

**STRUCTURAL ANALYSIS AND CONFORMATIONAL DYNAMICS OF
THE YEAST ISOPRENYLCYSTEINE CARBOXYL
METHYLTRANSFERASE, STE14**

by

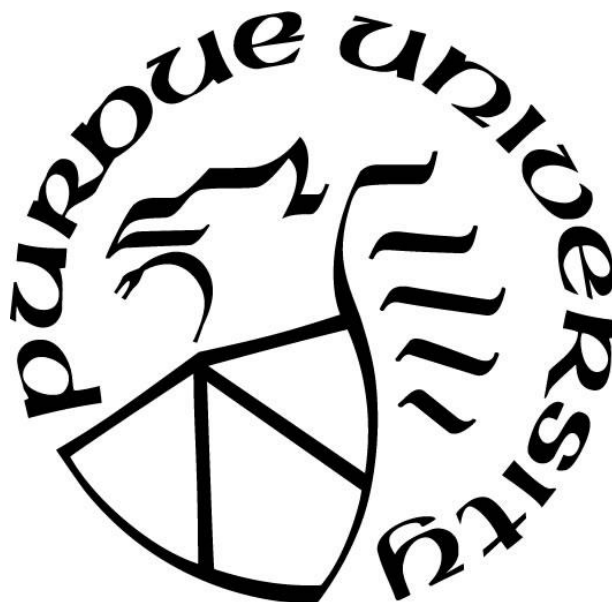
Anna C. Ratliff

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy



Department of Chemistry

West Lafayette, Indiana

December 2019

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Christine A. Hrycyna, Chair

Department of Chemistry

Dr. Chittaranjan Das

Department of Chemistry

Dr. Cynthia V. Stauffacher

Department of Biological Sciences

Dr. Angeline M. Lyon

Department of Chemistry

Approved by:

Dr. Christine A. Hrycyna

For Graham and my family

ACKNOWLEDGMENTS

There are many people I need to thank for their support, love, and understanding over the past five years. First, I need to recognize my family: Chris Ratliff, Elaine Ratliff, Kate Ratliff, and Carol Norgaard. Mom, thank you for always being available to talk to during the day. You were always there to help me with a problem or put a smile on my face. Dad, thank you for teaching me how to anticipate so successfully. I have always been impressed by your hard work and dedication to what you do, and I have continually wanted to have your work ethic. I appreciate the opportunity that both my parents gave me to have a great education and excel at York High School, Lawrence University, and Purdue University. Kate, thank you for always calling me on the way home from work and making me laugh. You and Alby were always there to teach me how to be an adult and the necessity of work-life balance. Grandma, thank you for your support my whole life. You and Grandpa have taught and shown me that hard work and kindness can take you a long way in life. My childhood would have been so much different if you had not lived eight blocks away and I would not have had it any other way.

Second, I would like to thank my biggest supporter and fiancé, Graham Jones and his family. Graham, I do not think I can thank you enough in this acknowledgment. The amount of times you have calmed me down when I was stressed, listened to presentations about biochemistry, and loved me no matter what, is amazing. You always gave me confidence when I did not feel smart enough or prepared. Je t'aime, Graham. On to our next adventures.

Third, I would like to thank my advisor, Dr. Christine Hrycyna. Thank you for your mentorship as a scientist, teacher, professional and friend. You taught me to be confident in myself and that women in science have no limits to their accomplishments as scientists, faculty and administrators. I truly wish the best for you, Alan, Oliver, Oscar and, of course, Eddie. Also, thank you to my thesis committee and Purdue University professors for all your advice over the past few years. Specifically, I need to thank Dr. Angeline Lyon for guidance towards my projects, allowing me to travel to Argonne National Laboratory and being an incredible person in which to confide. Additionally, thank you to Dr. Nick Noinaj and Runrun Wu for helping me with many biophysical techniques and all the advice on membrane protein purifications.

Finally, I would like to thank the Hrycyna lab, current and previous members. Everyone has helped me grow as a person and a scientist. I appreciate all the constructive criticism I was given because it helped to push me beyond what I thought I could do. Karen Olsen, thank you for your patience with teaching me about our research on Ste14p. Your precision toward your lab work and encouragement was always appreciated. Liz Garland-Kuntz, I cannot thank you enough. I have never met someone so intelligent, hardworking, and detail oriented as you. Thank you for teaching me new techniques, spending hours helping with my work before your own and, at the same time, being a great friend. I love talking to you and I am going to miss you so much. Ari Cardillo, I am so happy to have you as my desk mate since we are always on the same page with each other. Thank you for reading me the news, being a determined scientist, and one of my best friends at Purdue. I am so lucky to have met you and to have you and Zu as neighbors. Best of luck with where life takes you and I know we will always be friends. Sahej Bains, thank you for all the hard work you put into our manuscript and being my favorite person to talk to in lab. You always were willing to get coffee (or chai) and have a fun time when we both needed it. You are going to be the best M.D. Ph.D and I will always be here for you. Carly Schnoebelen, thank you for being an unofficial Hrycyna lab member. I have no idea how I would have managed being a teaching assistant without you for three years. You and Piper were the best running buddies, got me through my OP and were always the ones I wanted to go on adventures with in Indiana. I especially need to thank my undergraduate researcher, Alex Piroozi. You were always willing to help no matter the task or the amount of time it would take. You would always put a smile on my face even if I was stressed. You ended up not just becoming a knowledgeable researcher, but also a great friend. I wish the best to you and your future endeavors wherever they may take him. Lastly, Chelsea, Jason, Elias, Allison, Patty, and all other members, thank you for all the good times, bad times, laughs and complaining. I know you all will be successful and intelligent as scientists or with whatever you choose to pursue in your lives. Keep getting those gold stars.

TABLE OF CONTENTS

| | |
|--|----|
| LIST OF TABLES | 10 |
| LIST OF FIGURES | 11 |
| LIST OF ABBREVIATIONS | 13 |
| ABSTRACT..... | 17 |
| CHAPTER 1. INTRODUCTION: A BIOLOGICAL UNDERSTANDING OF THE INTERGRAL MEMBRANE PROTEIN, ISOPRENYLCYSTEINE CARBOXYL METHYLTRANSFERASE (ICMT) | 20 |
| 1.1 CaaX Proteins and Processing | 20 |
| 1.1.1 S-Isoprenylation | 20 |
| 1.1.2 Endoproteolysis | 21 |
| 1.1.3 Methyl Esterification | 23 |
| 1.2 Diseases Associated with CaaX Proteins..... | 23 |
| 1.3 Characterization of Human Icmt and the Yeast Homolog, Ste14..... | 25 |
| 1.3.1 General Characteristics, Localization, and Two-Dimensional Topology | 25 |
| 1.3.2 Hypothesized Kinetic Mechanism of Icmt | 27 |
| 1.4 Structural Analysis of Co-factor and Substrate Binding Sites of Icmt | 28 |
| 1.4.1 Icmts are Putative Metalloenzymes with Zinc Dependence..... | 28 |
| 1.4.2 First Icmt Structural Model in Eukaryotic Organism | 28 |
| 1.4.3 First Icmt Structural Model in Prokaryotic Organism..... | 29 |
| 1.4.4 Biochemical Analysis of Icmt Residues Important for Binding..... | 30 |
| 1.4.5 Substrate Specificity | 32 |
| 1.5 Inhibition of CaaX Proteins | 33 |
| 1.5.1 Challenges of Ras Inhibition | 33 |
| 1.5.2 Inhibition of Protein S-Isoprenylation | 34 |
| 1.5.3 Inhibition of Endoproteolysis | 34 |
| 1.5.4 Inhibition of Methyl Esterification | 35 |
| 1.6 References | 42 |
| CHAPTER 2. OPTIMIZATION OF EXPRESSION AND PURIFICATION OF HIS-STE14 UTILIZING THE BACULOVIRUS EXPRESSION SYSTEM | 57 |

| | | |
|--|--|----|
| 2.1 | Introduction..... | 57 |
| 2.2 | Materials and Methods..... | 58 |
| 2.2.1 | Materials | 58 |
| 2.2.2 | Yeast Strains, Transformation, and Yeast Crude Membrane Preparations | 59 |
| 2.2.3 | Expression and Membrane Preparation of His-Ste14 from Insect Cell Lines..... | 60 |
| 2.2.4 | <i>In vitro</i> Methyltransferase Vapor Diffusion Assay of Yeast and Insect Cell Crude Membranes | 61 |
| 2.2.5 | SDS-PAGE Staining and Immunoblot Analysis | 61 |
| 2.2.6 | Purification of His-Ste14 | 62 |
| 2.3 | Results and Discussion | 63 |
| 2.3.1 | Expression in Yeast vs Insect Cells Lines | 63 |
| 2.3.2 | Optimization of Size Exclusion Chromatography of His-Ste14..... | 64 |
| 2.3.3 | Detergent Screening of Pure His-Ste14..... | 65 |
| 2.4 | Future Directions | 67 |
| 2.5 | References..... | 79 |
| CHAPTER 3. ASSESSMENT OF STE14 CO-FACTOR BINDING DYNAMICS USING ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY | | 82 |
| 3.1 | Introduction..... | 82 |
| 3.2 | Materials and Methods..... | 84 |
| 3.2.1 | Materials | 84 |
| 3.2.2 | Cloning | 85 |
| 3.2.3 | Yeast Strains, Transformation, and Crude Membrane Preparations | 86 |
| 3.2.4 | <i>In vitro</i> Methyltransferase Vapor Diffusion Assay of Crude Membranes | 87 |
| 3.2.5 | Purification of His-Ste14 and His-Ste14 Mutants | 87 |
| 3.2.6 | Purification of His-Ste14 and His-Ste14 Mutants with the Addition of MTSL Spin Label | 88 |
| 3.2.7 | SDS-PAGE Staining and Immunoblot Analysis | 89 |
| 3.2.8 | <i>In vitro</i> Methyltransferase Vapor Diffusion Assay of Pure His-Ste14 and Mutants. | 90 |
| 3.2.9 | EPR Spectroscopy Conditions..... | 90 |
| 3.2.10 | TEMPO Standard Curve | 91 |
| 3.3 | Results and Discussion | 91 |

| | | |
|--|--|-----|
| 3.4 | Future Directions | 93 |
| 3.5 | References | 111 |
| CHAPTER 4. PRELIMINARY STRUCTURAL ANALYSIS OF HIS-STE14 UTILIZING VARIOUS BIOPHYSICAL TECHNIQUES | | 113 |
| 4.1 | Introduction..... | 113 |
| 4.2 | Materials and Methods..... | 114 |
| 4.2.1 | Materials | 114 |
| 4.2.2 | Expression and Membrane Preparation of His-Ste14 from Insect Cell Lines..... | 115 |
| 4.2.3 | <i>In vitro</i> Methyltransferase Vapor Diffusion Assay of Yeast and Insect Cell Crude Membranes | 116 |
| 4.2.4 | SDS-PAGE Staining and Immunoblot Analysis | 116 |
| 4.2.5 | Purification of His-Ste14 | 117 |
| 4.2.6 | Incorporation of His-Ste14 into Nanodiscs | 118 |
| 4.2.7 | Negative Stain EM Sample Preparation and Imaging | 118 |
| 4.2.8 | Cryo-EM Sample Preparation and Imaging | 119 |
| 4.2.9 | Introduction of His-Ste14 into Nanodiscs | 119 |
| 4.2.10 | Negative Staining with His-Ste14 Incorporated into Nanodiscs..... | 120 |
| 4.2.11 | Preliminary Cryo-EM Data Analysis | 121 |
| 4.3 | Future Directions | 121 |
| 4.4 | References | 125 |
| CHAPTER 5. CHARACTERIZATION OF THE OLIGOMERIC STATE OF STE14 AND THE GXXXG DIMERIZATION MOTIF..... | | 129 |
| 5.1 | Introduction..... | 129 |
| 5.2 | Materials and Methods..... | 131 |
| 5.2.1 | Materials | 131 |
| 5.2.2 | Expression and Membrane Preparation of His-Ste14 from Insect Cell Lines..... | 131 |
| 5.2.3 | <i>In vitro</i> Methyltransferase Vapor Diffusion Assay of Yeast and Insect Cell Crude Membranes | 132 |
| 5.2.4 | SDS-PAGE Staining and Immunoblot Analysis | 132 |
| 5.2.5 | Purification of His-Ste14 | 133 |
| 5.2.6 | SEC-MALS-SAXS data collection and analysis | 134 |

| | |
|----------------------------------|-----|
| 5.3 Results and Discussion | 135 |
| 5.4 Future Directions | 136 |
| 5.5 References..... | 142 |
| APPENDIX A..... | 145 |
| APPENDIX B | 148 |
| VITA..... | 150 |
| PUBLICATIONS..... | 151 |

