMAL-5, and Zwittergent 3-14, respectively/. (E) LT2 solubilized in CYMAL-6 crystallized in condition 26% v/v PEG 400, 0.1 M sodium citrate pH 5.5, 0.05 M NaCl. (F - G) Crystal condition (E) with additive ZnCl and CYMAL-7, respectively.

3.5 Discussion

The LT chimeras were successfully cloned, expressed, purified, and screened for oligomeric state, subunit exchange, and crystallization. The small size of the chimera library limits the information that can be obtained from the results, but in comparison to the SaEco chimeras, they can test previous findings.

Table 3-1. Summary Table of the Oligomeric State and Temperature of Exchange of the LT Chimeras in Comparison to Wild-Type and SaEco Constructs.

	Oligomeric State	T_E (°C)
EcMscL	5/6	44
LT0	5/6	37
30	5/6	37
LT1	4/5	37
50	5/6	37
MtMscL	5*	n/a
LT2	5	37
54	4/5	44
SaMscL	5	55
LT3	5	> 60
31	5	> 60

* Oligomeric State of MtMscL was obtained from Gandhi *et al.* [12]

OCAM

Comparison of the LT Chimeras to wild-type and SaEco constructs gives a fuller understanding of the effects of the structural regions of MscL. Comparing wild-type EcMscL to LT0 and SaEco30 shows that substitutions to the EcMscL C-term do not

affect oligomeric state. To further corroborate that finding, comparing wild-type SaMscL to LT3 and SaEco31 shows that substitutions to the C-term of SaMscL also do not alter oligomeric state. This implies that the C-term of MscL does not contribute to the determination of oligomeric state.

The role of the MscL loop is less clear. Comparing LT1 and SaEco50, it appears that substituting the loop and C-term of MtMscL into the EcMscL sequence alters the oligomeric state; however, the same substitutions of SaMscL do not. Comparing LT2 to SaEco54, substituting the loop and C-term of EcMscL into the MtMscL sequence does not change its oligomeric state, however, the substitution of the same regions into the sequence of SaMscL changes its oligomeric state. The loop of MscL is the least conserved region in the protein and therefore it is plausible that substitutions in that region from various homologues produces varied effects.

Subunit Exchange

The T_E of LT0, LT1, and LT2 are 37 °C and the T_E of LT3 is >60 °C. These values correspond well with the principal component analysis performed in Chapter 2 and compare well with the corresponding SaEco chimeras. LT0 and the corresponding SaEco30, LT1 and the corresponding SaEco50, and LT3 and the corresponding SaEco31, all have identical T_ES. LT2 and SaEco 54 have a T_E of 37 °C and 44 °C, respectively, and the difference in T_E could be because of the TM1 of SaMscL, which is strongly correlated to higher T_E.

Crystallization

In initial screens of the LT chimeras, only LT0 and LT2 gave crystal hits. LT0 was screened with various solubilization detergents, but it only crystallized when solubilized in DDM. LT2 was also screened and crystallized when solubilized in both DDM and CYMAL-6. None of the optimized crystals yielded high quality diffraction data.

References

1. Liu, Z., C.S. Gandhi, and D.C. Rees, *Structure of a tetrameric MscL in an expanded intermediate state.* Nature, 2009. **461**(7260): p. 120-4.

2. Chang, G., et al., *Structure of the MscL homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel.* Science, 1998. **282**(5397): p. 2220-6.

3. Steinbacher, S., et al., *Structures of the prokaryotic mechanosensitive channels MscL and MscS.* Mechanosensitive Ion Channels, Part A, 2007. **58**: p. 1-24.

4. Walton, T.A. and D.C. Rees, *Structure and stability of the C-terminal helical bundle of the E. coli mechanosensitive channel of large conductance.* Protein Sci, 2013. **22**(11): p. 1592-601.

5. Maurer, J.A., et al., *Confirming the Revised C-Terminal Domain of the MscL Crystal Structure.* Biophysical Journal, 2008. **94**(12): p. 4662-4667.

6. Yang, L.M., et al., *Three routes to modulate the pore size of the MscL channel/nanovalve.* ACS Nano, 2012. **6**(2): p. 1134-41.

7.   Maurer, J.A. and D.A. Dougherty, *Generation and Evaluation of a Large Mutational Library from the Escherichia coli Mechanosensitive Channel of Large Conductance, MscL.* Journal of Biological Chemistry, 2003. **278**(23): p. 21076-21082.

8.   Maurer, J.A., et al., *Comparing and contrasting Escherichia coli and Mycobacterium tuberculosis mechanosensitive channels (MscL). New gain of function mutations in the loop region.* J Biol Chem, 2000. **275**(29): p. 22238-44.

9.   Blount, P., et al., *Single residue substitutions that change the gating properties of a mechanosensitive channel in Escherichia coli.* Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11652-7.

10.   Park, K.H., et al., *Purification and functional reconstitution of N- and C-halves of the MscL channel.* Biophys J, 2004. **86**(4): p. 2129-36.

11.   Ajouz, B., et al., *Contributions of the different extramembranous domains of the mechanosensitive ion channel MscL to its response to membrane tension.* J Biol Chem, 2000. **275**(2): p. 1015-22.

12.   Gandhi, C.S., T.A. Walton, and D.C. Rees, *OCAM: a new tool for studying the oligomeric diversity of MscL channels.* Protein Sci, 2011. **20**(2): p. 313-26.

13.   Kozma, D., I. Simon, and G.E. Tusnady, *PDBTM: Protein Data Bank of transmembrane proteins after 8 years.* Nucleic Acids Res, 2013. **41**(Database issue): p. D524-9.

14.   Saint, N., et al., *A hexameric transmembrane pore revealed by two-dimensional crystallization of the large mechanosensitive ion channel (MscL) of Escherichia coli.* J Biol Chem, 1998. **273**(24): p. 14667-70.

15. Sukharev, S.I., M.J. Schroeder, and D.R. McCaslin, *Stoichiometry of the large conductance bacterial mechanosensitive channel of E. coli. A biochemical study.* J Membr Biol, 1999. **171**(3): p. 183-93.

16. Prive, G.G., *Detergents for the stabilization and crystallization of membrane proteins.* Methods, 2007. **41**(4): p. 388-97.

CHAPTER IV

SURFACE ENTROPY REDUCTION STRATEGY

4.1 Surface Entropy Reduction (SER) Strategy

An increasingly popular strategy to improve the crystallizability of proteins is Surface Entropy Reduction (SER) [1]. SER uses rational mutagenesis to alter the surface entropy of proteins. The surface entropy of a protein can affect crystallization since immobilization of the side chains of surface residues involved in lattice contacts is entropically unfavorable and may impede the formation of crystals. Crystal formation is a complex process that is affected by many factors; one of the many is the loss of conformational entropy of hydrophilic long chain residues on the protein surface. Surface residues with high conformational entropies (like lysine and glutamate) can extend into the solvent and have freedom of movement and flexibility; this freedom is restricted when they are involved in or near crystal contacts. The SER method aims to reduce the unfavorable loss of entropy due to side chain immobilization by formation of lattice contacts, thereby improving the probability of crystallization. The SER strategy is postulated to not only increase a protein's propensity to crystallize, but also to increase the probability of obtaining high diffraction quality crystals.

In SER, surface residues with high conformational entropies, like lysine, are mutated to amino acids with lower conformational entropies, like alanine [1]. Longenecker *et al.* conducted a study in which they systematically mutated clusters of lysines into alanines on the surface of truncated constructs of human guanine nucleotide dissociation inhibitor

(RhoGDI). Twelve SER mutants of RhoGDI(NΔ23) and RhoGDI(NΔ66), each with one or more lysines mutated to alanine (e.g., K113A or K98,99,105A), were screened using Crystal Screen 1 & 2 (Hampton Research, 96 conditions total). Neither of the two truncation mutants had previously been successfully crystallized, but the twelve SER mutants generated forty-two crystal hits. Four of those hits gave high diffraction quality crystals. Three of the crystal structures showed that the mutated residues were directly involved in crystal contacts and one structure showed that the single mutated residue was not in close proximity to any crystal contacts. The SER strategy has been updated to also include mutation of other high conformational entropy residues (glutamate and glutamine) to low conformational entropy amino acids (serine, tyrosine, threonine, and histidine) [2]. Again using RhoGDI(NΔ66), Cooper *et al.* made forty SER mutants and got hundreds of crystal hits, with many being of high diffraction quality. Their results showed that tyrosine mutants produced the most hits and although histidine mutants gave relatively fewer hits, their crystals were more likely to generate high resolution structures. Information drawn from early studies on SER has been used to create an SER prediction server (http://nihserver.mbi.ucla.edu/SER/) to identify possible sites for mutation. The primary amino acid sequence of a protein is input to the server and processed to output suggested mutation clusters, each with an SER score. To identify residues of interest, the primary sequence is run through a preliminary analysis, which involves secondary structure prediction, entropy profiling, and a PSI-BLAST search. The secondary structure is predicted using PSIPRED, and residues are scored based on their location on the protein. An entropy profile of the protein is determined using values from Sternberg's side chain entropy tables and averaged three residues at a time to identify sequence

regions of high side chain entropy. A PSI-BLAST search is subsequently performed to avoid mutations in highly conserved regions and highlight potential mutation sites that are aligned with low conformational entropy regions in homologous proteins. This preliminary analysis gives a list of residue clusters that are then scored based on several factors, including the length of the low entropy patch post mutation, the minimum number of mutations needed to effect the greatest entropy reduction, etc.

While the SER approach was motivated by side chain entropy considerations, it is possible, and even likely, that mutations of residues like lysine to alanine have consequences other than entropy reduction, and these other effects contribute to improved crystallization properties. For example, SER mutations result in a change in the electrostatic charge of the protein or can affect solubility, or result in the removal of side chains that alters steric effects. It can also be argued that the surface entropy reduction caused by the mutation of so few residues does not produce a large enough improvement in free energy to stimulate crystallization. The thermodynamics of crystallization is complex and not precisely understood, so it is difficult to absolutely prove or disprove the validity of SER. However the increasing repertoire of structures solved using the SER strategy demonstrates that, whatever the underlying mechanism(s), it represents an effective way to rationally design protein mutants for crystallization purposes. Prior to SER, truncations of large disordered regions, tryptic digests, and random mutagenesis were the main options for engineering proteins that crystallize better than wild-type. Truncations and tryptic fragments are major modifications that may significantly alter protein structure, and the possibilities for random mutagenesis are endless. SER has

opened up a new avenue for crystallographers to more rationally narrow down the target

sites for mutagenesis.

SER has helped to determine structures of many proteins that are recalcitrant to

crystallization, including the human proteins CUE:Ubiquitin complex, *Yersinia pestis*

LcrV Antigen [3], and choline acetyltransferase [4]. However, these are all soluble

proteins and the efficacy of the SER strategy in α-helical membrane protein

crystallography has not been well documented.

4.2 Design of SaMscL SER Mutants

The crystal structure of SaMscL(CΔ26) truncation mutant has been solved at 3.8 Å

resolution as a tetrameric channel in an expanded non-conducting intermediate state [5].