LIST OF TABLES

| | |
|---|-----|
| Table 2.1 Characteristics of detergents used in this study | 68 |
| Table 2.2 Percent (w/v) of detergents used for purification. | 69 |
| Table 3.1 Oligonucleotides used in this study to generate <i>STE14</i> mutants | 95 |
| Table 3.2 <i>S. cerevisiae</i> plasmids used in this study | 97 |
| Table 3.3 Specific methyltransferase activities of His-Ste14-TA mutants from crude membranes tested in this study | 100 |
| Table 3.4 Specific methyltransferase activities of purified His-Ste14-TA mutants tested in this study | 105 |
| Table 5.1 SAXS structural parameters of His-Ste14 | 140 |

LIST OF FIGURES

| | |
|---|-----|
| Figure 1.1 The CaaX processing pathway | 37 |
| Figure 1.2 Topology map of <i>Saccharomyces cerevisiae</i> Icmt (Ste14). | 38 |
| Figure 1.3 Sequence alignment of Icmt orthologs. | 39 |
| Figure 1.4 Sequentially ordered bi-bi reaction mechanism unique to the Icmt family..... | 40 |
| Figure 1.5 Crystal structures of <i>Methanosarcina acetivorans</i> Icmt (Ma-Icmt) and <i>Tribolium castaneum</i> Icmt (beetle) | 41 |
| Figure 2.1 Expression of His-Ste14 extracted from yeast and insect cell lines by immunoblotting. | 70 |
| Figure 2.2 Purification of His-Ste14 purified in various buffers. | 71 |
| Figure 2.3 Purification of His-Ste14 with DDM detergent at pH 7.2 and 6.0. | 72 |
| Figure 2.4 Purification of His-Ste14 with LMNG detergent at pH 7.2. | 73 |
| Figure 2.5 Purification of His-Ste14 with C12E7 detergent at pH 7.2. | 74 |
| Figure 2.6 Purification of His-Ste14 with DMNG detergent at pH 7.2. | 75 |
| Figure 2.7 Purification of His-Ste14 with C8E4 detergent at pH 7.2. | 76 |
| Figure 2.8 Purification of His-Ste14 with LDAO detergent at pH 7.2. | 77 |
| Figure 2.9 <i>In vitro</i> vapor diffusion methyltransferase activity assay of His-Ste14 solubilized in various detergents at pH 7.2..... | 78 |
| Figure 3.1 Expression levels of His-Ste14-TA and single cysteine mutants. | 99 |
| Figure 3.2 <i>In vitro</i> vapor diffusion methyltransferase activity assay of 46 non-conserved cytosolic His-Ste14-TA single cysteine mutants in crude membranes. | 102 |
| Figure 3.3 Expression levels and purity of purified His-Ste14-TA and single cysteine mutants. | 103 |
| Figure 3.4 Purity of His-Ste14-TA and single cysteine mutants. | 104 |
| Figure 3.5 <i>In vitro</i> vapor diffusion methyltransferase activity assay of nineteen purified non-conserved cytosolic His-Ste14-TA single cysteine mutants. | 106 |
| Figure 3.6 Heat map of single cysteine mutants in crude membranes..... | 107 |
| Figure 3.7 EPR spectral analysis of purified His-Ste14-V212C-TA. | 108 |
| Figure 3.8 EPR spectral analysis of purified His-Ste14-K231C-TA. | 109 |
| Figure 3.9 TEMPO standard curve. | 110 |

| | |
|--|-----|
| Figure 4.1 Confirmation of His-Ste14 incorporated into MSP 1E3D1 nanodiscs..... | 122 |
| Figure 4.2 Unadjusted micrograph field views of His-Ste14 in nanodiscs using negative staining EM..... | 123 |
| Figure 4.3 Unadjusted micrograph field views of His-Ste14 in nanodiscs using negative cryo-EM. | 124 |
| Figure 5.1 Topology of <i>Saccharomyces cerevisiae</i> Icmt, Ste14..... | 137 |
| Figure 5.2 Sequence alignment of Icmt orthologs. | 138 |
| Figure 5.3 Characterization of Ste14 via SEC-MALS-SAXS confirms dimerization state. | 139 |
| Figure 5.4 Predicted structural model of Ste14 dimerization. | 141 |

LIST OF ABBREVIATIONS

| | |
|--------------------------------|--|
| AEBSF | 4-(2-Aminoethyl)benzensulfonyl fluoride hydrochloride |
| AFC | <i>N</i> -acetyl- <i>S</i> -farnesyl- <i>L</i> -cysteine |
| AGGC | <i>N</i> -acetyl- <i>S</i> -geranylgeranyl- <i>L</i> -cysteine |
| <i>Ag</i> -Icmt | <i>Anopheles gambiae</i> Isoprenylcysteine carboxyl methyltransferase |
| ABCG2 | ATP binding cassette subfamily G member 2 |
| BFC | Biotin- <i>S</i> -farnesyl- <i>L</i> -cysteine |
| BMH | <i>Bis</i> -maleimidohexane |
| BMOE | <i>Bis</i> -maleimidoethane |
| BS ³ | Bi(sulfosuccinimidyl)suberate |
| C ₈ E ₄ | Tetraethylene glycol monododecyl ether |
| C ₁₂ E ₇ | Heptaethylene glycol monododecyl ether |
| CA | Carbonic anhydrase |
| Cdc42 | Cell division cycle 42 |
| CHAPS | 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate |
| CHAPSO | 3-[(3-cholamidopropyl)dimethylammonio]-2 hydroxyl-1-propanesulfonate |
| CIP | Calf intestinal alkaline phosphatase |
| CMC | Critical micelle concentration |
| Cryo-EM | single particle cryogenic electron microscopy |
| CW | Continuous-wave |
| DDM | <i>N</i> -Dodecyl- β -D-maltopyranoside |
| DEER | Double electron-electron resonance |
| DI | Double integral |
| D _{max} | Dimensional maximum |
| DMNG | 2,2-dioctylpropane-1,3-bis- β -D-maltopyranoside, Decyl maltose neopentyl glycol |
| DTT | Dithiothreitol |
| ECL | Enhanced chemiluminescence |
| EDTA | Ethylenediaminetetraacetic acid |
| EM | Electron microscopy |
| EPR | Electron paramagnetic resonance |

| | |
|------------------|---|
| ER | Endoplasmic reticulum |
| FPP | Farnesyl pyrophosphate |
| FTA | S-farnesyl thioacetic acid |
| FTase | Farnesyltransferase |
| FTIs | Farnesyltransferase inhibitors |
| FTA | S-farnesyl-thioacetic acid |
| FTP | S-farnesyl thiopropionic acid |
| FTS | S-farnesyl thiosalicylic acid |
| GGTP | Geranylgeranyl pyrophosphate |
| GGTase I | Geranylgeranyltransferase I |
| GGTase II | Geranylgeranyltransferase II |
| GGTase III | Geranylgeranyltransferase III |
| GGTIs | Geranylgeranyltransferase I inhibitors |
| GpA | Glycophorin A |
| High Five | <i>Trichoplusia ni</i> |
| HEPES | 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid |
| hIcmt | Human Isoprenylcysteine carboxyl methyltransferase |
| His-Ste14 | His ₁₀ -myc ₃ N-Ste14 |
| His-hIcmt | His ₁₀ -myc ₃ N-hIcmt |
| HRP | Horseradish peroxidase |
| IC ₅₀ | Inhibitory concentration 50% |
| Icmt | Isoprenylcysteine carboxyl methyltransferase |
| IgG | Immunoglobulin G |
| kDa | Kilodalton |
| LB | Luria-Bertani |
| LB-AMP | Luria-Bertani + ampicillin |
| LDAO | Lauryldimethylamine- <i>N</i> -oxide |
| LMNG | 2,2-didecylpropane-1,3-bis-β-D-maltopyranoside, Lauryl maltose neopentyl glycol |
| Ma-Icmt | <i>Methanosarcina acetivorans</i> Isoprenylcysteine carboxyl methyltransferase |
| MALS | Multi-angle light scattering |
| mAU | Milli absorbance unit |

| | |
|------------------------|--|
| MEFs | Mouse embryonic fibroblast |
| MSPs | Membrane scaffolding proteins |
| MTSL | 1-oxy1-2,2,5,5-tetramethylpyrroline-3-methyl)-methanethiosulfonate |
| MWCO | Molecular weight cutoff |
| NMR | Nuclear magnetic resonance |
| OD | Optical density |
| OG | <i>N</i> -octyl- β -D-glucopyranoside |
| OTG | <i>N</i> -octyl- β -D-thioglucopyranoside |
| PAGE | Polyacrylamide gel electrophoresis |
| PBS | Phosphate-buffered saline |
| PBST | Phosphate-buffered saline with Tween-20 |
| PCR | Polymerase chain reaction |
| PDAC | Pancreatic ductal adenocarcinomas |
| PGK | Phosphoglycerate kinase promoter |
| P(r) | Pair-distance distributions |
| p.s.i | Pounds per square inch |
| RabGGTase | Rab geranylgeranyl transferase |
| Rce1 | Ras converting enzyme 1 |
| R _g | Radius of gyration |
| <i>S. cerevisiae</i> | <i>Saccharomyces cerevisiae</i> |
| SAH | <i>S</i> -adenosyl-L-homocysteine |
| SAM | <i>S</i> -adenosyl-L-methionine |
| [¹⁴ C]-SAM | <i>S</i> -adenosyl-L-[methyl- ¹⁴ C] methionine |
| SAXS | Small angle light scattering |
| SC-URA | Synthetic complete medium without uracil |
| S.D. | Standard deviation |
| SDS | Sodium dodecyl sulfate |
| SDSL | Sight-directed spin labeling |
| SEC | Size exclusion chromatography |
| Sf9 | <i>Spondoptera frugiperda</i> |
| TA | Triple A |

| | |
|-------|--------------------------------------|
| TCA | Trichloroacetic acid |
| TEM | Transmission electron microscope |
| TEMPO | 2,2,6,6-tetramethyl-1-piperidinyloxy |
| TM | Transmembrane |
| UV | Ultraviolet |
| v/v | Volume per volume |
| WT | Wild-type |
| Wt | Weight |
| wt/v | Weight per volume |
| YPD | Yeast peptone dextrose medium |

ABSTRACT

CaaX proteins are involved in many key cellular processes such as proliferation, differentiation, trafficking, and gene expression. CaaX proteins have four specific C-terminal amino acids designated as a CaaX motif, where the “C” is a cysteine, “a” are aliphatic residues, and “X” represents one of several amino acids. Proteins with this motif undergo three post-translational modifications: isoprenylation of the cysteine residue, endoproteolysis of the –aaX residues and methylation of the isoprenylated cysteine, which is necessary for their localization in the cell and function. Due to involvement of CaaX proteins in many critical signaling pathways, mutations in CaaX proteins can result in a wide variety of disorders and carcinomas. Most notably, mutants in the *KRAS* gene are associated with 90% of pancreatic cancers and 30% of all cancers. Isoprenylcysteine carboxyl methyltransferase (Icmt), an integral membrane protein in the endoplasmic reticulum, is the only known protein responsible for the post-translational α -carboxyl methylesterification of the C-terminus of CaaX proteins. Cells with Icmt deficiency causes the small G-protein, K-ras, to be mislocalized and decreases downstream signaling of K-ras. Thus, our goal is to better understand the structure and methylation mechanism of Icmt in order to inhibit mutant K-ras in oncogenic cells and aid in the creation of a chemotherapeutic for pancreatic cancer.

Icmt studies have focused on the founding member of the Icmt family, Ste14. Ste14 is expressed in *Saccharomyces cerevisiae* (*S. cerevisiae*) and shares high homology with the human Icmt (hIcmt), which has yet to be functionally purified. Specifically, hIcmt and Ste14 share 63% similarity and 41% identity, mostly within the C-termini of the proteins. First, we optimized expression and purification of Ste14 in order to generate a larger yield of protein, which is necessary for many biophysical techniques. Infection of Sf9 cells with a baculovirus expressing an N-terminally 10-His-tagged and 3-myc-tagged Ste14 (His-Ste14), increased protein expression between four and five-fold compared to our yeast model and used significantly less starting materials. We also performed a detergent screen for the purification of His-Ste14 from insect cell expression. We concluded that *n*-Dodecyl- β -D-maltopyranoside (DDM), lauryl maltose neopentyl glycol (LMNG), and heptaethylene glycol monododecyl ether (C₁₂E₇) were detergents that stabilize His-Ste14 for further biophysical techniques. Additionally, we found 1xEQ buffer at pH 6.0 resulted in the most homogenous His-Ste14 sample.

Second, we sought to elucidate the SAM binding/ SAH release mechanism of His-Ste14 by utilizing a combinatorial method of site-directed spin labeling and electron paramagnetic resonance (EPR) spectroscopy analysis. We used SDSL-EPR to determine the conformational dynamics of His-Ste14 with and without SAM. EPR is an attractive method to study conformational changes of proteins as it is done in solution and requires relatively small amounts of protein. We generated a library of 46 non-conserved single cysteine mutants introduced into cysteine-less His-Ste14 (His-Ste14-TA). The cysteine residues engineered into His-Ste14-TA were in the cytosolic portion of the protein to ensure efficient labeling and were tested for methyltransferase activity levels. From crude membranes, only nineteen mutants retained activity levels of $\geq 50\%$ of His-Ste14-TA, which were then purified and tested for methyltransferase activity levels. Eight purified mutants were selected as candidates for EPR with activity levels of $\geq 50\%$ of His-Ste14-TA. Once optimized, we introduced a nitroxide spin label, 1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl-methanethiosulfonate (MTSL), to several of the purified single cysteine mutants. Then, we evaluated protein dynamics during the methylation reaction by monitoring mobility of the MTSL-labelled residue upon addition of SAM. Overall, our structural and biochemical analyses will be used to ascertain the structural dynamics associated with SAM binding of this unique methyltransferase.

Additionally, we were able to incorporate His-Ste14 in nanodiscs. Nanodiscs mimic the membrane of a cell and are a more native-like environment than detergent micelles or liposomes. Since nanodiscs are conducive to many biophysical techniques, unlike detergents, we have begun preliminary studies to better understand the structure of Ste14. Techniques we have begun to pursue are negative stain electron microscopy (EM), single particle cryo-electron microscopy (cryo-EM), and X-ray crystallography.

Finally, we previously showed Ste14 functions as a dimer or higher order oligomer. Ste14 is comprised of six transmembrane (TM) domains in which TM1 contains a putative dimerization motif, $G_{31}XXXG_{35}XXXG_{39}$, where G is a glycine amino acid residue and X is a subset of hydrophobic amino acids. Using cysteine-scanning mutagenesis, we characterized TM1 cysteine mutants for their effects on protein expression, activity, and stability. We determined residues S27, Y28, L30, G31, G35, and G39 are critical for maintaining activity levels. Additionally, residues M25, T26, Y28, F41, P42, and Q43 were found to form strong dimers through the addition of sulfhydryl specific cross-linkers and immunoblot analysis. Recently, the purification of dimeric

Ste14 from aggregated protein components via size exclusion chromatography (SEC) was improved for further experimentation. The purified, monodispersed, His-Ste14 underwent size exclusion chromatography (SEC), multi-angle light scattering (MALS) and small-angle X-ray scattering (SAXS) to confirm the dimerization state of Ste14. Together, we have used many biochemical and biophysical methods to gain insight about the structure, function, and mechanism of Ste14. Ultimately, our studies will be utilized to design more potent therapeutics to minimize K-Ras signaling in cancer cells.

CHAPTER 1. INTRODUCTION: A BIOLOGICAL UNDERSTANDING OF THE INTEGRAL MEMBRANE PROTEIN, ISOPRENYLCYSTEINE CARBOXYL METHYLTRANSFERASE (ICMT)

1.1 CaaX Proteins and Processing

Many proteins undergo covalent modifications after translation, which significantly changes the chemical characteristics of the protein.¹ Common examples of such modifications include: lipidation, phosphorylation, ubiquitination, and methylation. CaaX protein processing is a series of sequential post-translational modifications that are essential for CaaX protein function because the hydrophobic modifications properly localize the proteins to the plasma or organelle membranes.²⁻⁷ CaaX proteins have four specific C-terminal amino acids designated as a CaaX motif, where the “C” is a cysteine, “a” are aliphatic residues, and “X” represents one of several amino acids.^{3,8} Although these proteins were first discovered in fungi, the family of known CaaX proteins expanded to include the γ subunits of heterotrimeric GTP-binding proteins (heterotrimeric G proteins), small GTPases, nuclear lamins, yeast **a**-factor mating pheromone, protein kinases, phosphatases, and phosphodiesterases.⁹⁻¹³ CaaX proteins are ubiquitously expressed and involved in regulatory functions such as cell growth, migration, protein trafficking, and many other biological processes.^{4,14-17} Currently, there are approximately 300 known eukaryotic CaaX proteins, where over the past decades, the Ras superfamily is the most well studied and best characterized.^{4,14,18} The Ras superfamily consists of the Rho, Ras, and Rab protein families.

1.1.1 S-Isoprenylation

The first step of CaaX processing is the irreversible addition of a 15-carbon farnesyl group or 20-carbon geranylgeranyl group to the cysteine residue of the CaaX protein with a thioester linkage (Figure 1.1).¹⁹⁻²¹ The X amino acid dictates the type of lipidation that occurs. If the X residue is alanine, serine, methionine, or glutamine, a farnesyl isoprenoid is usually attached by the cytosolic enzyme farnesyltransferase (FTase), whereas if the X residue is leucine or phenylalanine, a geranylgeranyl isoprenoid is attached by geranylgeranyl transferase I (GGTase-I).^{4,19-25} Due to their CaaX box sequences, Ras GTPases, lamin B, and Rheb proteins are farnesylated, while Rac and Rho proteins are geranylgeranylated.^{4,19} The addition of a lipid group

is necessary for the subsequent two steps of CaaX processing to occur, leading to proper localization of the substrate to membranes in the cell.^{4,14,18,23,26-28} The third type of prenylation enzyme is Rab geranylgeranyl transferase (RabGGTase) or geranylgeranyltransferase II (GGTase-II), which recognizes and lipidates non-CaaX sequences. Although less defined, GGTase-II has been shown to modify proteins containing two cystine residues (XXCC, XCCX, CCXX, CCXXX, or XCXC).²⁹ Very recently, a fourth prenylation enzyme was discovered, geranylgeranyltransferase III (GGTase-III).³⁰ This enzyme was found to geranylgeranilate FBXL2, an F-box protein necessary for the assembly of an active SCF ubiquitin ligase complex. Interestingly, FBXL2 contains a C-terminal CaaX box, previously believed to be prenylated by GGTase-I. Currently, a new topic of interest is to determine additional ubiquitin ligases or other enzymes being modified by GGTase-III.

The CaaX prenylation proteins, FTase and GGTase-I have similar structural components composed of α/β -heterodimers. Both enzymes share a similar α -subunit, which is encoded by the same gene, *FNTA*.²⁸ However, these enzymes differ more significantly in their β -subunits with only 30% sequence similarity.³¹⁻³³ The β -subunits are necessary for lipid and substrate binding, which accounts for their preference of specific protein substrates. The ordered reaction mechanism is comprised of two steps: FTase or GGTase binds to a farnesyl pyrophosphate (FPP) or geranylgeranyl pyrophosphate (GGPP), respectively, followed by binding to the CaaX protein.^{28,31,34} The prenylation reaction is fast, but the release of the lipidated CaaX protein is the rate-limiting step of the reaction.^{35,36} Additionally, Zn^{2+} is needed to coordinate the cysteine thiol within the CaaX motif of the substrate, which lowers the pK_a of the thiol. The lower pK_a causes the metal ion to activate the substrate for nucleophilic attack.^{20,24,28,37,38} X-ray crystal structures of various FTases and GGTases have been determined, which have been essential for characterization of these enzymes.^{39,40}

1.1.2 Endoproteolysis

Lipidation of CaaX proteins allows for their localization to the membrane of the endoplasmic reticulum (ER) due to the addition of a prenyl group that favors the hydrophobic environment of the ER. At the ER membrane, the second post-translational modification occurs as an endoproteolysis reaction (Figure 1.1).⁴¹ In mammals, the Ras converting enzyme 1 (Rce1) is the integral membrane protein that cleaves the terminal -aaX peptides from CaaX proteins.

Previous researchers have suggested that the cleavage of the -aaX motif is necessary for intracellular trafficking and coordination of the CaaX protein with the methyltransferase responsible for the last modification of CaaX processing.^{15,42-44}

In yeast, there are two functionally redundant homologs of this proteolytic protein, Rce1 and Ste24. Ste24 is located both in the ER and in the nuclear envelope.^{41,45,46} In addition to CaaX protein proteolysis, Ste24 performs an additional cleavage to the substrate **a**-factor, a yeast peptide mating pheromone. This second cleavage step removes the first seven amino acids from the N-terminus of the protein and after several more steps, releases the pheromone from the cell and signals for the initiation of mating with an **α**-cell.^{45,47-49} Due to its short peptide sequence, **a**-factor is commonly used as a model substrate for studying the isoprenylation pathway.⁵⁰

The mammalian homolog of Ste24, ZMPSTE24, plays an important role in prelamin A maturation to lamin A.^{41,49,51} Nuclear lamins are intermediate filament proteins that form mesh-like structures underlying the nuclear envelope, providing an environment near the nuclear membrane for organization of transcription factors, nuclear pore complexes, and heterochromatin.^{52,53} ZMPSTE24 cleaves the -aaX motif from CaaX proteins, but also performs an additional N-terminal cleavage on prelamin A maturation, which results in the removal of the whole lipidated C-terminal tail. Interestingly, ZMPSTE24 and Ste24 can each perform both cleavage reactions on **a**-factor and prelamin A.^{54,55} Human ZMPSTE24 shows 36% identity and 62% similarity to Ste24 of *Saccharomyces cerevisiae*.⁵⁶

Interestingly, X-ray crystallography of Ste24 and ZMPSTE24 identified a unique structure in both enzymes, that has not been seen in any other membrane protein; a large, hollow, barrel-shaped chamber formed within seven transmembrane α -helices.^{57,58} This rare chamber-like structure is predicted to be large enough to accommodate a ~10kDa protein or ~450 molecules of water.⁵⁶ First, it is predicted that the CaaX motif of the substrate enters this chamber for the first cleavage step. After endoproteolysis, the substrate may completely or partially leave the chamber using one of the side portals for methylation by Icmt. Later, the CaaX protein must position itself in the chamber again for the second cleavage reaction. Currently, this two-step mechanism remains unclear, particularly how ZMPSTE24 binds the substrate for each cleavage step.

1.1.3 Methyl Esterification

The third step of the CaaX processing pathway is the methyl esterification of the exposed α -carboxylate group of the C-terminal cysteine residue following endoproteolysis (Figure 1.1).¹⁹ Regardless of the type of prenylation (farnesylation or geranylgeranylation), the methyl esterification is performed by the ER integral membrane protein, isoprenylcysteine carboxyl methyltransferase (Icmt), before the substrate is localized to the plasma or organelle membrane for proper signaling.^{4,7,14,18,59-61} Currently, there are several proposed roles of the methylation modification. First, the methyl group is believed to cap the hydrophilic, negatively charged amino acid moiety, making the CaaX protein more hydrophobic and more favorable to associate with the membrane.⁶² Second, the addition of the methyl group serves as a recognition site for protein-protein interactions in the CaaX proteins signal transduction pathways, including examples like lamin B interactions at the nuclear envelop and K-Ras with microtubules.^{15,62-67} Finally, methylation may protect the protein from any additional proteolytic cleavages at the C-terminus by covering the charge of the exposed cysteine carboxylic acid.⁶⁸

The mating pheromones from the jelly fungi, *Tremella brasiliensis* and *Tremella mesenteric*, were the first proteins shown to contain an isoprenylated and methylated C-terminal cysteine residue.^{12,13,70} Later, Michaelis *et al.* identified the first isoprenylcysteine carboxyl methyltransferase gene in budding yeast, *STE14*.⁴⁴ Clarke *et al.* found that membrane extracts of sterile *S. cerevisiae*, which lacked *STE14*, also lacked farnesyl cysteine C-terminal carboxyl methyltransferase activity typically found in yeast **a**-cells and **α** -cells.⁴³ Interestingly, other sterile strains of *S. cerevisiae* lacking *STE6* and *STE16* retained carboxyl methyltransferase activity.⁴³ By performing genetic analysis on *Δste14* yeast strains, the first Icmt protein was identified in *Saccharomyces cerevisiae* as Ste14.⁴³ Later, Icmt was found to be ubiquitously expressed in eukaryotes and all mammalian Icmt homologs contain highly conserved sequences.^{43,44,59} The Icmt family is of interest due to the discover of its substrates, more notably H-Ras, which were found to be methylated and farnesylated in its mature form.⁶⁹⁻⁷³ To date, Ste14 is used as a model system to better characterize human Icmt (hIcmt) and the Icmt family of proteins.