The crystal structure of wild-type MtMscL has been solved at 3.5 Å resolution as a

pentameric channel in a "closed" non-conducting state [6, 7]. While these structures have

given insight into the gating behavior and function of MscL, the use of the CΔ26

truncation of SaMscL has complicated the comparison to MtMscL since it is not clear

how deletion of the C-terminal domain has influenced the structure. Obtaining a high

resolution structure of a full length SaMscL SER mutant that is closer to wild-type

(differing by only one to three mutations) could help address important questions

concerning differences in their observed gating, function, and oligomeric state, such as:

- Will the full length SaMscL SER mutants have a conformation similar to the expanded non-conducting intermediate state of SaMscL(CΔ26) or a "closed" state similar to that of MtMscL?

- Will it be a tetramer like SaMscL(CΔ26), a pentamer like MtMscL, or have a unique oligomeric state?

The amino acid sequence of wild-type SaMscL was submitted to the SER prediction server (http://services.mbi.ucla.edu/SER/) and the output of this analysis identified three mutation sites: the charge cluster composed of the three residues 92-94 (KKE92AAA), residues 54-55 (KE54AA), and residue 62 (K62A), with SER scores of 7.88, 3.83, and 1.11, respectively. A higher SER score indicates a higher predicted reduction in surface entropy. Figure 4-1 shows the location of the three suggested SER mutation sites highlighted on a single subunit of the SaMscL(CΔ26) crystal structure. K62 and KE54 are located on the periplasmic loop and KKE92 is located on the C-terminal region, interestingly right before the truncation point.
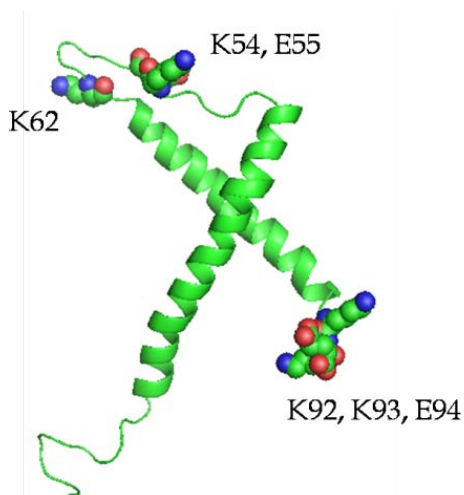
Figure 4-1. Crystal Structure of a Single Subunit of SaMscL(CΔ26) with the SER Mutation Sites Highlighted. Mutation sites are highlighted by depicting the molecules of the side chains as space filling spheres. K62 and KE54 are located on the periplasmic loop and KKE92 is located on the C-terminal region.

Cloning, Protein Expression and Purification

SaMscL SER mutations were cloned by site-directed mutagenesis. A standard quick-change PCR protocol was used to incorporate the mutations into the wild type gene. All SaMscL SER mutants have an N-terminal 6His-tag with a thrombin recognition sequence between the tag and target protein. All genes were incorporated into a pET15b plasmid and were transformed into and expressed in the *E. coli* BL21Gold(DE3) strain. Cells were cultured in 1% glucose TB-AMP media and expression was induced with IPTG. Proteins were extracted and solubilized in dodecyl-β-D-maltoside (DDM) detergent and lysed using either a sonicator or a microfluidizer. They were subsequently purified by both nickel affinity and size exclusion chromoatography. The 6His tag was then removed by thrombin digestion followed by a final purification step on a size exclusion column to remove thrombin and cleavage fragments. Purified samples were then concentrated to 12-18 mg/ml for subsequent crystallization screening.

4.3 Crystallization

Crystal trays were set up either by hand or using the automated Mosquito™ (Rigaku/TTP LabTech). Trays set up by hand were vapor diffusion hanging drops with a protein to reservoir solution ratio of 1 ul: 1 ul, set up at room temperature and stored at 20 °C or 4 °C. Trays set up by the Mosquito™ were vapor diffusion sitting drops with a protein to reservoir solution ratio of 0.2 ul: 0.2 ul, set up at room temperature and stored at 20 °C or 4 °C. All constructs were screened at a minimum with the commercial screens MemGold 1 & 2 (Molecular Dimensions, MD), Crystal 1 & 2 (Hampton Research, HR), and Index 1 & 2 (HR). These screens were chosen because they cover a comprehensive range of precipitants, salts, pHs, and additives. The MemGold screens were of particular interest, because they are specifically designed for screening membrane proteins and in the past have produced crystal hits for MscL homologues and their mutants. Other commercial screens used were MemSys (MD), MemStart (MD), MemPlus (MD), the PEGs Suites 1 & 2 (Qiagen), Peg/Ion (HR), MembFac (HR), Wizard 1&2 (Emerald BioSystems), and JCSG+ (MD).

Four of the SaMscL SER constructs produced crystal hits, with SaMscL(KKE92AAA) being the most successful (ten hits), while SaMscL(K62H) yielded three hits and SaMscL(KE54HH) and SaMscL(KKE92HHD) produced one hit each. Table 4-1 lists all the SaMscL SER mutants screened, the commercial crystallization screens used, and the number of initial crystal hits obtained. Figure 4-2 shows a few sample images of these crystal hits and their conditions. The SaMscL SER mutants crystallized in a variety of forms including, needles, rods, and plates. MemGold was the most successful screen.

Wizard and Crystal also gave multiple hits. The trends among successful crystallization conditions were low molecular weight polyethylene glycols (PEGs 400-1000), sodium and magnesium salts, and pH 5.4 - 7.0 buffer. Crystallization conditions that produced crystal hits were optimized until they generated harvestable crystals. These crystals were looped, frozen under liquid nitrogen, and shipped to Stanford Synchrotron Radiation Lightsource (SSRL) for X-ray diffraction screening at Beamline 12-2.

Table 4-1. Table of the SER Mutants Screened For Crystallization. A total of ten SaMscL SER mutants were cloned, expressed, purified, and screened for crystallization conditions. SaMscL(KKE92AAA) produced ten hits, SaMscL(K62H) yielded three hits and SaMscL(KE54HH) and SaMscL(KKE92HHD) produced one hit each.

SaMscL SER mutant	Commercial Screens Tested	Crystal Hits
K62A	MemGold 1&2, MemPlus, MemSys, Crystal 1&2, Index 1&2, JCSG+, Wizard 1&2, PEG Ion, MembFac	None
KE54AA	MemGold 1&2, MemPlus, MemSys, Crystal 1&2, Index 1&2, JCSG+, Wizard 1&2, PEG Ion, MembFac	None
KKE92AAA	MemGold 1&2, MemPlus, MemSys, Crystal 1&2, Index 1&2, JCSG+, Wizard 1&2, PEG Suites 1&2, PEG Ion, MembFac	10
K62Y	MemGold 1&2, MemPlus, MemSys, Crystal 1&2, Index 1&2, JCSG+, Wizard 1&2, PEG Suites 1&2, PEG Ion, MembFac	None
KE54YY	MemGold 1&2, MemPlus, MemSys, Crystal 1&2, Index 1&2, JCSG+, Wizard 1&2, PEG Suites 1&2, PEG Ion, MembFac	None
KKE92YYY	MemGold 1&2, MemPlus, MemSys, Crystal 1&2, Index 1&2, JCSG+, Wizard 1&2, PEG Suites 1&2, PEG Ion, MembFac	None
K62H	MemGold 1&2, MemPlus, MemSys, Crystal 1&2, Index 1&2, JCSG+, Wizard 1&2, PEG Suites 1&2, PEG Ion, MembFac	3
KE54HH	MemGold 1&2, MemPlus, MemSys, Crystal 1&2, Index 1&2, JCSG+, Wizard 1&2, PEG Suites 1&2, PEG Ion, MembFac	1
KKE92HHD	MemGold 1&2, MemPlus, MemSys, Crystal 1&2, Index 1&2, JCSG+, Wizard 1&2, PEG Suites 1&2, PEG Ion, MembFac	1
KE54LL	MemGold 1&2, MemPlus, MemSys, Crystal 1&2, Index 1&2, JCSG+, Wizard 1&2, PEG Suites 1&2, PEG Ion, MembFac	None
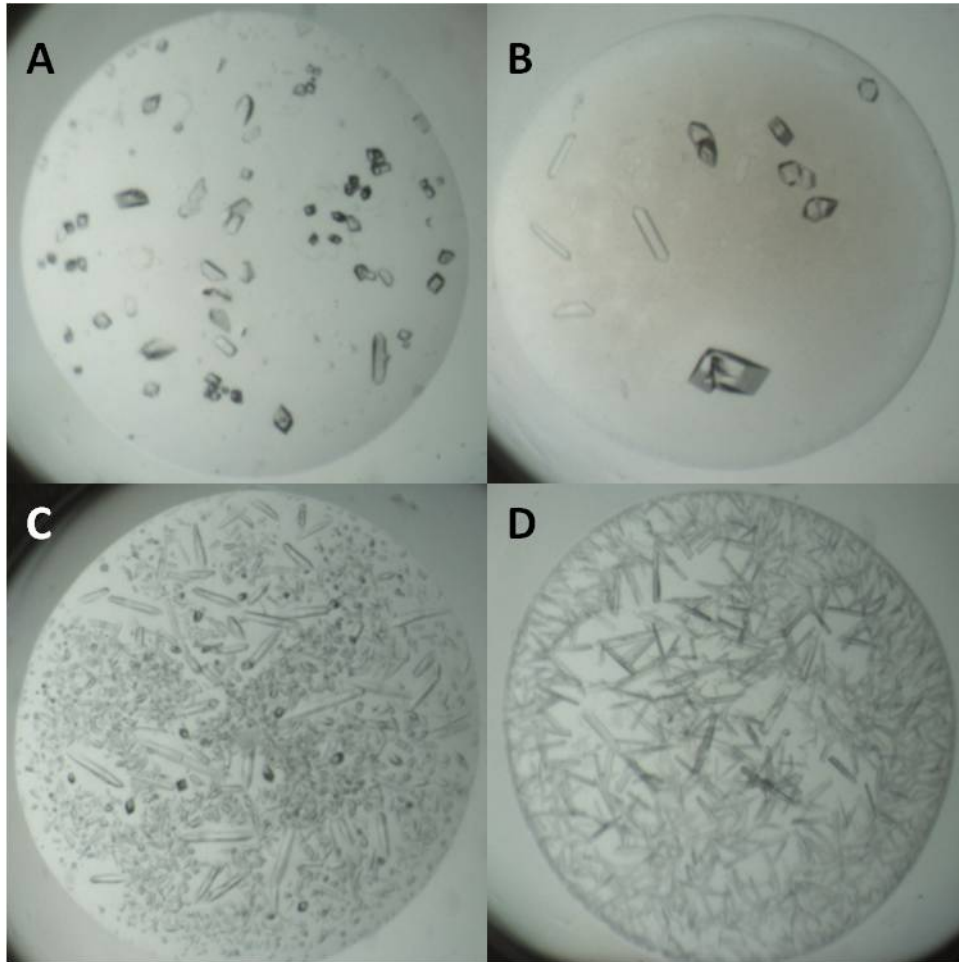
Figure 4-2. SaMscL SER Mutants Produced Crystal Hits. (A) SaMscL(K62H), condition 0.1 M sodium chloride, 0.1 M bicine pH 9.0, 20% w/v polyethylene glycol monomethyl ether 550. (B) SaMscL(KKE92AAA), condition 0.05 M sodium chloride, 0.02 M magnesium chloride, 0.1 M sodium citrate pH 6.0, 22% v/v polyethylene glycol 400. (C) SaMscL(KKE92AAA), condition 0.1 M sodium chloride, 0.1 M bicine pH 9.0, 20% w/v polyethylene glycol monomethyl ether 550. (D) SaMscL(KE54HH) 0.05 M magnesium chloride, 22% v/v polyethylene glycol 400, 0.1 M glycine pH 9.0.

Optimization of Crystal Conditions

SaMscL(K62H) crystallized in a solution of 0.05 M magnesium acetate, 24% v/v polyethylene glycol 400, and 0.05 M sodium acetate pH 5.4. This initial condition generated thin needle clusters [Fig 4-3A]. The first round of optimization was a 2-D screen of PEG 400 concentrations and buffer pH. A new condition of 24% PEG 400, 0.05 M sodium acetate pH 4.6, 0.05 M magnesium acetate, gave larger needles, but they were still clustered. The new optimized condition was then screened with commercial additive and detergent screens (Hampton Research). The addition of several additives, in particular strontium chloride, calcium chloride, hexanediol, and hexamine cobalt (III) chloride, produced needles and plates that were harvestable [Fig 4-3B, C & D]. Crystals from these drops were harvested and sent for X-ray diffraction screening.

Several harvested crystals were screened and they all diffracted poorly and could not be used for structural analysis. The best crystals diffracted out to ~11 Å resolution in the best direction and had well defined spots. SaMscL(KKE92AAA) SaMscL(K62H), SaMscL(KE954HH) and SaMscL(KKE92HHD) were exhaustively screened for optimized crystallization conditions as well as cryoprotectants, but there was no improvement in diffraction quality.
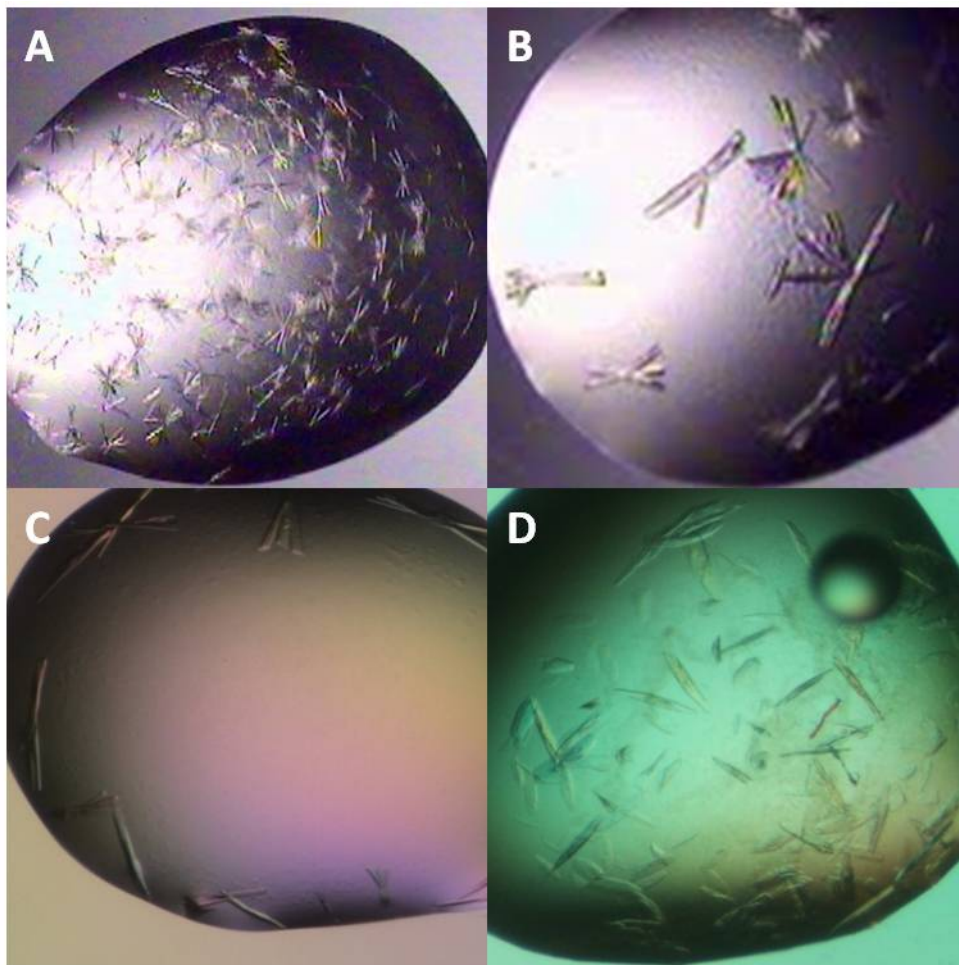
Figure 4-3. Optimization of SaMscL(K62H) Crystals. (A) SaMscL(K62H) 24 % v/v polyethylene glycol 400, 0.05 M magnesium acetate, and 0.05 M sodium acetate pH 5.4. (B) SaMscL(K62H) 24 % v/v polyethylene glycol 400, 0.05 M magnesium acetate, 0.05 M sodium acetate pH 4.6, and 0.15 M calcium chloride. (C) SaMscL(K62H) 24 % v/v polyethylene glycol 400, 0.05 M magnesium acetate, 0.05 M sodium acetate pH 4.6, and 0.02 M hexamine cobalt (III) chloride. (D) SaMscL(K62H) 24 % v/v Polyethylene Glycol 400, 0.05 M magnesium acetate, 0.05 M sodium acetate pH 4.6, and 6% v/v hexanediol.

4.4 Discussion

The SER strategy was used to design mutants of SaMscL, and these constructs were successfully cloned, expressed, purified, and screened for crystallization conditions. SaMscL(KKE92AAA) SaMscL(K62H), SaMscL(KE954HH) and SaMscL(KKE92HHD) gave over ten initial crystal hits, which is more than had been recorded for wild-type

SaMscL. Crystallization conditions were optimized extensively, but failed to yield high quality diffraction patterns for structural analysis. The best crystals only diffracted out to ~ 11 Å resolution in the strongest direction. While the SER strategy appears to have improved the crystallizability of SaMscL, unfortunately the diffraction qualities of these crystals were not significantly improved.

References

1.   Longenecker, K.L., et al., *Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of RhoGDI.* Acta Crystallogr D Biol Crystallogr, 2001. **57**(Pt 5): p. 679-88.

2.   Cooper, D.R., et al., *Protein crystallization by surface entropy reduction: optimization of the SER strategy.* Acta Crystallogr D Biol Crystallogr, 2007. **63**(Pt 5): p. 636-45.

3.   Derewenda, Z.S., *Rational protein crystallization by mutational surface engineering.* Structure, 2004. **12**(4): p. 529-535.

4.   Kim, A.R., et al., *Surface-entropy reduction used in the crystallization of human choline acetyltransferase.* Acta Crystallographica Section D-Biological Crystallography, 2005. **61**: p. 1306-1310.

5.   Liu, Z.F., C.S. Gandhi, and D.C. Rees, *Structure of a tetrameric MscL in an expanded intermediate state.* Nature, 2009. **461**(7260): p. 120-U132.

6.   Chang, G., et al., *Structure of the MscL homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel.* Science, 1998. **282**(5397): p. 2220-6.

7.   Steinbacher, S., et al., *Structures of the prokaryotic mechanosensitive channels MscL and MscS.* Mechanosensitive Ion Channels, Part A, 2007. **58**: p. 1-24.

CHAPTER V

FUTURE DIRECTIONS

5.1 Functional Characterization of SaEco Chimeras

The most critical next step is the functional characterization of the SaEco chimeras to elucidate the roles of the structural elements in channel physiology. There are several assays that can be used to characterize MscL function of which the most widespread are cell survival under downshock conditions and electrophysiology. Osmotic down-shock assays can be used to test cell survivability after exposure to hypo-osmotic shock conditions [1-4]. Electrophysiology experiments can characterize single channel traces to evaluate changes in the pressure threshold of gating, channel open dwell times, and conductance. From the results of these studies, chimeras may be found to behave similarly to wild-type, or they may exhibit altered channel function which can be assigned to two basic phenotypes: gain of function (GOF) and loss of function (LOF) phenotypes. A GOF phenotype can be caused by channels that have increased sensitivity and therefore gate at lower tensions, including the extreme case of channels that can spontaneously open in the absence of applied tension. A LOF phenotype can result from channels that either have decreased sensitivity and gate at higher tensions, including channels that no longer can open.

Cell Survival Assay

There are several cell survival assays that have been used to test the activity of MscL variants and their ability to rescue cells from osmotic down-shock. Preliminary studies

were conducted using a simple osmotic down-shock assay outlined by Liu *et al.* to assess the function of MscL [4]. First, plasmids containing MscL variant genes were transformed into the *mscl/mscs* knock out *E. coli* strain MJF465(DE3) and cells were cultured in LB media overnight. Pre-cultures were used to seed a fresh culture of LB media and were grown to $OD_{600} = 0.3$. Cultures were mixed at a 1:1 ratio with LB media supplemented with NaCl (to final concentration of 1 M) and IPTG. Samples were incubated for 1 h after which they were normalized to $OD_{600} = 0.1$. These normalized samples were then diluted 1:500 into isotonic LB media (mock-shock) or pure distilled water (down-shock). Samples were incubated for 20 minutes and then aliquots were spread on regular LB agar plates. After overnight incubation, colonies on each plate can be counted. A comparison of the number of colonies on the mock-shock plates to those on down-shock plates indicates whether the MscL variant was able to rescue cells from osmotic down-shock.

Preliminary trials were run using this protocol to test both negative and positive controls [Fig. 5-1]. Negative controls were carried out using the MJF465(DE3) cells, cells transformed with an empty vector, and cells expressing the soluble protein MLP-Tv. The positive control was done on cells expressing wild-type SaMscL. The results show that only cells expressing SaMscL were able to survive osmotic down-shock, evidenced by the large number of colonies on the down-shock plates. The results of this assay were not reproducible, however, and we were not able to obtain consistent results from this deceptively simple analysis.