1.2 Diseases Associated with CaaX Proteins

CaaX proteins are involved in many essential signaling pathways including cell proliferation, trafficking, gene expression, and differentiation.^{4,9-11,14,15} Since many of these

processes promote cell growth and division, mutations in CaaX proteins and alterations in their regulation can result in many different diseases and disorders including malaria, Alzheimer's disease, various cancers, Hutchinson-Gilford Progeria Syndrome (HGPS) and amyotrophic lateral sclerosis.⁷⁴⁻⁸⁰ The Ras super family is the most studied family of CaaX proteins because of their roles in all human cancers.^{3,81-84} Specific single point mutations in *RAS* genes causes the growth of carcinomas and metastasis.^{85,86}⁸¹ Additionally, if CaaX proteins do not undergo all three post-translational modifications of CaaX processing, they accumulate in the cell, which can lead to other health problems like kidney disorders.⁸⁷

Ras proteins (K-Ras4A, K-Ras4B, H-Ras, and N-Ras) are categorized as small GTPases that are involved in transmitting signals within cells.⁸⁸⁻⁹⁰ All isoforms of Ras are ubiquitously expressed across eukaryotic species and partake in many signaling pathways, including the mitogen-activated protein kinase pathway (MAPK) and phosphoinositide-3 kinase (PI3K) pathways.^{6,91,92} Active Ras is bound to GTP, but can be deactivated by GTPase-activating proteins (GAPs). GAPs hydrolyze GTP to GDP, causing a conformational change in Ras preventing downstream signaling. Ras can be reactivated by, guanine nucleotide exchange factors (GEFs), which remove GDP and replace it with GTP in this same region of the protein, known as switch 1 and switch 2.^{88,93-96}

Approximately 30% of all human cancers contain a mutation in the *RAS* gene, which results in reduced hydrolysis of GTP to GDP.^{81,84,97-100} Since GTP is not removed, Ras becomes constitutively active and leads to oncogenic transformation. In all Ras isoforms, 99.2% of oncogenic mutations occur at codons 12, 13, and 61.^{81,101} *KRAS* was shown to be the most frequently mutated isoform; it is mutated in 95% of pancreatic ductal adenocarcinomas (PDAC) and 30% of all cancers.^{81,102,103} In pancreatic cancer, the most common *KRAS* mutations occurs at codon 12, where the glycine residue is typically mutated to valine or aspartic acid, G12V and G12D, respectively.¹⁰⁴ Currently there is no cure for pancreatic cancer and the 5-year survival rate is ~7%. As a result of this poor prognosis, CaaX processing proteins are being utilized as potential targets to inhibit oncogenic K-Ras by mislocalizing the protein and preventing downstream effector signaling. A full analysis of CaaX protein inhibitors will be discussed in length in section

1.3 Characterization of Human Icmt and the Yeast Homolog, Ste14

1.3.1 General Characteristics, Localization, and Two-Dimensional Topology

The Icmt family of enzymes is the only known class of membrane-associated methyltransferases. These enzymes do not contain any of the main four motifs characteristically found in soluble methyltransferases. As a result, the Icmt family members form the class VI methyltransferases.⁶³ Interestingly, in a primarily hydrophobic environment, Icmt has the ability to accommodate both a hydrophilic co-factor, *S*-adenosyl-L-methionine (SAM), as a methyl donor and a lipophilic methyl acceptor.

To date, one of the most characterized members of the Icmt family is its founding member Ste14, found in *S. cerevisiae*. Ste14 was functionally purified and reconstituted in detergent for further characterization, while hIcmt has yet to be purified actively.¹⁰⁵ hIcmt and Ste14 share 63% similarity and 41% identity, mostly within the C-termini of the proteins. Most importantly, when expressed in yeast in a Δ *ste14* deletion strain, hIcmt was able to resume the sterile phenotype and methyltransferase activity, suggesting a similar catalytic mechanism of the two enzymes.⁵⁹

Ste14, the 26kDa integral membrane protein, was found localized to the ER membrane with six transmembrane (TM) spans and one unique helical hairpin between TM5 and TM6. (Figure 1.2).^{60,68,106-108} Determination of the Ste14 sequence and production of a hydropathy plot helped determine the membrane spanning regions of Ste14.^{59,60,106} A protease protection assay was first used to show Ste14 with its N- and C- termini facing the cytosolic side of the ER, along with most of the soluble loops.⁶⁰ Similar topology results were confirmed with hIcmt in live cells with a fluorescent probe selectively placed on the N- and C- termini.¹⁰⁸

To better understand the function, mechanism, and structure of Ste14, the amino acid sequence was compared to known methyltransferases and matched most closely with the C-terminal end of two phospholipid methyltransferases, Pem1p and Pem2p.¹⁰⁶ Shortly after, Romano *et al.* determined that the sequence of Ste14 was highly conserved with ten gene products including three sterol reductases and seven open reading frames of unknown bacterial enzymes.¹⁰⁶ In 2011, Court *et al.* aligned the sequences of Icmt from fifteen different species and identified 42 amino acids that were identical and 35 amino acids that had similar characteristics (Figure 1.2).⁵⁶ The C-termini had the most conserved residues among the fifteen species and confirmed the potential importance of several conserved motifs. Amongst all these species was the consensus motif, RHPxY-30 hydrophobic-EE, where 30 hydrophobic residues were surrounded by regions A

(RHPxY) and B (EE) (Figure 1.2).¹⁰⁶ In Ste14, the RHPxY-30 hydrophobic-EE motif spans from amino acid 161 to 217. This hydrophobic stretch, called the helical hairpin motif, was shown to be embedded in the ER membrane and was necessary for methyltransferase activity (Figure 1.2).¹⁰⁶ To understand the role of the amino acids in this tripartite conserved motif, site-directed mutagenesis was performed on the two glutamic acid residues in region B, E213D and E214D. Both residues were essential for enzyme function.¹⁰⁶ Other residues were mutated in this study and lost significant activity compared to wild-type Ste14 as well, including G31E, L81F, G132R, S148F, P173L, and L217S. Logue and Olsen performed an alanine scan of all the conserved residues in Ste14 via site-directed mutagenesis.^{109,110} Within the RHPxY-30 hydrophobic-EE motif, 31 mutants completely lost activity.¹⁰⁹ Additionally, the forty-four conserved, cytosolic residues that comprise Ste14 were mutated to alanine in crude membranes and resulted in 29 mutants retaining 10% or less activity compared to wild-type.¹¹⁰ Mutant activities were not rescued when the conserved residues were mutated to amino acids similar to wild-type. For example, tyrosine residues were also mutated to phenylalanine. These studies indicated that the conserved residues amongst Icmts are important for proper function.

Further mutational analysis of Icmts from *Methanosarcina acetivorans* (Ma-Icmt), *Saccharomyces cerevisiae* (Ste14), and *Anopheles gambiae* (Ag-Icmt) confirmed the findings from Romano *et al.* and identified other residues and motifs conserved for proper structure and function.^{63,64,105,106,111} Sequence alignment of these Icmt orthologs was performed to accurately compare residues and motifs of interest (Figure 1.3). One conserved motif was the repeating GXXXG of the GXXXG-like motif found near the N-terminal end of all Icmts. The “G” represents a glycine residue or another small amino acid like serine or alanine, and an “X” motif which can be one of variety of amino acids, but are typically hydrophobic in order to span the membrane, including valine, leucine, and isoleucine.¹¹²⁻¹¹⁴ The GXXXG motif has been shown to be important in the dimerization of several membrane proteins including glycophorin A (GpA), ATP binding cassette subfamily G member 2 (ABCG2), yeast ATP synthase, human carbonic anhydrase (CA), *Helicobacter pylori* vacuolating toxin, and yeast α -factor receptor.¹¹⁵⁻¹²² Interestingly, the GXXXG motif is a highly conserved pattern of amino acids and commonly found in integral membrane proteins that form dimers.¹¹² In the GXXXG repeating motif, the spacing of the glycine residues allows the amino acids to be positioned on the same face of the helix and provide a flat surface to permit the close packing of TM1 of two Icmt

proteins.^{113,114,123} Additionally, it has been shown that the residues surrounding each glycine, such as serine, threonine, and tyrosine, are also important for dimerization due to the formation of potential intramolecular hydrogen bonds.¹¹⁶ Moreover, leucine, valine, and isoleucine residues can form van der Waals interactions, which can further stabilize the protein dimer.¹¹⁶

Ste14 contains a GXXXG-like sequence, ²⁷SYILGILLGIFVG³⁹ (Figure 1.2).^{68,107,111,124} Therefore, Ste14 was proposed to be a naturally occurring dimer or higher order oligomer.¹²⁴ Griggs *et al.* supported this hypothesis by showing that Ste14 was a dimer or oligomer in the presence of the crosslinker, bis(sulfosuccinimidyl)suberate (BS³), in crude membranes and when purified.¹²⁴ Additionally, Griggs *et al.* performed dominant-negative activity assay experiments with wildtype Ste14 and two mutants with negligible enzymatic activity.¹²⁴ When purified and reconstituted His-Ste14 in lipids was expressed with untagged Ste14, methyltransferase activity levels were higher than activity with reconstituted His-Ste14 alone.¹²⁴ When catalytic mutants were purified and reconstituted in lipids, little to no activity was detected. Yet, activity levels increased when untagged Ste14 and the mutants were expressed together.¹²⁴ Together, these experiments provided evidence that the interaction between Ste14 monomers were important for methyltransferase function.¹²⁴ Less is known about the GXXXG-like motif in TM3 of hIcmt, ⁷⁰ACFLGFVFG⁷⁸, but Diver *et al.* (2018) predicted it stabilizes a similar GXXXG-like motif in TM1.¹⁰⁸ In chapter 5 of this dissertation, we further characterize the GXXXG-like motif in TM1 and how it affects the dimerization of Ste14.

1.3.2 Hypothesized Kinetic Mechanism of Icmt

The kinetics of the Icmt family are still relatively unknown. Early studies identified the Icmt catalyzes the methylation of CaaX proteins through a bi-bi reaction mechanism, not shared by any other type of methyltransferase (Figure 1.4).^{125,126} It was proposed that SAM binds to Icmt first, causing a large conformational change, followed the prenylated substrate as the methyl acceptor. After the reaction occurs, the fully processed CaaX protein is released, followed by the reaction product S-adenosyl-L-homocysteine (SAH). Both products display product inhibition with Icmt.^{125,127} Shi *et al.* first performed these kinetic studies with rod outer segment (ROS) isoprenylated protein methyltransferase.¹²⁵ Subsequently, Baron *et al.* supported these claims with kinetic experiments using recombinant hIcmt.¹²⁶ However, photoreactive substrate analogs were

slow to bind to both hIcmt and Ste14 in reactions lacking SAM.¹²⁸ Therefore, more kinetic analyses need to be performed on Icmts and their unique methylation reaction in the membrane.

1.4 Structural Analysis of Co-factor and Substrate Binding Sites of Icmt

1.4.1 Icmts are Putative Metalloenzymes with Zinc Dependence

All mammalian Icmts were demonstrated to be metalloenzymes.^{105,129,130} Metalloenzymes are biologically necessary for cell survival by playing critical roles in oxygen transport, signal transduction, and many other processes. Purified Ste14 was shown to be inhibited by 1, 10-phenanthroline and 2-carboxyl-2'-hydroxy-5'-sulfoformazylbenzene (Zincon), metal chelators specific for inhibition of only zinc (Zn^{2+}) or zinc and copper, respectively.^{105,129} Interestingly, Ste14 was unaffected by water-soluble chelators, which suggested that the more hydrophilic chelators could not enter the hydrophobic environment of the protein and ultimately prevented access to the metal ion.¹⁰⁵ When hydrophobic cholesterol groups were added to the metal chelator nitrolotri-acetic acid (NTA) via a lysine linker (cholesterol-Lys-NTA), the inhibitor became more potent, further suggesting that the metal ion in Icmt was buried in a hydrophobic region of the protein and was essential for correct structure assembly or catalytic activity.¹³⁰ Although, the three endogenous cysteine residues (C99, C121, C126) were predicted to interact with Zn^{2+} , the activity and expression of Ste14 was unaffected *in vitro* and *in vivo*, when the cysteine residues were mutated to alanine or serine.¹³⁰ Therefore, the cysteine residues were not necessary for metal coordination or methyltransferase activity. To date, residues interacting with the divalent zinc cation remains unknown.

1.4.2 First Icmt Structural Model in Eukaryotic Organism

In 2018, Diver *et al.* solved the crystal structure for the red flour beetle Icmt, *Tribolium castaneum* (Figure 1.5B, PDB: 5VG9).¹³² This structure contains eight transmembrane domains and a short cytosolic cap helix. Importantly, this was the first eukaryotic Icmt to be crystallized, with ~59% identity and ~76% similarity with hIcmt. The 2.4Å structure was co-crystallized with a monobody inhibitor based on a randomized fibronectin protein domain and *N*-acetyl-*S*-geranylgeranyl-*L*-cysteine (AGGC). They proposed that the ER membrane dips at the C-terminus of the protein, to allow for SAM and the CaaX substrate to enter the active site, while remaining in their energetically preferred environment. Diver *et al.* (2018) proposed a methylation

mechanism based on their model.¹³² First, the negatively charged, nucleophilic carboxy group of the exposed, isoprenylated cysteine coordinates with two arginine residues R173 and R246, located on TM6 and TM8, respectively. The connector between TM6 and TM7 and the C-terminal cap region are predicted to act as a hinge region controlled by two residues, G181 and S193, to mediate SAM binding and release. The hinge closes around SAM positioning the co-factor for the methylation reaction and formed several hydrogen bonds to stabilize the transition state at F184, Y212, and Q185. Yet, these bonds are considered controversial because the bond with Y212 was 5Å long (typical a hydrogen bond does not exceed 3.5Å).