94
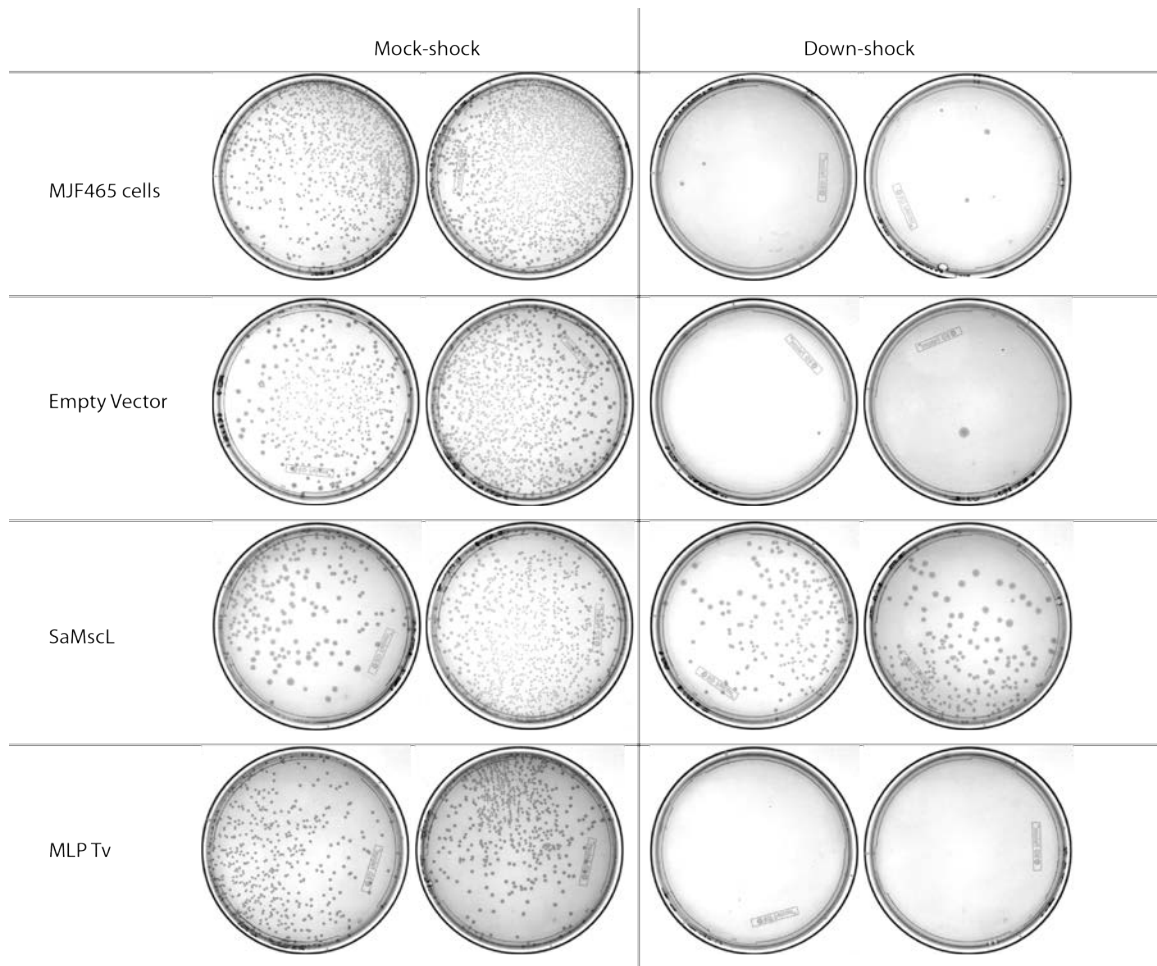
Figure 5-1. Preliminary Results of a Cell Survivability Osmotic Down-Shock Assay. Negative controls were carried out using the MJF465(DE3) cells, cells transformed with an empty vector, and cells expressing the soluble protein MLP-Tv. The positive control was done on cells expressing wild-type SaMscL. The results show that only cells expressing SaMscL were able to survive osmotic down-shock, evidenced by the large number of colonies on the down-shock plates.

Maurer *et al.* developed a high-throughput assay that modifies the Molecular Probes' Live/Dead® *Bac*Light™ bacterial viability assay to characterize GOF and LOF MscL constructs [2]. In this assay, cells are grown in the presence of two DNA binding fluorescent dyes, the membrane-impermeable red fluorescent dye propidium iodide and the membrane permeable green fluorescent dye SYTO 9. The fluorescence of both SYTO 9 and propidium iodide increases when bound to DNA. Monitoring the green to red fluorescence signal ratio is an indicator of the ratio of live to dead cells. Cell cultures

with live cells will have a high green to red fluorescence ratio, since while SYTO 9 can penetrate the membrane and bind to intracellular DNA, propidium iodide cannot. Cultures with a high population of dead cells will have a lower green to red ratio, because once the cell membrane is compromised, propidium iodide can now interact with DNA, with the consequent increase in red fluorescence.

The experiments are performed with the MJF465(DE3) strain and the MscL variants are introduced via an inducible plasmid. First, cells are grown in minimal media for 14 h until they reach steady state. These cultures are used to inoculate a fresh culture and are grown in the presence of IPTG for 7.5 h. Samples are diluted 20-fold in solutions of varying osmotic strength containing propidium iodide and SYTO 9 and incubated for 75 minutes. Fluorescence signals of these samples are read at excitation wavelengths of 480 and 490 nm and emission wavelengths of 500 and 635 nm. There are three possible outcomes: for GOF MscL variants, the green to red fluorescent levels will be low at all osmotic strengths, because the cells died upon induction of protein expression; for wild type activity MscL variants, the green to red fluorescence ratio should remain high at most osmotic strengths except at very high down-shocks; LOF MscL variants should have a high green to red fluorescence ratio at low down-shock shocks, but the fluorescence ratio should drop at intermediate down-shocks. This assay was successfully used by Maurer and Dougherty (2001), and is a promising approach for large scale screening of chimeras.

A second fluorescence assay has also been described, which uses the poorly membrane permeable DNA binding dye ethidium bromide to measure release of DNA upon cell death [3]. MJF465(DE3) cells carrying plasmids with the desired MscL variant gene are cultured for 16 h in LB media supplemented with NaCl (to a final concentration of 500 mM) (LB-500) in the presence of IPTG. These pre-cultures are used to seed fresh cultures of the same media and these cultures grown to an $OD_{600} = 0.5$. Cells are pelleted by centrifugation at a low spin and resuspended in 500 mM NaCl solution. Samples are diluted 50-fold into six shock solutions containing ethidium bromide and ranging from 500 mM NaCl to 0 M NaCl. Samples are incubated for 45 minutes in the dark at room temperature and centrifuged at low spin to pellet the cells. The fluorescence of the supernatant can then be measure at 254 nm excitation wavelength and 632 nm emission wavelength. After correction for background, the fluorescence intensity is proportional to the amount of DNA in solution due to DNA release after cell death. Cells expressing MscL variants with wild type activity will have a low fluorescence signal at all osmotic strengths. Cells expressing LOF MscL variants are expected to have a low fluorescence signal at moderate osmotic strengths and show an increase in fluorescence intensity at low osmotic strengths. Cells expressing non-lethal GOF MscL variants will have a high fluorescence signal at all osmotic strengths. While the apparent simplicity of this assay is attractive, in our experience, it was not sensitive enough to detect the difference in the fluorescence levels between samples with live and dead cells. All cells showed significant fluorescence levels, perhaps due to high background lipid fluorescence noise, or cells becoming transiently permeable to ethidium bromide through the open mechanosensitive channels during osmotic changes.

It is important to ensure that all protocols are thoroughly tested using controls (SaMscL, EcMscL, empty vector, and inducible non-mechanosensitive protein) for reproducibility of results before variants are characterized. Osmotic down-shock assays are notoriously finicky and protocols are often difficult to reproduce. Minute variables such as the temperature of the LB-agar plates used in the experiments, to the method by which cells are spread (beads vs. spatula), have been shown to have significant effects on results. Extensive efforts were made to develop a stringent protocol for a down-shock assay and this still yielded inconsistent results. Further efforts need to be made to find a protocol that gives consistent results.

Electrophysiology

Electrophysiological experiments on MscL variants will yield a more detailed assessment of the properties of the channels. There are two instruments that have been used to take electrophysiological measurements of MscL: traditional patch clamp rig and nan]i[on Port-a-Patch® [4, 5]. Both systems are compatible with *E. coli* giant spheroplasts and reconstituted proteoliposomes. With traditional patch clamp rigs, the patch pipette tip is manipulated under the microscope and touched to the surface of a vesicle. Light suction is applied and the pipette tip forms a giga-ohm seal with the membrane and the membrane patch is excised. A voltage is applied across the membrane and negative pressure is applied to the patch to induce channel gating. Current across the membrane is measured as pressure is applied and discreet jumps in current occur upon channel opening. With the Port-a-Patch®, a planar chip with an aperture is used instead of a

pipette, and vesicle solution is loaded onto the chip. Light suction is applied and a vesicle should land on the aperture, forming a giga-ohm seal. Measurements are then taken the same way as with a traditional patch-pipette.

E. coli Giant Spheroplast Preparation

Detailed below is a protocol that I have used to create *E. coli* giant spheroplasts expressing MscL. MJF465(DE3) cells transformed with plasmids containing MscL variants were cultured overnight in LB media and used to inoculate a fresh culture. Cultures were grown to an OD_{600} = 0.3, then inoculated with IPTG and cultured for another 30 minutes. Cells were pelleted by centrifugation at low speed and resuspended in phosphate buffer. Cells were pelleted again and resuspended in phosphate buffer containing sucrose and lysozyme. Cells were incubated in a shaker at 37 °C for 2 h. Samples were diluted with phosphate buffer containing EDTA at 1:1 ratio and returned to the shaker for another 15 minutes. Cells were pelleted, resuspended in phosphate buffer containing sucrose, pelleted again, and finally resuspended in the same buffer. The size and shape of the cells can be assessed under a microscope. While I and others in the group have used this protocol to prepare sphaeroplasts for electrophysiology, the general experience was that the preparations were difficult to work with. There was a lot of cell debris in the samples which clogged the aperture of the chip of the Port-a-Patch® and also reduced the chances of a viable sphereoplast landing on the aperture. The spheroplasts were quite fragile and when using both the Port-a-Patch® and traditional patch clamp rig, excised patches formed weak seals.

Proteoliposome Preparation

A straightforward protocol was used by Liu *et al.* to create proteoliposomes for analysis of MscL activity [4]. Purified protein is added to azolectin suspension in phosphate buffer. The mixture is incubated at room temperature for 1 h after which biobeads are added (to remove detergent) and incubation is continued for another 3 h. Biobeads are removed by centrifugation at low spin and the lipid/protein mixture is pelleted by ultracentrifugation at high speed. The pellet is washed once with DR buffer (200 mM KCl, 10 mM HEPES pH 7.4), then resuspended in the same buffer. A drop of the suspension is deposited on a slide and allowed to dry overnight in a vacuum desiccator. The resulting lipid film was spotted with DR buffer and rehydrated overnight by vapor diffusion equilibration over a sealed reservoir. The result is a suspension of unilamellar and multilamellar giant vesicles containing the reconstituted protein that can be used for patch clamp experiments. While used successfully in our group, in general, this approach was plagued by formation of multilamellar giant vesicles that could not be used for electrophysiology, since the multiple membranes prevent the flow of ions through the patch pipette, even when MscL is open.

5.2 Identification of Interfaces Necessary for Subunit Exchange

Using the subunit exchange protocol, different MscL constructs can be screened for their ability to form hetero-oligomeric complexes. This will give insight into the interfaces that mediate complex formation. The simple protocol will mix two samples together, one containing MscL variant A and the other containing MscL variant B-sGFP fusion. This

mixture will be initially heated at 4°C, 37°C, 50°C, and 60°C for 4 h and run on a BN-PAGE gel. Subsequent studies can look more carefully at time and temperature dependence of this process. If MscL variant A and B complexes exchange subunits, there will be a ladder of intermediate species. It will be particularly interesting to test if EcMscL and SaMscL can form heterogeneous complexes. The first set of experiments can test various MscL variants against wild-type EcMscL or SaMscL. A preliminary experiment was performed on SaEco30 and SaEco50-sGFP. After incubation at 34 °C for 4 h, results show that they are able to exchange subunits [Fig. 5-2]. SaEco30 and SaEco50 share a similar sequence, differing only in the sequence of the loop. Their ability to form a heterogeneous complex implies that the loop is not critical for complex formation. Previous studies have postulated an intersubunit interaction between adjacent loops, and mutagenesis experiments showed a role of the loop in channel open dwell times and tension sensitivity [6, 7]. Even though the loop is functionally relevant, it may not be important for formation of oligomers. A subunit exchange study of MscL variants will test this hypothesis and reveal more about the role of the other sequence elements in complex formation.
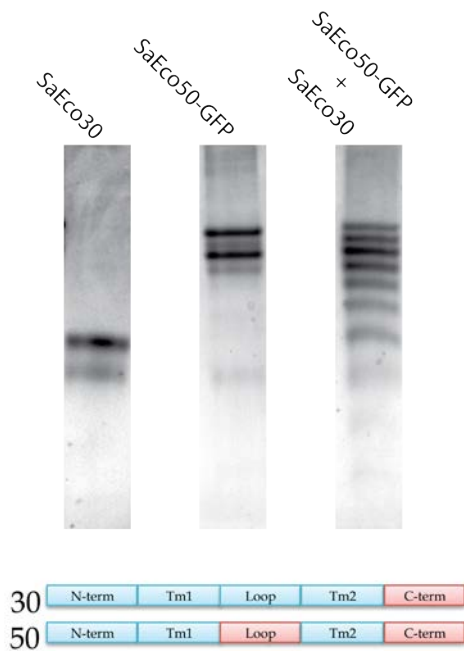
Figure 5-2. Preliminary Experiments Reveal Subunit Exchange Can Occur Between SaEco30 and SaEco50-sGFP Complexes. Purified SaEco30 and SaEco50-sGFP were mixed together and incubated at 37°C for 4 h and run on a BN-PAGE. The results show a ladder of intermediate species as a product of intercomplex subunit exchange.

5.3 In-Membrane Subunit Exchange

Bimolecular fluorescence complementation (BiFC) analysis can be used to probe subunit exchange between MscL complexes embedded in the cell membrane. BiFC uses fluorescent protein fragments fused to two interacting species to probe protein-protein interaction. Here one MscL subunit is fused to YN (residues 1-172, the N-terminal fragment of YFP) and another MscL subunit fused to YC (residues 173-238, the C-terminal fragment of YFP). If these two subunits are in close proximity, the two YFP fragments will reconstitute and regain the ability to fluoresce [8, 9]. Therefore, presumably if two samples, one with homogeneous Mscl-YN complexes and another

with MscL-YC complexes, are mixed and subunit exchange occurs, the YFP fragments will reconstitute and a fluorescent signal will arise.

First, the optimal position for a fluorophore fragment fusion on the MscL subunit must be determined. To get fluorescence when the MscL subunits interact, the fluorescent protein fragment on each subunit must be close enough to interact with each other. The fluorescent protein fragment must also not be placed so that fragments fused to subunits in two different complexes can interact. Trial and error experiments for the placement of the fluorescent protein fragments and the length of the linker must be done. A C-terminal fusion like that used on the MscL-sGFP constructs will probably be effective. *In vitro* subunit exchange experiments on purified protein can be used as a control to test the hypothesis of fluorescence upon subunit exchange. Once this is optimized, the experiment can be carried out in cell membranes.

E. coli membrane vesicles containing over expressed MscL fusions can be prepared using the method detailed by Lee *et al*. [10]. MscL-YN and MscL-YC can be cloned into separate plasmids and transformed into cells (also separately). These cells are then cultured and protein expression induced using a standard protocol. Cells are lysed using a microfluidizer and unlysed cells removed by centrifugation at a low spin. The membranes are pelleted by ultracentrifugation at a high speed and resuspended at 20 mg/ml concentration. At this point, the membranes from cells containing MscL-YN and membranes from cells containing MscL-YC can be mixed at a 1:1 ratio. It may be necessary to run a quantitative western blot to estimate the amount of target protein in

each membrane solution so that equivalent amounts of MscL-YN and MscL-YC are present in the vesicles. The membrane solution is made into vesicles by gentle sonication on ice. Vesicles are subjected to five freeze thaw cycles to randomize the orientation of MscL in the membrane and extruded through a 400 nm polycarbonate filter to form evenly sized vesicles. A fluorescence measurement can be taken as a baseline for subsequent experiments. Vesicles can then be incubated at various temperatures and fluorescence signal monitored over time. If subunit exchange occurs between MscL-YN complexes and MscL-YC complexes, over time the fluorescence intensity should increase. MscL-YN - MscL-YC complex will be visualized at 512 nm excitation wavelength and 529 nm emission wavelength [8].

There are two potential pitfalls in applying this technique to MscL. One is that part of the vesicle preparation protocol calls for five freeze-thaw cycles. These temperature spikes may induce MscL subunit exchange during vesicle preparation and bias the starting fluorescence levels. It may be possible to skip the freeze-thaw step of the protocol, because randomizing MscL orientation in the membrane is not critical for the qualitative purposes of the assay. The second potential pitfall is that once the YC and YN fusions of two MscL subunits reconstitute, the ability of those subunits to dissociate for further subunit exchange may be hindered. Potential pitfalls withstanding, this in-membrane subunit exchange assay can possibly provide an answer to the physiological relevance on subunit exchange.

5.4 Crystallography

Additional SaEco chimeras will be identified as targets for crystallographic studies. Four SaEco chimeras have been screened for crystallization and the observation that SaEco28 ($T_E = 60$ °C) and SaEco31 ($T_E > 60$ °C) crystallized while SaEco32 ($T_E = 44$ °C) and SaEco52 ($T_E = 37$ °C) did not produce crystals provides new criteria to identify promising targets. A high T_E may be correlated to a more stable homogenous population, which is favorable for crystal formation and growth. The next targets could be SaEco38 ($T_E = 60$ °C) and SaEco39 ($T_E > 60$ °C). The best diffracting crystals screened were those of SER mutants SaMscLK62H and SaMscLKKE92AAA. Therefore, a combination of these mutations may improve crystal quality even more. SaMscLK62H,KKE92AAA can be easily cloned by quick change mutagenesis and screened for crystallization.

References

1.   Levina, N., et al., *Protection of Escherichia coli cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity.* EMBO J, 1999. **18**(7): p. 1730-7.

2.   Maurer, J.A. and D.A. Dougherty, *A high-throughput screen for MscL channel activity and mutational phenotyping.* Biochim Biophys Acta, 2001. **1514**(2): p. 165-9.

3.   Powl, A.M., J.M. East, and A.G. Lee, *Lipid-protein interactions studied by introduction of a tryptophan residue: the mechanosensitive channel MscL.* Biochemistry, 2003. **42**(48): p. 14306-17.

4.   Liu, Z., C.S. Gandhi, and D.C. Rees, *Structure of a tetrameric MscL in an expanded intermediate state.* Nature, 2009. **461**(7260): p. 120-4.

5.   Farre, C., et al., *Port-a-patch and patchliner: high fidelity electrophysiology for secondary screening and safety pharmacology.* Comb Chem High Throughput Screen, 2009. **12**(1): p. 24-37.

6.   Yang, L.M., D. Zhong, and P. Blount, *Chimeras reveal a single lipid-interface residue that controls MscL channel kinetics as well as mechanosensitivity.* Cell Rep, 2013. **3**(2): p. 520-7.

7.   Park, K.H., et al., *Purification and functional reconstitution of N- and C-halves of the MscL channel.* Biophys J, 2004. **86**(4): p. 2129-36.

8.   Grinberg, A.V., C.D. Hu, and T.K. Kerppola, *Visualization of Myc/Max/Mad family dimers and the competition for dimerization in living cells.* Mol Cell Biol, 2004. **24**(10): p. 4294-308.

9.   Robida, A.M. and T.K. Kerppola, *Bimolecular fluorescence complementation analysis of inducible protein interactions: effects of factors affecting protein folding on fluorescent protein fragment association.* J Mol Biol, 2009. **394**(3): p. 391-409.

10. Lee, J.Y., et al., *Structural basis for heavy metal detoxification by an Atm1-type ABC exporter.* Science, 2014. **343**(6175): p. 1133-6.

APPENDIX

A.1 Methods and Materials

Cloning – SaEco Chimeras

SaEco and SaEco-sGFP constructs were cloned using homologous recombination and site directed mutagenesis techniques. These constructs were cloned into a pET15b vector under the T7 promoter. All SaEco-sGFP constructs had a short linker (SASGENLYFQSL) with a TEV protease recognition site between the SaEco and sGFP sequences. A pET15b vector containing EcMscL-sGFP was used as a template and was linearized by either a restriction digest at restriction sites NdeI and NheI or PCR amplification using Phusion Polymerase®. The linearized vector had both the N-terminal 6-His tag (MGSSHHHHHHSSGLVPRGSH) and the C-terminal linker and sGFP fusion intact. Gene fragments were PCR amplified from template plasmids containing wild-type EcMscL and SaMscL using Phusion polymerase®. These fragments were then amplified to attach 18 base pair overhangs that had a complementary sequence to the adjacent sequence in the assembled SaEco gene. Assembly of the linearized vector and fragments was done using the GeneArt Seamless Cloning Kit®. The plasmids were then transformed into *E. coli* TOP10® cells, cultured in LB-AMP, extracted with a Qiagen mini-Prep kit®, and sequence verified.

The GeneArt® cloning kit advertises the ability to assemble five fragments, however three fragments was the maximum successfully assembled. Therefore the SaEco chimeras were cloned in a step-wise manner. The first set of chimeras were cloned with two gene fragments, one from EcMscL and the other from SamscL. The second set of chimeras

were cloned with fragments from the first set and each fragment had both the EcMscL and SaMscL sequence. After extended effort, some chimeras proved very difficult to clone, therefore SaEco36-sGFP, SaEco37-sGFP, SaEco41-sGFP, SaEco43-sGFP, SaEco44-sGFP, SaEco45-sGFP, SaEco46-sGFP, SaEco47-sGFP, SaEco49-sGFP, SaEco51-sGFP, SaEco54-sGFP, and SaEco55-sGFP were ordered from GenScript™. To create the SaEco (non-fusion) constructs, a simple quick-change mutagenesis protocol was used to insert a stop codon (TAA) at the end of the SaEco gene.