Diver *et al.* (2018) also predicted the prenylated CaaX protein enters a deep cavity formed by TM4, TM5, TM7, and TM8. This cavity connects the co-factor binding site to the membrane where the substrate could enter, bringing the methyl donor and acceptor together. Additionally, the cavity is ~22Å long by 6Å wide, making Icmt large enough to accommodate farnesylated, geranylgeranylated, or double geranylgeranylated substrates. The access tunnel is lined with aromatic residues, allowing the hydrophobic prenylation group to enter the beetle Icmt. When the aromatic residues were mutated activity levels decreased compared to WT, suggesting an importance in substrate binding. However, in the X-ray crystal structure, the electron density in the cavity was at a very low resolution and could not be identified with certainty.

In addition, zinc was never modeled in the beetle Icmt structure, even though it has been suggested to be a metalloenzyme (discussed above). Several mutations were also made to the protein in order to be crystallized, and the roles of the residues of the lipidated substrate that interact with Icmt were not clearly elucidated. More importantly, there were no insights about the dynamics of Icmt, necessary to fully understand this unique methylation reaction.

1.4.3 First Icmt Structural Model in Prokaryotic Organism

Yang *et al.* crystallized the first Icmt from an Archaea, *Methanosarcina acetivorans* Icmt (*Ma-Icmt*) (Figure 1.5A, PDB: 4A2N).⁶³ *Ma-Icmt* is an ortholog of the eukaryotic Icmt and showed high sequence conservation at the C-terminus. In addition, *Ma-Icmt* can methylate the substrate, *N*-acetyl-S-farnesyl-L-cysteine (AFC). The crystal structure of *Ma-Icmt* in complex with SAH was determined at resolution of 3.4Å. A tunnel for prenylated substrates to enter the catalytic site of *Ma-Icmt* was proposed, but the Icmt ortholog was never crystallized with a substrate or substrate mimic. Since proteins are not prenylated in prokaryotes, *Ma-Icmt* has no known substrates. Yet,

this first structural model was helpful for predicting important regions for catalytic and structural motifs.

Ma-Icmt shares only 14% sequence homology with *hIcmt* and only ~40% with *Ste14*. Amino acid similarity was only found within the C-terminal end of *Ma-Icmt* with other enzymes in the *Icmt* family. Interestingly, this region of the protein is proposed to interact with the co-factor, SAM, since SAH co-crystallized in this region of the enzyme. When *Ma-Icmt* was superimposed on *Ste14* using the program SWISS-MODEL Version 8.05, only C-terminal residues 115-239 of *Ste14* were aligned, which includes the region A, region B, and the helix-turn-helix motif described in section 1.3.1.

The 3D model, along with the analysis of *Ma-Icmt* by Yang *et al.* has helped predict the roles of various *Ste14* residues believed to be necessary for SAM binding.⁶³ In *Ste14*, Y167, S174, and Y175, interact with the carboxylate group of SAH. In addition, H172 and E213 formed hydrogen bonds with the carboxylate group and amino group of SAH, respectively. The models identified H159 as interacting with the adenine ring of SAH, whereas Y225 and Y228 interacts with water molecules that mediate contact with the adenine ring. Logue found all of these residues lost activity when mutated to alanine.¹⁰⁹ *Ste14* residues A143, L161, V162, and L217 are conserved, nonpolar residues, also predicted to orient the adenine ring of SAM correctly in the binding site. Other residues are predicted to be important for stabilizing the protein: (1) L161 stabilized F221 with hydrophobic interactions and (2) R171 and E214 both stabilized H172 to interact with SAH. Site-directed mutagenesis of these residues led to enzyme inactivation. P173 was involved in stabilizing TM5 of *Ste14* and is conserved in all *Icmt* enzymes. When this proline was mutated to alanine, all activity was lost, suggesting this residue may be necessary for the structural integrity of the protein. Although these studies aided in the understanding of SAM/SAH binding, the *Ma-Icmt* structure only has conserved homology with eukaryotic *Icm*ts in the C-terminus and therefore, less was known about the N-terminal end of the protein.

1.4.4 Biochemical Analysis of *Icmt* Residues Important for Binding

To obtain a better understanding of important regions in the enzyme, Diver *et al.* (2014) analyzed the amino acid sequence of the *Icmt* ortholog from *Anopheles gambiae*, *Ag-Icmt*.¹¹¹ By utilizing the EMBL-EBI EMBOSS Stretcher Global Alignment tool, *Ag-Icmt* had ~41% percent identity and 82% similarity with *Ste14* and ~49% identity and ~67% similarity to *hIcmt*. *Ag-Icmt*

has higher sequence conservation with hIcmt since both of these Icmts have eight transmembrane segments. Diver *et al.* (2014) performed methyltransferase activity assays utilizing biotin-S-farnesyl-L-cysteine (BFC) as the methyl acceptor, SAM as the methyl donor, and GFP labelled Icmt enzymes (wild-type and mutants).¹¹¹ Through site-directed mutagenesis of 137 residues (out of 283) of Ag-Icmt and measurement of methyltransferase activity retained, 62 mutants observed reduced or undetectable catalytic activity. Specifically, all the residues in the proposed SAM binding motif identified by Yang *et al.*, lost significant activity when mutated to alanine, except for the residue equivalent to the Ste14 H159.⁶³ Also, this model supports that E213 and E214 in Ste14 form hydrogen bonds with SAH. The negative charge of each of these residues are predicted to stabilize histidine 172. The hydrophobic residues equivalent to Ste14 residues F80, I81, and L82 in Ag-Icmt were identified to be necessary for wild-type activity levels, however, their role is not fully understood.¹¹¹

Additionally, this study wanted to identify which residues are important for CaaX protein binding since there is higher N-terminal conservation of residues in Ag-Icmt to hIcmt. The percent activity compared to wild-type Ag-Icmt mutants significantly decreased in the following mutants: L62A, Y64A, F179A, T185A, Q186A, W203A, F205A (Ste14 amino acid sites).¹¹¹ All these residues reside in transmembrane domains of Icmt, are highly conserved amongst Icmt family members. Diver *et al.* (2014) concluded that these mutations had lower affinity than wild type for SAM due to their increased K_m values. Activity assays were performed on these mutants with increasing amounts of SAM and BFC as a substrate. Inhibition by BFC was observed. Thus, the isoprenylated substrate, but not SAM, inhibits these mutants, suggesting that inhibition is independent of SAM concentration and the mutations are preventing substrate binding. They also hypothesize that R173 and R246 in Ste14 aid in substrate binding due to their loss of methyltransferase activity when mutated. Although Diver *et al.* (2014) had many insights on SAM and substrate binding in Ag-Icmt, their percent activity levels for several of their mutants were significantly different than those found by Funk (2017) in Ste14. For example, Diver *et al.* (2014) found T185A, Q186A and W203A had methyltransferase activity levels between 5-30% of wild-type, yet Funk (2017) found the activity levels to be 68%, 90%, and 98%, respectively.¹³¹ Due to the conflicting views about the activity of Icmt mutants, new biochemical techniques are needed in order to reveal more information on the exact amino acids participating in substrate recognition.

1.4.5 Substrate Specificity

It is important to note that the Icmt enzyme can recognize a vast number of CaaX protein substrates, which can be farnesylated or geranylgeranylated.^{62,133,134} Several studies have probed for substrate specificity of the Icmt family. First in 1990, Stephenson *et al.* developed and used the peptide LARYKC coordinated with a farnesyl or geranylgeranyl group to study the substrate specificity of Icmt in rat liver microsomal membranes.¹³⁴ The LARYKC sequence originated from the C-terminus of Ras from *Drosophila*.¹³⁴ The peptide alone did not act as a substrate, but both farnesylated and geranylgeranylated substrates could be modified. Additionally, *S*-farnesyl-LARYKC was found to inhibit methylation of bovine ROS proteins with nearly identical K_m and V_{max} values.¹³⁴ Thus, their studies confirmed that prenylation must occur for Icmt to recognize the CaaX protein.

As discussed earlier, Stephenson *et al.* and Clarke *et al.* demonstrated that the CaaX protein substrate must be prenylated for recognition by Icmt.^{19,134,135} Specifically, they found that only the first isoprene group was necessary for this process.^{19,134,135} Later, Tan *et al.* demonstrated that a sulfoxide derivative of AFC was a poor substrate of Icmt. Thus, the sulfur heteroatom of the substrate was necessary for the methylation reaction with Icmt to proceed.¹³⁵ Another poor substrate of Icmt was an AFC homocysteine analog, suggesting the distance between the carboxyl and sulfur functional groups dictated substrates of Icmt. This hypothesis was supported by activity assays performed with *S*-farnesyl-thioacetic acid (FTA) and *S*-farnesyl thiopropionic acid (FTP), where FTP retained activity, but FTA did not.¹³⁵ Both molecules are identical except FTA has a shorter carbon backbone between its sulfur and carboxyl group.

Further investigations revealed that bovine, rat, mosquito and yeast Icmts methylate *N*-acetyl-*S*-farnesyl-*L*-cysteine (AFC) and *N*-acetyl-*S*-geranylgeranyl-*L*-cysteine (AGGC), as well as derivatives of each including biotin-*S*-farnesyl-*L*-cysteine (BFC).^{44,127,135-137} AFC and AGGC are identical molecules that only differ in the number of isoprene groups in their lipid portion, three or four isoprene groups, respectively. Several non-peptide-based substrates were also identified including FTP and *S*-geranylgeranyl thiopropionic acid (GGTP) -triazole compounds.^{105,133} Anderson *et al.* examined substrate specificity of Ste14 in crude membranes, as well as, purified protein in detergent followed by functional reconstitution in lipids.¹⁰⁵ Ste14 recognized both AFC and AGGC with equal turnover rates and demonstrated that Ste14 does not prefer one substrate

over the other. In terms of stereochemistry, the D isomer of AFC is a slightly poorer substrate than the L-isomer, with a K_m of $\sim 70\mu\text{M}$ opposed to $\sim 20\mu\text{M}$.

Additionally, Funk (2017) observed differences in substrate specificity of single Ste14 mutants. Activity assays were performed with 76 conserved His-Ste14 residues mutated to alanine or glycine in crude membrane extracts. It was shown that N-terminal and TM1 alanine mutants preferred the AGGC substrate over AFC, unlike wild-type Ste14.¹³¹ Other regions of Ste14 also had varying preferences of AFC and AGGC. Thus, further studies are needed to understand the substrate, isoprenoid binding pocket of this unique enzyme.

1.5 Inhibition of CaaX Proteins

1.5.1 Challenges of Ras Inhibition

Mutations in CaaX proteins can lead to deregulation or changes in cell proliferation, cell migration, protein trafficking, and assembly of the nuclear envelope.^{4,9-11,14,15} CaaX proteins have been linked to many diseases, but the most studied is the relationship between the Ras superfamily and cancer.⁸⁰ One of the most severe types of cancer involves mutations in the *KRAS* isoform; it is mutated in 95% of pancreatic ductal adenocarcinomas (PDAC) and 30% of all cancers.^{81,102,103} PDAC has been predicted to surpass breast and colorectal cancer to become the second most common cause of cancer-related deaths, partially due to the failure to create successful drugs and treatments after years of drug design.¹³⁸ Importantly, K-Ras knockdown in mouse models with lung adenocarcinoma or pancreatic cancer caused tumor regression and knockdown in humans showed various results.¹³⁹⁻¹⁴¹ When drug design began, Ras was the primary target. However, despite decades of effort, Ras has yet to be inhibited by a clinical drug and was even termed “undruggable” for a period of time. Ras has been an elusive target due to its lack of a druggable binding pocket, picomolar affinities to GTP/GDP, and the inability to target specific Ras isoforms.^{142,143} Currently, there are five main strategies to inhibit Ras: (1) Disruption of the regulator or effector interactions, (2) Inhibit association to the plasma membrane, (3) Inhibition of downstream effectors, (4) Inhibition of synthetic lethal interactions, and (5) Inhibition of metabolism.¹⁴⁴ A plethora of research has focused on the second strategy by inhibiting the three CaaX processing enzymes as potential targets for development of Ras inhibitors for the treatment of Ras driven cancers.