Cloning – LT Chimeras

The codon optimized genes of LT1, LT2, and LT3 were ordered from GenScript™. LT0 was cloned by Dr. Zhenfeng Liu.

Cloning – SaMscL SER Mutants

All SaMscL SER mutants were cloned by simple site-directed mutagenesis. The quick change PCR protocol used designed primers to incorporate a 1 - 4 base pair mismatch into the target gene to create the desired amino acid mutation. The template DNA, wild-type SaMscL in a pET15b vector, was PCR amplified using Pfu Ultra HF Polymerase®, Dpn1 digested, transformed into *E. coli* Nova Blue® cells, cultured in LB-AMP, extracted with a Qiagen mini-Prep kit®, and sequence verified.

Protein Expression and Purification – Small Scale for OCAM and Subunit Exchange

For small-scale preparations, transformed cells were cultured in 5 ml of ZYM5052-AMP auto-induction media and cultured at 225rmp at 37 °C for 18 - 24 h. Cells were pelleted

at 3,000 rpm for 10 minutes, the supernatant was discarded, and the pellet was resuspended in 500 μl of lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.25 mg/ml Lysozyme, 0.1 mg/ml DNAse, 1x Protease Inhibitor). The solution was homogenized by agitation for 30 minutes at room temperature. Cells were lysed by three freeze/thaw cycles (liquid nitrogen/42°C water bath). Solubilization was done by the addition of DDM to a final concentration of 1% w/v and rotation at 4°C for 2 h. Debris was removed by centrifugation at 3000 rpm for 15 minutes. To further remove cell debris, the supernatant was filtered through a 0.22 μm filter (filter plate).

Ni-NTA Superflow® resin was loaded onto 96-well filter plates to give a column volume (CV) of 100 μl per well. Eluate was removed by centrifugation at 1,000 rpm for 10 minutes. The resin was washed with 5 CV of equilibration buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM imidazole, 0.05% DDM), before the cell extract was loaded. The resin was then washed with 5 CV of equilibration buffer, followed by a wash with 5 CV of high salt buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 25 mM imidazole, 0.05% DDM), followed by another wash with 5 CV of low imidazole buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 82.5 mM imidazole, 0.05% DDM). Finally, the target protein was eluted by a wash with 3 CV of elution buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 300 mM imidazole, 0.05% DDM). Samples were loaded into dialysis tubes (10,000 MW cut-off) and allowed to equilibrate overnight in 2 L of dialysis buffer (10 mM Tris HCl pH 7.5, 150 mM Nacl, 0.02% DDM) to remove imidazole. Samples were used immediately or flash frozen and stored at -80 °C for later use.

Protein Expression and Purification – Large Scale for Crystallization

For large-scale preparations, transformed cells were pre-cultured in TB-AMP media overnight at 37°C, 225 rpm. Pre-cultures were used to inoculate 2 L of fresh TB-AMP media and cells were grown to an optical density at 600 nm of 2.0. Expression was then induced with 1 mM IPTG and allowed to culture for another $2 - 3$ hours. Cells were pelleted at 6,000 rpm for 15 minutes, the supernatant was discarded, and the pellet was resuspended at 1:10 dilution of lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1x Protease Inhibitor, 1% DDM). Cells were then lysed using either a sonicator or a microfluidizer. Cell debris was removed by centrifugation at 14,000 rpm for 45 minutes.

Gravity columns were loaded with 8 ml of Ni-NTA Superflow® resin and then washed with 5 CV of equilibration buffer before the cell extract was loaded. The resin was then washed with 5 CV of equilibration buffer, followed by a wash with 5 CV of high salt buffer, followed by another wash with 5 CV of low imidazole buffer. Finally the target protein was eluted by a wash with 3 CV of elution buffer. To remove imidazole, the sample was further purified on a size exclusion column (Tricorn™ High Performance column packed with Superdex™, GE Healthcare). If removal of the 6His tag was desired, samples were incubated with Thrombin for 42 hours, followed by another size exclusion column to remove Thrombin and cleavage fragments. The protein solution was concentrated down to a final concentration of 12-18 mg/ml.
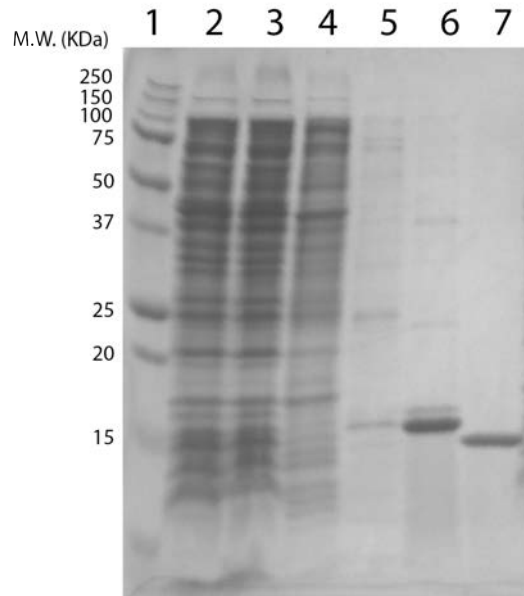
Figure A-1. SDS-PAGE Gel Tracking the Protein Purification Steps. (1) Precision Plus Protein Dual Color Standards (Bio-Rad). (2) Eluate from nickel column after loading cell lysate. (3) Eluate from nickel column after a wash with high salt buffer. (4) Eluate from nickel column after wash with low imidazole buffer. (5) Eluate from nickel column after a wash with elution buffer. (6) Eluate from size exclusion column to remove imidazole. (7) Eluate from from size exclusion column after cleavage of 6His tag.

OCAM

Reaction buffer (final concentration 20 mM Tris pH 7.5, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT) was added to MscL-sGFP samples and proteolysis was initiated by the addition of AcTEV protease at a concentration of 1 U/μg of protein. The reaction was incubated at 34°C (the optimal temperature of TEV protease activity) and aliquots were quenched at 15, 60, and 240 minute time points using 10x quench/loading buffer (50% Glycerol, 0.1% Ponceau-S, 0.1 M Iodoacetamide). Negative controls were also run

111

without added TEV and sampled at the 0 minutes time point and after 240 minutes incubation at 34°C. The products of the reaction were visualized by Blue Native PAGE.

Subunit Exchange

Equivalent amounts of SaEco-sGFP constructs and their corresponding SaEco non-fusion constructs were mixed in a PCR tube and incubated at 4 °C, 37 °C, 50 °C, and 60 °C for 4 h. Samples were then prepared for Blue Native PAGE by addition of 10x loading buffer (50% Glycerol, 0.1% Ponceau-S).

Blue Native PAGE Gels

BN-PAGE gels were run using Bio-Rad™ Criterion precast 4-20% Tris-HCl gradient gels. The first phase of the gel was run with Anode Buffer (50 mM Bis-Tris pH 8.0) and Cathode Buffer B (50 mM Tricine pH 8.0, 15 mM Bis-Tris pH 8.0, 0.02% Coomassie brilliant blue G-250, 2 mM 6-aminocaprooic acid) at 150 V for 60 minutes, after which Cathode Buffer B was removed and replaced with Cathode Buffer A (50 mM Tricine pH 8.0, 15 mM Bis-Tris pH 8.0) and the second phase of the gel was run at 75 V for an additional 999 minutes. All buffers were kept at 4 °C and the gels were run at 4 °C. Gels were stained in in 40% Ethanol/10% Acetic Acid/0.1% Coomassie Brilliant Blue R-250, destained in 40% Ethanol/10% Acetic Acid, and rehydrated in deionized water. Gels were imaged using the Bio-Rad™ ChemiDoc MP imaging system and densitometry analysis was performed using GE Healthcare™ ImageQuant TL v8.1 software. For display purposes, densitometry traces were background subtracted and traces smoothed using the rolling ball algorithm.
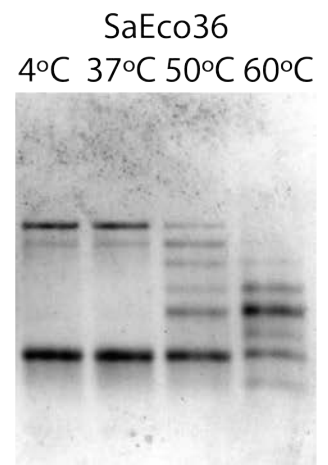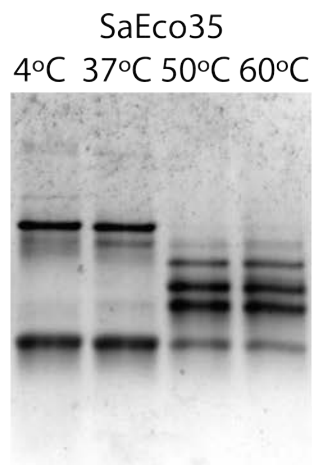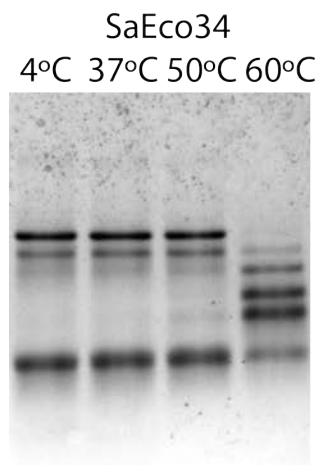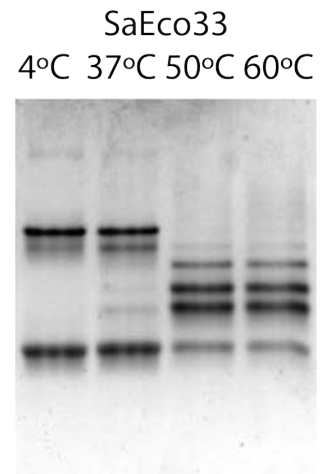
A.2 OCAM Gel Images