1.5.2 Inhibition of Protein S-Isoprenylation

Prenylation was the first CaaX processing step to be targeted to inhibit oncogenic Ras. The addition of a prenyl group is essential for plasma membrane association of the protein, thus the development of farnesyl transferase inhibitors (FTIs) could be used to delocalize oncogenic Ras and prevent its continuous function.¹⁴ Development of FTIs were determined through structure-based analysis of the enzyme and several high throughput methods.²⁵ Pre-clinical trials appeared promising *in vitro* and in *in vivo* mouse models.¹⁴⁵⁻¹⁴⁷ Thus, some FTIs made it to human clinical trials, most notably tipifarnib and lonafarnib.¹⁴⁸ However, FTIs had no survival advantages for patients with advanced solid cancers or acute myeloid leukemia (AML).¹⁴⁹⁻¹⁵² Overall, FTIs were not successful because these inhibitors only affected the isoform H-Ras, while N-Ras, K-Ras4A, and K-Ras4B were still prenylated by GGTase-I.^{153,154} When geranylgeranylated, the CaaX proteins still undergo proteolysis and methylation by the subsequent CaaX processing enzymes and leads to normal signaling.^{153,154}

Very recently there has been a resurgence to design FTIs and geranylgeranyl transferase inhibitors (GGTIs). In 2019, Sun *et al.* designed a guanidyl-based bivalent peptidomimetic, FTI-227, that disrupts the interaction between K-Ras and FTase or GGTase to prevent its subsequent interactions with c-Raf.¹⁵⁵ Although these drugs look promising for treatment, they have not undergone clinical trials. Therefore, due to the strong need for a chemotherapeutic to target K-Ras, the next two steps of CaaX processing were proposed as alternative chemotherapeutic targets.

1.5.3 Inhibition of Endoproteolysis

Rce1 was the next enzyme in CaaX processing to gain higher interest as a drug target for Ras-driven cancers. First, Bergo *et al.* found that deletion of Rce1 would be beneficial to inhibit Ras signaling.¹⁵⁶ Mice with a *Cre*-mediated excision of *Rce1* (*Rce1*^{Δ/Δ}) eliminated the endoproteolytic cleavage and methylation of Ras. The *Rce1* deletion within cells caused partial mislocalization of K-Ras and H-Ras proteins.¹⁵⁶ Additionally, the conditional lack of *Rce1* expression in skin carcinoma cells that expressed activated H-Ras retarded cell growth, limited Ras-induced transformation, and sensitized tumor cell to FTIs like lonafarnib.¹⁵⁶ Although promising, the Rce1 deficiency had only modest effects on cancerous Ras transformation. Many drugs were designed to inhibit Rce1 based on its substrates.^{157,158} Unfortunately, these inhibitors had high enzymatic degradation and low cell permeability, making them poor choices for clinical

trials. To fix this problem, Schlitzer *et al.* designed non-peptidic, non-prenylic inhibitors.¹⁵⁹ To date, the most potent compound is an 8-hydroxyquinoline-based inhibitor, NSC1011.¹⁶⁰ However, as Rce1 became more characterized, it was found to cause severe detrimental effects when inhibited. Bergo *et al.* found that Rce1 deficiency led to the acceleration of K-Ras induced myeloproliferative disease in oncogenic K-Ras driven hematopoietic cells.¹⁶¹ Additionally, inhibition of this enzyme was shown to promote the development of lethal cardiomyopathy in mice.¹⁶² Finally, studies found that the accumulation of isoprenylated proteins that could not complete CaaX processing, lead to symptoms similar to progeroid disorders, including muscle weakness and bone fractures.^{55,163,164}

1.5.4 Inhibition of Methyl Esterification

Finally, Icmt was examined as a chemotherapeutic target to inhibit the CaaX processing pathway. Since mammalian genomes only encode for one Icmt family member and Icmt is in a unique class of methyltransferases that lacks homology to other protein methyltransferases, it is a more specific drug target than Rce1.⁵⁹ Bergo *et al.* used genetic mouse studies to show that Icmt contributes to K-Ras localization to the plasma membrane and function.^{165,166} In addition, inactivation of Icmt in fibroblasts with constitutively active K-Ras demonstrated a decrease in cellular growth, K-Ras and B-Raf oncogenic transformation, and methylation of Ras.¹⁶⁵ Additionally, Lau *et al.* created Icmt loss-of-function isogenic cell lines through CRISPR/Cas9 genome editing.¹⁶⁷ Ras-transformed human mammary epithelial cells (HME1) and human cancer cell lines with naturally occurring mutant K-Ras (MiaPaca-2 and MDA-MB-231) were used to demonstrate the role of Icmt in Ras-driven cancers both *in vivo* and *in vitro*.¹⁶⁷ Lack of Icmt abolished tumor initiation of all isoforms of mutant Ras, not just K-Ras4B, and tumor maintenance in all cell lines. However, there have been some conflicting results. In 2008, Wahlstrom *et al.* found that inactivating Icmt in mice ameliorates K-Ras induced myeloproliferative disease and Court *et al.* found deficiencies in Icmt exacerbates K-Ras-driven pancreatic neoplasia via Notch suppression.^{168,169}

Icmt inhibitors fall into three classes: (1) substrate analogs, (2) natural products, and (3) synthetic compounds. The reaction product, SAH, was first examined as an inhibitor. It inhibited the methyl esterification reaction with Icmt, but showed no specificity, since they act as non-selective inhibitors of all methyltransferases, limiting their potential as a chemotherapeutic.¹²⁵

Then, some of the first synthesized inhibitors mimicked CaaX protein substrates including the minimal substrates AFC, AGGC, FTA, and GGTA, along with variations of these substrate backbones.¹⁷⁰⁻¹⁷² Specifically, the Hrycyna and Gibbs laboratories designed the AFC minimal substrates that all contained a lipid group, sulfur heteroatom, cysteine amino acid, and N-terminus of a peptide.^{62,171} Other groups determined additional characteristics necessary for these inhibitors. Clarke *et al.* found that the isoprene length does not change the strength of inhibition, as long as the first isoprene group was present.¹⁹ Tan *et al.* showed that the replacement of the sulfur moiety with oxygen or selenium made for poor substrates.¹³⁵ All these inhibitors competed with prenylated Ras in the binding pocket and decreased activation of the MAPK pathway, some of which showed effects in the low micromolar range.^{172,173} Common modifications made to these compounds were the addition of bulky hydrophobic groups and modified amide regions.

Second, in 2008, natural extracts were determined as inhibitors via high-throughput screenings. Most prominently, two derivatives from a Verongida sponge, *Pseudoceratina* sp., known as spermatinamine and aplysamine 6 were found to have IC₅₀ values of 1.9 μ M and 14 μ M respectively.^{174,175} Additionally, three natural compounds from the plant *Hovea parvicalyx* inhibited Icmt with IC₅₀ values ranging between 17 μ M and 30 μ M.¹⁷⁶

Finally, nonpeptide-based, synthetic derivatives have been designed to inhibit Icmt as well. Most notably is cysmethynil, which was discovered by high-throughput screening from a library of ~10,000 compounds. Cysmethynil has an IC₅₀ value of 2.4 μ M and was shown to decrease Ras localization and down-stream signaling.^{177,178} Although it demonstrated good antitumor efficacy in several human cell lines, this Icmt inhibitor had low water solubility and poor affinity for the plasma membrane.¹⁷⁹ Thus, it was a poor candidate for clinical use. Modifications were later made to cysmethynil, which led to a more potent and bioavailable inhibitor, compound 8.12. In PC3 prostate and HepG2 liver cancer cells, this compound induced autophagy, cell death, attenuated tumor growth, and abolished anchorage-independent colony formation.¹⁷³ Moreover, there has been a resurgence of the compound Salirasib (farnesyl thiosalicylic acid, FTS) as an inhibitor of Icmt. It has undergone clinical trials for pancreatic, colon, lung, and breast cancer, and more recently, a new antimalarial agent.^{180,181} Overall, rational design of better inhibitors has been hampered by a lack of ICMT structural information. Therefore, the focus of this dissertation is to better understand the structural and conformational dynamics of the yeast isoprenylcysteine carboxyl methyltransferase, Ste14.

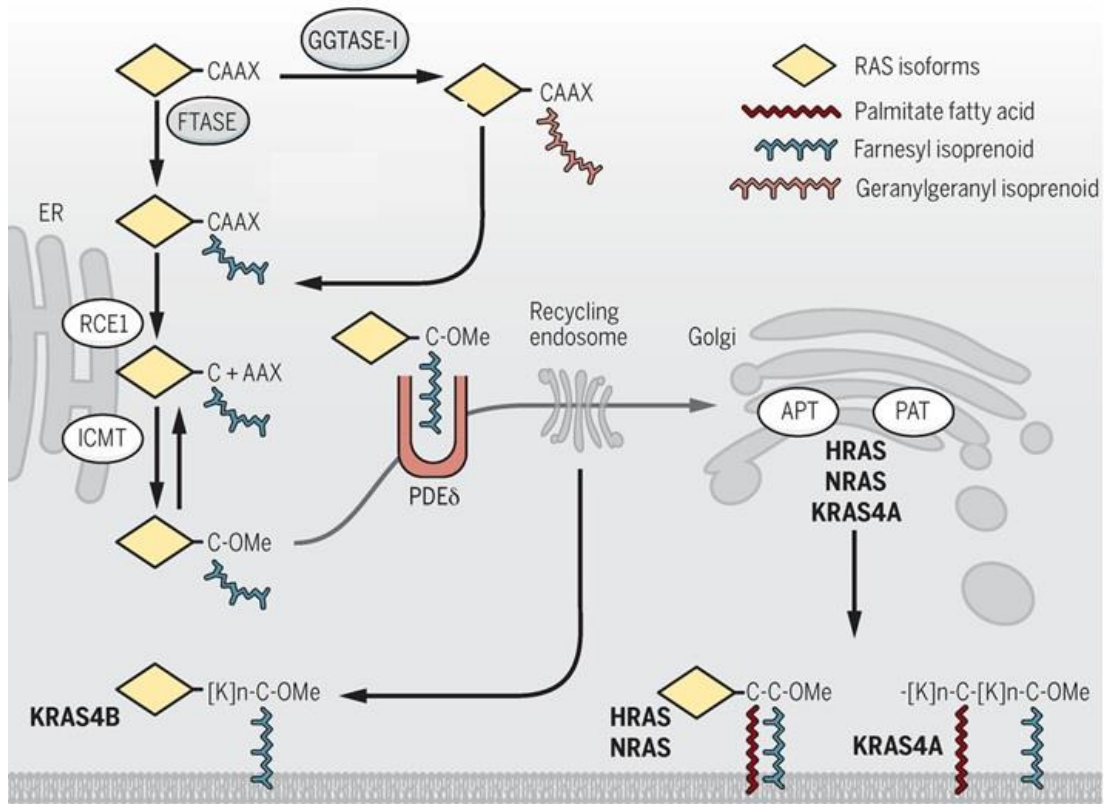


Figure 1.1 The CaaX processing pathway.¹⁴⁴

The C-terminal CaaX motif is recognized by the farnesyltransferase (FTase) or geranylgeranyltransferase (GGTase-I) in the cytosol of the cell. These enzymes covalently add a farnesyl or geranylgeranyl isoprenoid group to the cysteine residue, increasing the hydrophobicity of the protein and facilitating association to the membrane of the endoplasmic reticulum (ER). Next, the Ras converting enzyme 1 (Rce1) or Ste24 cleaves off the three C-terminal amino acids, which is followed by the methyl esterification of the exposed carboxyl group of the cysteine residue by isoprenylcysteine carboxyl methyltransferase (Icmt). These three modifications are necessary for the CaaX protein to become associated to the plasma or nuclear membrane and have proper downstream signaling. The Ras superfamily undergoes CaaX processing, but H-Ras, N-Ras, and K-Ras4A undergo an additional palmitoylation step at a different cysteine residue, executed by the Golgi-associated palmitoyl acyltransferases (PATs). Acyl-protein thioesterases (APTs) can reverse this process in the Golgi. K-Ras4B skips this step due to its polybasic sequence (the beginning ten residues from the end of the C-terminus). PDE δ recognizes the farnesyl or geranylgeranyl group of Ras and transports the protein from the cytosol to the recycling endosome or the Golgi, from which it is shuttled to the plasma membrane. Me, Methyl. From: Papke, B. & Der, C. J. Drugging RAS: Know the enemy. *Science* **355**, 1158-1163, doi:10.1126/science.aam7622 (2017). Reprinted with permission from AAAS.

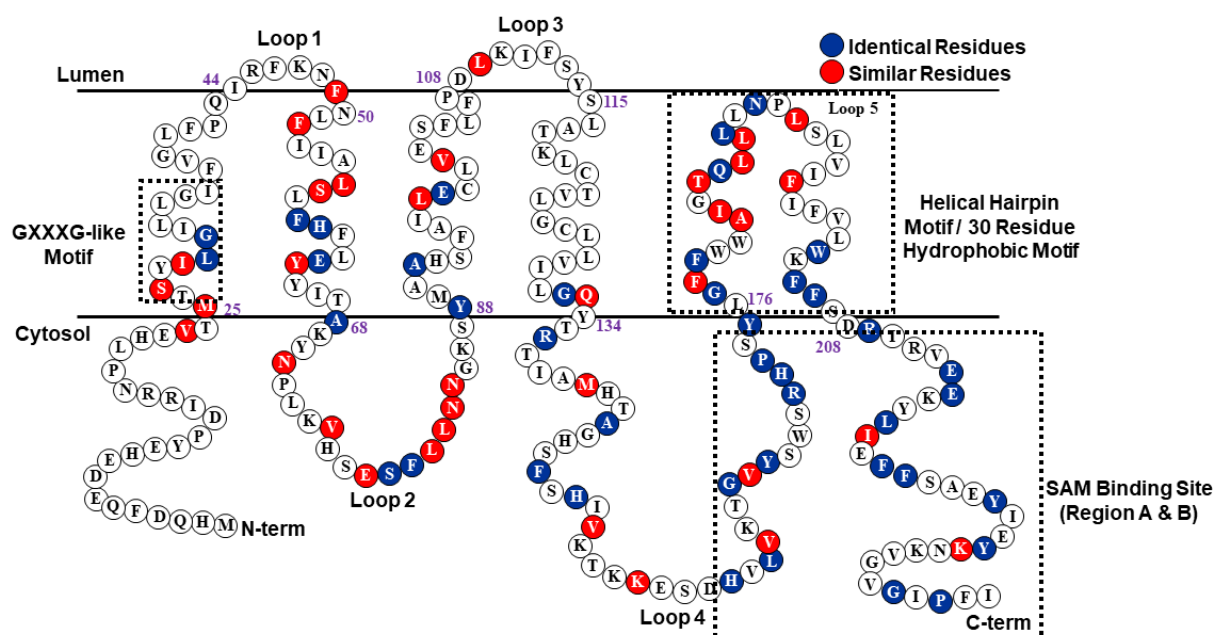


Figure 1.2 Topology map of *Saccharomyces cerevisiae* Icmt (Ste14).

Similar residues across fifteen species of Icmts are in red and identical residues are blue. The GXXXG-like motif and tripartite motif (comprised of regions A and B surrounding 30 hydrophobic residues/ the helical hairpin motif) are identified by dashed boxes. The most conserved region is the C-terminus of Ste14, where SAM is predicted to bind. The topology map of Ste14 was previously mapped by Romano *et al.* and showed the enzyme has six transmembrane domains with the N- and C- termini exposed to the cytosol.¹⁰⁶ Adapted from: Court, H., Hahne, K., Philips, M. R. & Hrycyna, C. A. Biochemical and Biological Functions of Isoprenylcysteine Carboxyl Methyltransferase. *Enzymes, Vol 30: Protein Prenylation, Pt B* **30**, 71-90, doi:10.1016/b978-0-12-415922-8.00004-5 (2011).

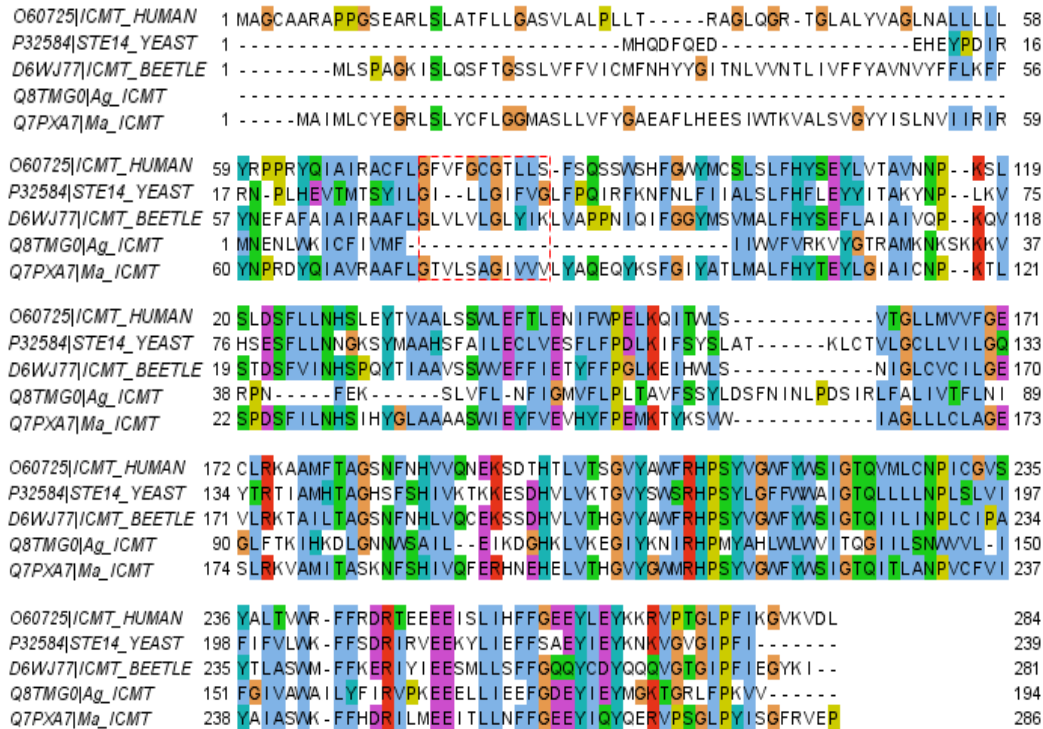


Figure 1.3 Sequence alignment of Icmt orthologs.

Sequence alignment of human Icmt (hIcmt; 284 residues), *Saccharomyces cerevisiae* Icmt (Ste14; 239 residues), *Tribolium castaneum* Icmt (beetle, 281 residues), *Anopheles gambiae* Icmt (Ag-Icmt; 283 residues), and the prokaryotic enzyme from *Methanosarcina acetivorans* (Ma-Icmt; 194 residues). UniProt accession numbers for the Icmt ortholog sequences are as follows: O60725, P32584, D6WJ77, Q8TMG0, and Q7PXA7, respectively. Alignment was made using Clustal Omega and Jalview 2.10.5. Colored boxes denote residues with similar properties: glycine is orange, acidic amino acids are purple, proline is yellow, nonpolar amino acids are light blue, histidine and tyrosine are turquoise, basic amino acids are red, and remaining polar, neutrally charged amino acids are green. The red dashed box surrounds the GXXXG motif in Ste14.

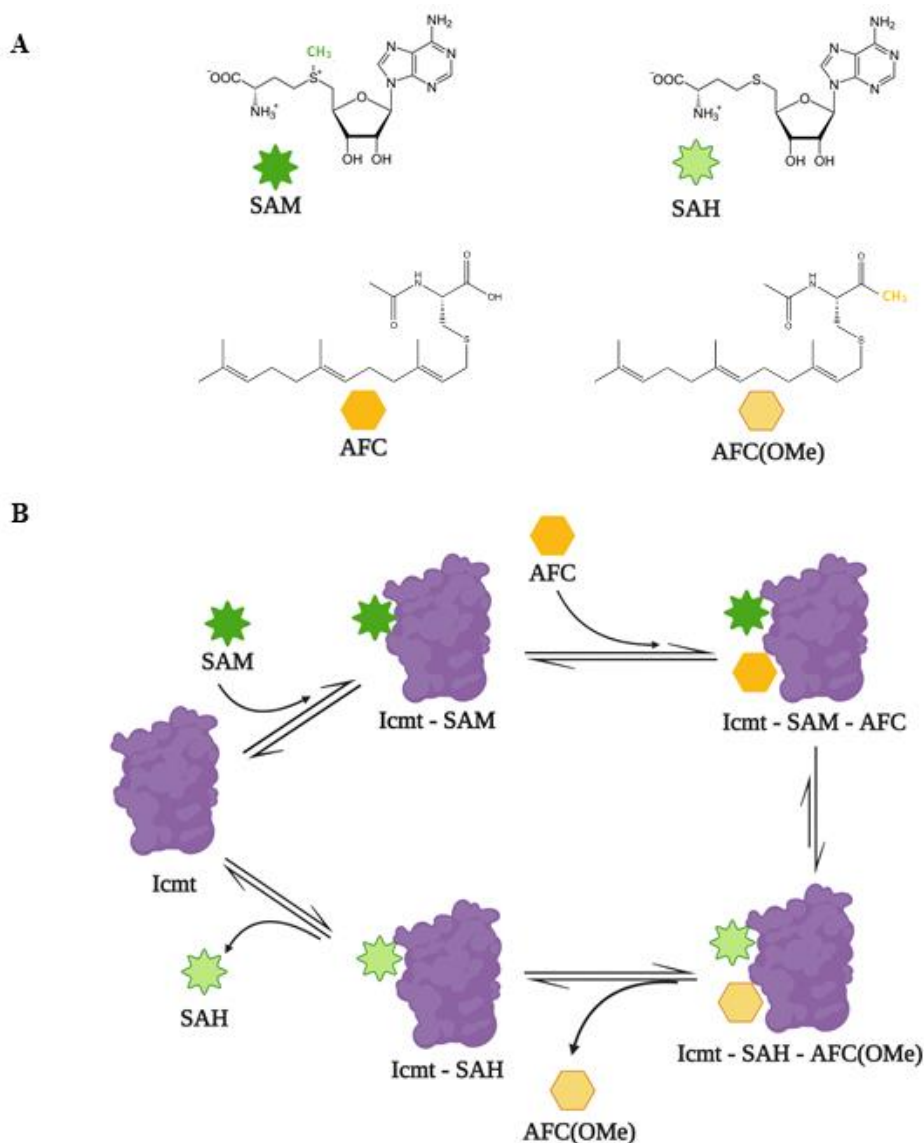


Figure 1.4 Sequentially ordered bi-bi reaction mechanism unique to the Icmt family. In the methyl esterification reaction, Icmt uses the hydrophilic co-factor, *S*-adenosyl-L-methionine (SAM), as a methyl donor and a lipophilic CaaX protein as the methyl acceptor.^{125,126} (A) SAM is depicted as a dark green star, *S*-adenosyl-L-homocysteine (SAH) is depicted as a light green star, AFC is depicted as a darker yellow hexagon, and AFC(OMe) is depicted as a light yellow hexagon. (B) First, SAM binds to Icmt followed by the prenylated CaaX protein. After the reaction proceeds, the fully processed CaaX protein is released, followed by the reaction product *S*-adenosyl-L-homocysteine (SAH). *N*-acetyl-*S*-farnesyl-*L*-cysteine (AFC) is a minimal substrate and represents a lipophilic CaaX protein in this kinetic mechanism. Created by biorender.com.

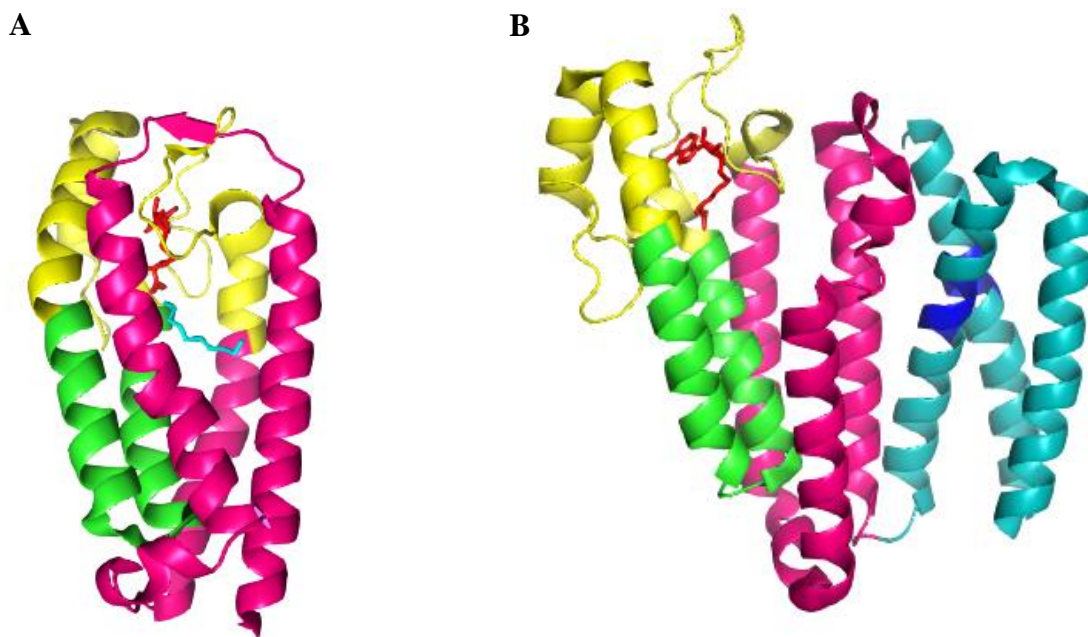


Figure 1.5 Crystal structures of *Methanosarcina acetivorans* Icmt (Ma-Icmt) and *Tribolium castaneum* Icmt (beetle).^{63,132}

PDB accession numbers 4A2N and 5VG9, respectively. Both Icmt structures include the co-factor reaction product, *S*-adenosyl-L-homocysteine (SAH) binding site and predict an access tunnel for the hydrophobic substrate. (A) In *Ma*-Icmt, the tripartite motif (comprised of regions A and B surrounding 30 hydrophobic residues/ the helical hairpin motif) where SAM/SAH are predicted to interact with Icmt is yellow and the helix-turn-helix motif is green. The remainder of the protein is pink. SAH shown in red sticks. Additionally, the crystal structure proposes a lipid substrate access tunnel recognizing prenylated proteins within the hydrophobic pocket. A cyan, 10 carbon lipid (in stick formation) demonstrates where the prenylated CaaX protein substrate is predicted enter the active site. (B) In the beetle Icmt crystal structure, the first three alpha helices are teal, the core helices are pink, the GXXXG motifs are dark blue, and the helix-turn-helix motif that acts as part of the substrate access tunnel is green. Additionally, the predicted SAM/SAH binding site which includes the hinge region and region A and B of the tripartite motif, is yellow. SAH is shown as red sticks. Images created with The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.