SaEco28

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco29

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco30

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco31

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco32

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco33

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco34

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco35

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco36

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco37

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco38

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco39

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco40

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco41

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco42

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco43

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco44

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco45

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

114

SaEco46

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco47

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco48

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco49

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco50

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco51

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco52

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco53

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco54

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

115

SaEco55

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaEco56

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

SaMscL

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

EcMscL

TEV	-	+	+	+	-
Time (min)	-	15	60	240	240

A.3 Subunit Exchange Gel Images

SaEco28
4°C 37°C 50°C 60°C

SaEco29
4°C 37°C 50°C 60°C

SaEco30
4°C 37°C 50°C 60°C

SaEco31
4°C 37°C 50°C 60°C

SaEco32
4°C 37°C 50°C 60°C

SaEco33
4°C 37°C 50°C 60°C

SaEco34
4°C 37°C 50°C 60°C

SaEco35
4°C 37°C 50°C 60°C
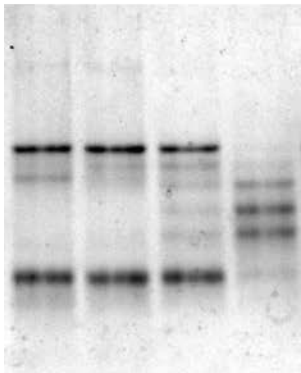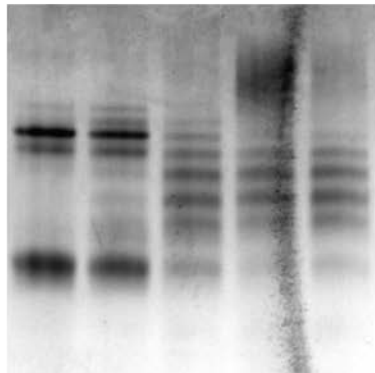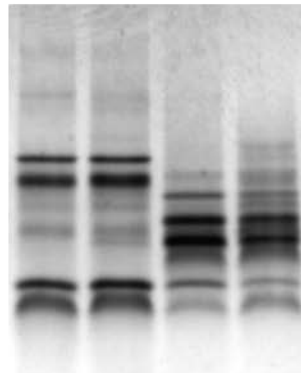
SaEco36
4°C 37°C 50°C 60°C
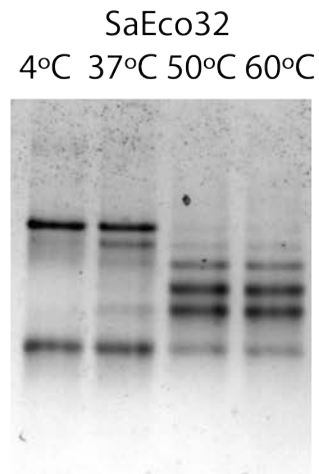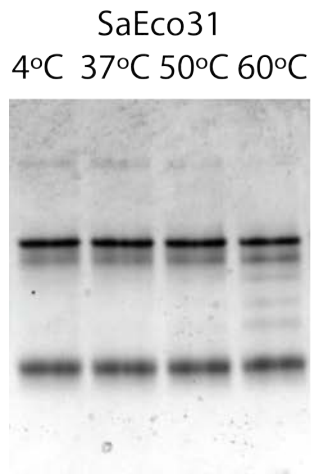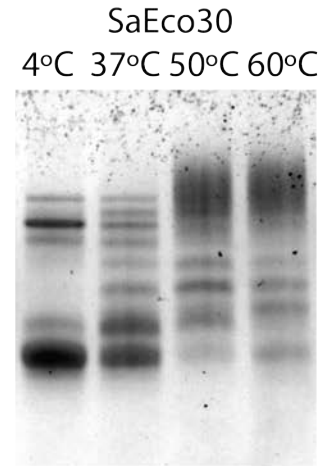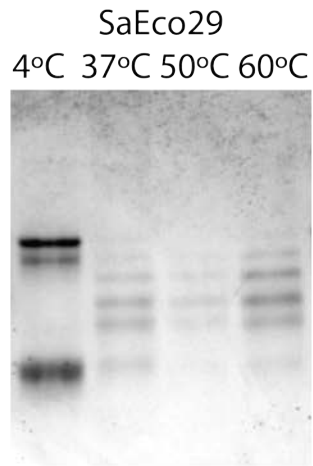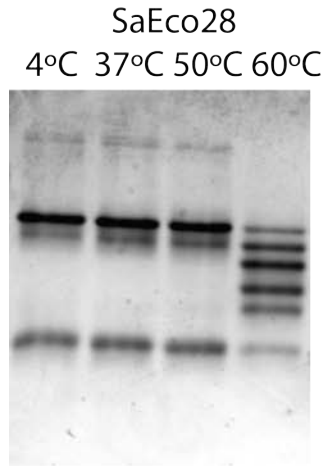
SaEco37
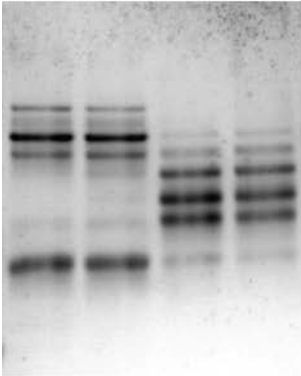
4°C 37°C 50°C 60°C

SaEco38

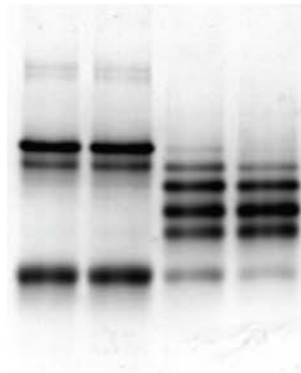4°C 37°C 50°C 60°C
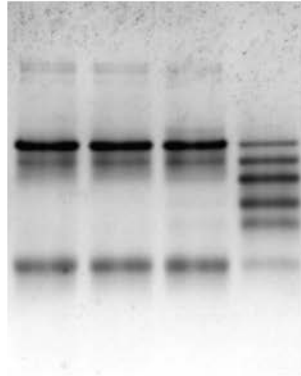
SaEco39

4°C 37°C 50°C 60°C
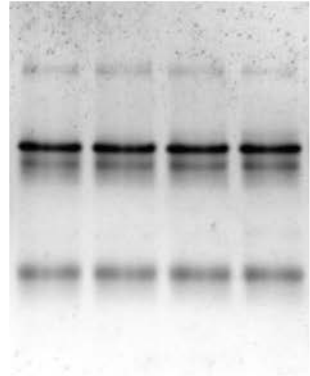
SaEco40

4°C 37°C 50°C 60°C
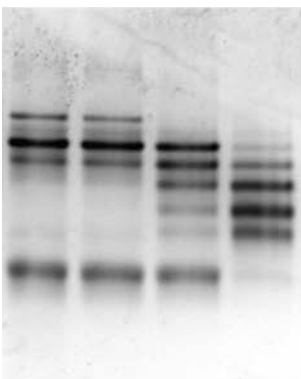
SaEco41

4°C 37°C 50°C 60°C
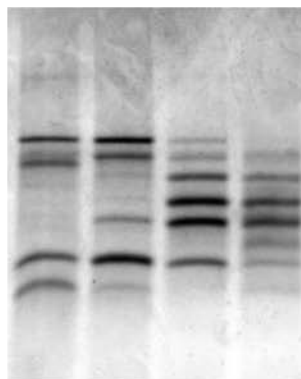
SaEco42

4°C 37°C 50°C 60°C
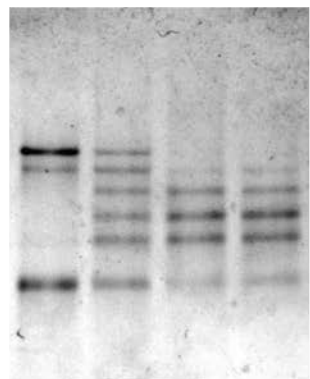
SaEco43

4°C 37°C 50°C 60°C
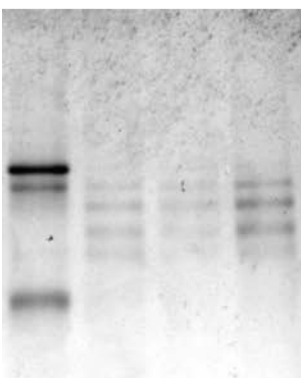
SaEco44

4°C 37°C 50°C 60°C
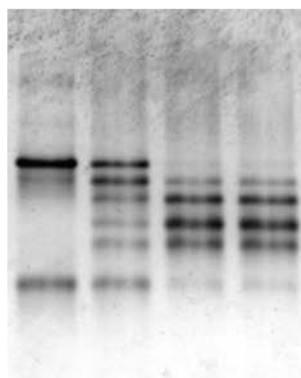
SaEco45

4°C 37°C 50°C 60°C
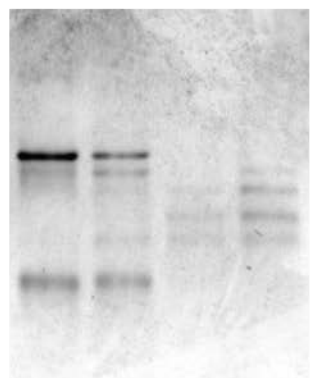
SaEco46
4°C 37°C 50°C 60°C
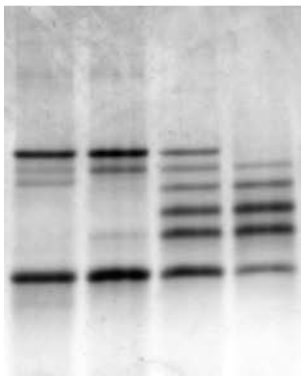
SaEco47
4°C 37°C 50°C 60°C
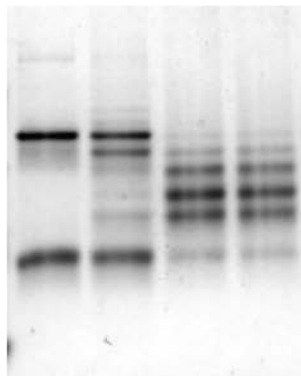
SaEco48
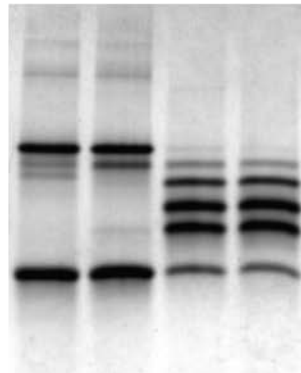4°C 37°C 50°C 60°C
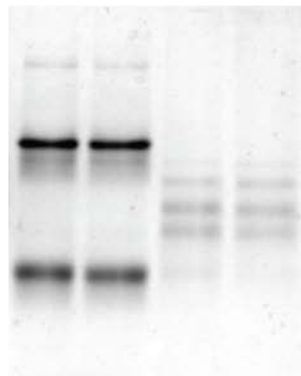
SaEco49
4°C 37°C 50°C 60°C

SaEco50
4°C 30°C 37°C 50°C 60°C

SaEco51
4°C 37°C 50°C 60°C

SaEco52
4°C 37°C 50°C 60°C
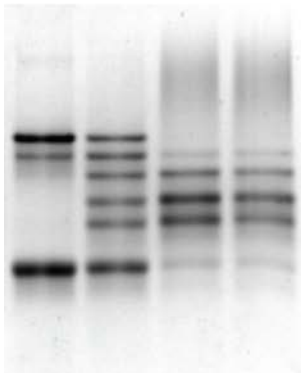
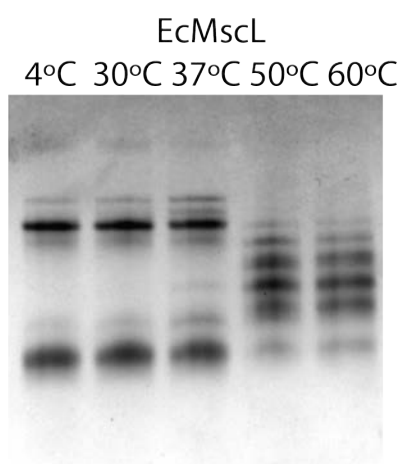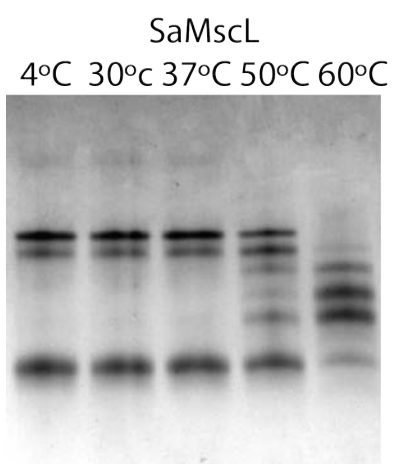SaEco53
4°C 37°C 50°C 60°C

SaEco54
4°C 37°C 50°C 60°C

SaEco55

4°C 37°C 50°C 60°C
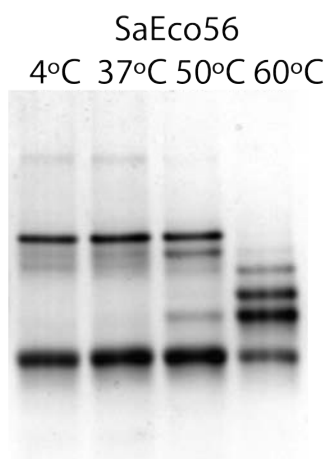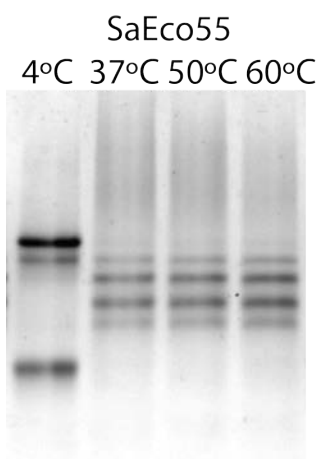
SaEco56

4°C 37°C 50°C 60°C

SaMscL

4°C 30°c 37°C 50°C 60°C

EcMscL

4°C 30°C 37°C 50°C 60°C

A.4 Protein Sequences

Wild-Type MscL
Escherichia coli
MSIIKEFREFAMRGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGGIDFKQF
AVTLRDAQGDIPAVVMHYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEPAAA
PAPTKEEVLLTEIRDLLKEQNNRS

Staphylococcus aureus
MLKEFKEFALKGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIFGSVDFAKEW
SFWGIKYGLFIQSVIDFIIIAFALFIFVKIANTLMKKEEAEEEAVVEENVVLLTEIRD
LLREKK

Mycobacterium tuberculosis
MLKGFKEFLARGNIVDLAVAVVIGTAFTALVTKFTDSIITPLINRIGVNAQSDVGIL
RIGIGGGQTIDLNVLLSAAINFFLIAFAVYFLVVLPYNTLRKKGEVEQPGDTQVVL
LTEIRDLLAQTNGDSPGRHGGRGTPSPTDGPRASTESQ

SaEco Chimeras
27
MSIIKEFREFAMRGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGGIDFKQF
AVTLRDAQGDIPAVVMHYGLFIQSVIDFIIIAFALFIFVKIANTLMKKEEAEEEAVV
EENVVLLTEIRDLLREKK

28
MLKEFKEFALKGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIFGSVDFAKEW
SFWGIKYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEPAAAPAPTKEEVLLTEI
RDLLKEQNNRS

29
MSIIKEFREFAMRGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLFGSVDFAK
EWSFWGIKYGLFIQSVIDFIIIAFALFIFVKIANTLMKKEEAEEEAVVEENVVLLTEI
RDLLREKK

30
MSIIKEFREFAMRGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGGIDFKQF
AVTLRDAQGDIPAVVMHYGVFIQNVFDFLIVAFAIFMAIKLINKLMKKEEAEEEA
VVEENVVLLTEIRDLLREKK

31
MLKEFKEFALKGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIFGSVDFAKEW
SFWGIKYGLFIQSVIDFIIIAFALFIFVKIANTLNRKKEEPAAAPAPTKEEVLLTEIRD
LLKEQNNRS

32

MLKEFKEFALKGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIIGGIDFKQFAV
TLRDAQGDIPAVVMHYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEPAAAPA
PTKEEVLLTEIRDLLKEQNNRS

33

MLKEFKEFALKGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGGIDFKQFA
VTLRDAQGDIPAVVMHYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEPAAAP
APTKEEVLLTEIRDLLKEQNNRS

34

MSIIKEFREFAMRGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIFGSVDFAKE
WSFWGIKYGLFIQSVIDFIIIAFALFIFVKIANTLMKKEEAEEEAVVEENVVLLTEIR
DLLREKK

35

MSIIKEFREFAMRGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIIGGIDFKQFA
VTLRDAQGDIPAVVMHYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEPAAAP
APTKEEVLLTEIRDLLKEQNNRS

36

MSIIKEFREFAMRGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIIGGIDFKQFA
VTLRDAQGDIPAVVMHYGLFIQSVIDFIIIAFALFIFVKIANTLMKKEEAEEEAVVE
ENVVLLTEIRDLLREKK

37

MSIIKEFREFAMRGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIIGGIDFKQFA
VTLRDAQGDIPAVVMHYGVFIQNVFDFLIVAFAIFMAIKLINKLMKKEEAEEEAV
VEENVVLLTEIRDLLREKK

38

MSIIKEFREFAMRGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIFGSVDFAKE
WSFWGIKYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEPAAAPAPTKEEVLLT
EIRDLLKEQNNRS

39

MSIIKEFREFAMRGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIFGSVDFAKE
WSFWGIKYGLFIQSVIDFIIIAFALFIFVKIANTLNRKKEEPAAAPAPTKEEVLLTEI
RDLLKEQNNRS

40

MSIIKEFREFAMRGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIFGSVDFAKE
WSFWGIKYGVFIQNVFDFLIVAFAIFMAIKLINKLMKKEEAEEEAVVEENVVLLT
EIRDLLREKK

41

122

MLKEFKEFALKGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGGIDFKQFA
VTLRDAQGDIPAVVMHYGLFIQSVIDFIIIAFALFIFVKIANTLMKKEEAEEEAVVE
ENVVLLTEIRDLLREKK

42
MLKEFKEFALKGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGGIDFKQFA
VTLRDAQGDIPAVVMHYGVFIQNVFDFLIVAFAIFMAIKLINKLMKKEEAEEEAV
VEENVVLLTEIRDLLREKK

43
MLKEFKEFALKGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLFGSVDFAKE
WSFWGIKYGLFIQSVIDFIIIAFALFIFVKIANTLMKKEEAEEEAVVEENVVLLTEIR
DLLREKK

44
MLKEFKEFALKGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLFGSVDFAKE
WSFWGIKYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEPAAAPAPTKEEVLLT
EIRDLLKEQNNRS

45
MLKEFKEFALKGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLFGSVDFAKE
WSFWGIKYGLFIQSVIDFIIIAFALFIFVKIANTLNRKKEEPAAAPAPTKEEVLLTEI
RDLLKEQNNRS

46
MLKEFKEFALKGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGGIDFKQFA
VTLRDAQGDIPAVVMHYGLFIQSVIDFIIIAFALFIFVKIANTLNRKKEEPAAAPAP
TKEEVLLTEIRDLLKEQNNRS

47
MSIIKEFREFAMRGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLFGSVDFAK
EWSFWGIKYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEPAAAPAPTKEEVLL
TEIRDLLKEQNNRS

48
MSIIKEFREFAMRGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLFGSVDFAK
EWSFWGIKYGLFIQSVIDFIIIAFALFIFVKIANTLNRKKEEPAAAPAPTKEEVLLTE
IRDLLKEQNNRS

49
MSIIKEFREFAMRGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGGIDFKQF
AVTLRDAQGDIPAVVMHYGLFIQSVIDFIIIAFALFIFVKIANTLNRKKEEPAAAPA
PTKEEVLLTEIRDLLKEQNNRS

50

MSIIKEFREFAMRGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLFGSVDFAK
EWSFWGIKYGVFIQNVFDFLIVAFAIFMAIKLINKLMKKEEAEEEAVVEENVVLL
TEIRDLLREKK

51
MLKEFKEFALKGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIIGGIDFKQFAV
TLRDAQGDIPAVVMHYGLFIQSVIDFIIIAFALFIFVKIANTLMKKEEAEEEAVVEE
NVVLLTEIRDLLREKK

52
MLKEFKEFALKGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIIGGIDFKQFAV
TLRDAQGDIPAVVMHYGVFIQNVFDFLIVAFAIFMAIKLINKLMKKEEAEEEAVV
EENVVLLTEIRDLLREKK

53
MLKEFKEFALKGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIFGSVDFAKEW
SFWGIKYGVFIQNVFDFLIVAFAIFMAIKLINKLMKKEEAEEEAVVEENVVLLTEI
RDLLREKK

54
MLKEFKEFALKGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIIGGIDFKQFAV
TLRDAQGDIPAVVMHYGLFIQSVIDFIIIAFALFIFVKIANTLNRKKEEPAAAPAPT
KEEVLLTEIRDLLKEQNNRS

55
MLKEFKEFALKGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLFGSVDFAKE
WSFWGIKYGVFIQNVFDFLIVAFAIFMAIKLINKLMKKEEAEEEAVVEENVVLLT
EIRDLLREKK

56
MSIIKEFREFAMRGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIIGGIDFKQFA
VTLRDAQGDIPAVVMHYGLFIQSVIDFIIIAFALFIFVKIANTLNRKKEEPAAAPAP
TKEEVLLTEIRDLLKEQNNRS

LT Chimeras
LT0
MSIIKEFREFAMRGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGGIDFKQF
AVTLRDAQGDIPAVVMHYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEVEQP
GDTQVVLLTEIRDLLAQTNGDSPGRHGGRGTPSPTDGPRASTESQ

LT1
MSIIKEFREFAMRGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGVNAQSD
VGILRIGIGGGQTIDYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEVEQPGDTQ
VVLLTEIRDLLAQTNGDSPGRHGGRGTPSPTDGPRASTESQ

LT2
MLKGFKEFLARGNIVDLAVAVVIGTAFTALVTKFTDSIITPLINRIGGIDFKQFAVT
LRDAQGDIPAVVMHLNVLLSAAINFFLIAFAVYFLVVLPYNTLRKKGEPAAAPAP
TKEEVLLTEIRDLLKEQNNRS

LT3
MLKEFKEFALKGNVLDLAIAVVMGAAFNKIISSLVENIIMPLIGKIFGSVDFAKEW
SFWGIKYGLFIQSVIDFIIIAFALFIFVKIANTLRKKGEVEQPGDTQVVLLTEIRDLL
AQTNGDSPGRHGGRGTPSPTDGPRASTESQ

Fusion Proteins
6his tag
MGSSHHHHHHSSGLVPRGSH

sGFP Fusion (+ TEV Linker)
SASGENLYFQSLSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTL
KFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIS
FKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITA
DKQKNGIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLS
KDPNEKRDHMVLLEFVTAAGITHGMDELYK