1.6 References

- 1 Alberts, B., Johnson, A. & Lewis, J. *Molecular Biology of the Cell*. 4th edn, (Garland Science, 2002).
- 2 Wang, M. & Casey, P. J. Protein prenylation: unique fats make their mark on biology. *Nature Reviews Molecular Cell Biology* **17**, 110-122, doi:10.1038/nrm.2015.11 (2016).
- 3 Winter-Vann, A. M. & Casey, P. J. Opinion - Post-prenylation-processing enzymes as new targets in oncogenesis. *Nature Reviews Cancer* **5**, 405-412, doi:10.1038/nrc1612 (2005).
- 4 Zhang, F. L. & Casey, P. J. Protein prenylation: Molecular mechanisms and functional consequences. *Annual Review of Biochemistry* **65**, 241-269, doi:10.1146/annurev.bi.65.070196.001325 (1996).
- 5 Young, S. G., Ambroziak, P., Kim, E. & Clarke, S. in *The Enzymes: Protein lipidation* Vol. 21 (Academic Press, 2000).
- 6 Hancock, J. F. Ras proteins: Different signals from different locations. *Nature Reviews Molecular Cell Biology* **4**, 373-384, doi:10.1038/nrm1105 (2003).
- 7 Casey, P. J. Protein lipidation in cell signaling. *Science* **268**, 221-225, doi:10.1126/science.7716512 (1995).
- 8 Reid, T. S., Terry, K. L., Casey, P. J. & Beese, L. S. Crystallographic analysis of CaaX prenyltransferases complexed with substrates defines rules of protein substrate selectivity. *Journal of Molecular Biology* **343**, 417-433, doi:10.1016/j.jmb.2004.08.056 (2004).
- 9 Collins, S. P., Reoma, J. L., Gamm, D. M. & Uhler, M. D. LKB1, a novel serine/threonine protein kinase and potential tumour suppressor, is phosphorylated by cAMP-dependent protein kinase (PKA) and prenylated in vivo. *Biochemical Journal* **345**, 673-680, doi:10.1042/0264-6021:3450673 (2000).
- 10 Schwindinger, W. F. & Robishaw, J. D. Heterotrimeric G-protein beta gamma-dimers in growth and differentiation. *Oncogene* **20**, 1653-1660, doi:10.1038/sj.onc.1204181 (2001).
- 11 Cates, C. A. *et al.* Prenylation of oncogenic human PTPCAAX protein tyrosine phosphatases. *Cancer Letters* **110**, 49-55, doi:10.1016/s0304-3835(96)04459-x (1996).
- 12 Sakagami, Y., Yoshida, M., Isogai, A. & Suzuki, A. Peptidal sex-hormones inducing conjugation tube formation in compatible mating-type cells of tremella-mesenterica. *Science* **212**, 1525-1527, doi:10.1126/science.212.4502.1525 (1981).
- 13 Ishibashi, Y., Sakagami, Y., Isogai, A. & Suzuki, A. Structures of tremmerogen-a-9291-i and tremmerogen-a-9291-vii - peptidyl sex-hormones of tremella-brasiliensis. *Biochemistry* **23**, 1399-1404, doi:10.1021/bi00302a010 (1984).

- 14 Casey, P. J. & Seabra, M. C. Protein prenyltransferases. *Journal of Biological Chemistry* **271**, 5289-5292, doi:10.1074/jbc.271.10.5289 (1996).
- 15 Hrycyna, C. A. & Clarke, S. Modification of eukaryotic signaling proteins by c-terminal methylation reactions. *Pharmacology & Therapeutics* **59**, 281-300, doi:10.1016/0163-7258(93)90071-k (1993).
- 16 Wilson, K. L., Zastrow, M. S. & Lee, K. K. Lamins and disease: Insights into nuclear infrastructure. *Cell* **104**, 647-650, doi:10.1016/s0092-8674(01)00261-6 (2001).
- 17 Hutchison, C. J. Lamins: Building blocks or regulators of gene expression? *Nature Reviews Molecular Cell Biology* **3**, 848-858, doi:10.1038/nrm950 (2002).
- 18 Gao, J. H., Liao, J. & Yang, G. Y. CAAX-box protein, prenylation process and carcinogenesis. *American Journal of Translational Research* **1**, 312-325 (2009).
- 19 Clarke, S. Protein isoprenylation and methylation at carboxyl-terminal cysteine residues. *Annual Review of Biochemistry* **61**, 355-386, doi:10.1146/annurev.bi.61.070192.002035 (1992).
- 20 Park, H. W. Crystal structure of protein farnesyltransferase at 2.25 Å resolution *Science* **276**, 21-21 (1997).
- 21 Chen, W. J., Andres, D. A., Goldstein, J. L. & Brown, M. S. Cloning and expression of a cDNA encoding the alpha subunit of rat p21ras protein farnesyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 11368-11372, doi:10.1073/pnas.88.24.11368 (1991).
- 22 Xu, N. *et al.* Protein prenylation and human diseases: a balance of protein farnesylation and geranylgeranylation. *Science China-Life Sciences* **58**, 328-335, doi:10.1007/s11427-015-4836-1 (2015).
- 23 Sinensky, M. Functional aspects of polyisoprenoid protein substituents: roles in protein-protein interaction and trafficking. *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids* **1529**, 203-209, doi:10.1016/s1388-1981(00)00149-9 (2000).
- 24 Moores, S. L. *et al.* Sequence dependence of protein isoprenylation. *Journal of Biological Chemistry* **266**, 14603-14610 (1991).
- 25 Berndt, N., Hamilton, A. D. & Sefti, S. M. Targeting protein prenylation for cancer therapy. *Nature Reviews Cancer* **11**, 775-791, doi:10.1038/nrc3151 (2011).

- 26 Willumsen, B. M., Norris, K., Papageorge, A. G., Hubbert, N. L. & Lowy, D. R. Harvey murine sarcoma-virus p21 ras protein - biological and biochemical significance of the cysteine nearest the carboxy terminus. *Embo Journal* **3**, 2581-2585, doi:10.1002/j.1460-2075.1984.tb02177.x (1984).
- 27 Willumsen, B. M., Christensen, A., Hubbert, N. L., Papageorge, A. G. & Lowy, D. R. The p21 ras c-terminus is required for transformation and membrane association. *Nature* **310**, 583-586, doi:10.1038/310583a0 (1984).
- 28 Maurer-Stroh, S., Washietl, S. & Eisenhaber, F. Protein prenyltransferases: Anchor size, pseudogenes and parasites. *Biological Chemistry* **384**, 977-989, doi:10.1515/bc.2003.110 (2003).
- 29 Zhang, H., Seabra, M. C. & Deisenhofer, J. Crystal structure of Rab geranylgeranyltransferase at 2.0Å resolution. *Structure with Folding & Design* **8**, 241-251, doi:10.1016/s0969-2126(00)00102-7 (2000).
- 30 Kuchay, S. *et al.* GGTase3 is a newly identified geranylgeranyltransferase targeting a ubiquitin ligase. *Nature Structural & Molecular Biology* **26**, 628-+, doi:10.1038/s41594-019-0249-3 (2019).
- 31 Reiss, Y., Goldstein, J. L., Seabra, M. C., Casey, P. J. & Brown, M. S. Inhibition of purified p21ras farnesyl - protein transferase by cys-aax tetrapeptides. *Cell* **62**, 81-88, doi:10.1016/0092-8674(90)90242-7 (1990).
- 32 Zhang, F. L. *et al.* cDNA cloning and expression of rat and human protein geranylgeranyltransferase type-I. *Journal of Biological Chemistry* **269**, 3175-3180 (1994).
- 33 Mazur, P. *et al.* Purification of geranylgeranyltransferase I from *Candida albicans* and cloning of the CaRAM2 and CaCDC43 genes encoding its subunits. *Microbiology-Uk* **145**, 1123-1135, doi:10.1099/13500872-145-5-1123 (1999).
- 34 Yokoyama, K., McGeady, P. & Gelb, M. H. Mammalian protein geranylgeranyltransferase-I - substrate-specificity, kinetic mechanism, metal requirements, and affinity labeling. *Biochemistry* **34**, 1344-1354, doi:10.1021/bi00004a029 (1995).
- 35 Pais, J. E., Bowers, K. E., Stoddard, A. K. & Fierke, C. A. A continuous fluorescent assay for protein prenyltransferases measuring diphosphate release. *Analytical Biochemistry* **345**, 302-311, doi:10.1016/j.ab.2005.07.040 (2005).
- 36 Tschantz, W. R., Furfine, E. S. & Casey, P. J. Substrate binding is required for release of product from mammalian protein farnesyltransferase. *Journal of Biological Chemistry* **272**, 9989-9993 (1997).
- 37 Caplin, B. E., Hettich, L. A. & Marshall, M. S. Substrate characterization of the *Saccharomyces-cerevisiae* protein farnesyltransferase and type-I protein geranylgeranyltransferase. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **1205**, 39-48, doi:10.1016/0167-4838(94)90089-2 (1994).

- 38 Long, S. B., Casey, P. J. & Beese, L. S. The basis for K-Ras4B binding specificity to protein farnesyl-transferase revealed by 2 angstrom resolution ternary complex structures. *Structure with Folding & Design* **8**, 209-222, doi:10.1016/s0969-2126(00)00096-4 (2000).
- 39 Basso, A. D., Kirschmeier, P. & Bishop, W. R. Farnesyl transferase inhibitors. *Journal of Lipid Research* **47**, 15-31, doi:10.1194/jlr.R500012-JLR200 (2006).
- 40 Nguyen, U. T. T., Goody, R. S. & Alexandrov, K. Understanding and exploiting protein prenyltransferases. *Chembiochem* **11**, 1194-1201, doi:10.1002/cbic.200900727 (2010).
- 41 Schmidt, W. K., Tam, A., Fujimura-Kamada, K. & Michaelis, S. Endoplasmic reticulum membrane localization of Rcelp and Ste24, yeast proteases involved in carboxyl-terminal CAAX protein processing and amino-terminal a-factor cleavage. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 11175-11180, doi:10.1073/pnas.95.19.11175 (1998).
- 42 Hrycyna, C. A. & Clarke, S. Maturation of isoprenylated proteins in *Saccharomyces cerevisiae* - multiple activities catalyze the cleavage of the 3 carboxyl-terminal amino-acids from farnesylated substrates in vitro. *Journal of Biological Chemistry* **267**, 10457-10464 (1992).
- 43 Hrycyna, C. A. & Clarke, S. Farnesyl cysteine C-terminal methyltransferase activity is dependent upon the Ste14 gene-product in *saccharomyces-cerevisiae*. *Molecular and Cellular Biology* **10**, 5071-5076, doi:10.1128/mcb.10.10.5071 (1990).
- 44 Hrycyna, C. A., Sapperstein, S. K., Clarke, S. & Michaelis, S. The *Saccharomyces cerevisiae* Ste14 gene encodes a methyltransferase that mediates c-terminal methylation of a-factor and ras proteins. *Embo Journal* **10**, 1699-1709, doi:10.1002/j.1460-2075.1991.tb07694.x (1991).
- 45 Barrowman, J., Wiley, P. A., Hudon-Miller, S. E., Hrycyna, C. A. & Michaelis, S. Human ZMPSTE24 disease mutations: residual proteolytic activity correlates with disease severity. *Human Molecular Genetics* **21**, 4084-4093, doi:10.1093/hmg/ddc233 (2012).
- 46 Boyartchuk, V. L., Ashby, M. N. & Rine, J. Modulation of Ras and a-factor function by carboxyl-terminal proteolysis. *Science* **275**, 1796-1800, doi:10.1126/science.275.5307.1796 (1997).
- 47 Tam, A. *et al.* Dual roles for Ste24 in yeast a-factor maturation: NH₂-terminal proteolysis and COOH-terminal CAAX processing. *Journal of Cell Biology* **142**, 635-649, doi:10.1083/jcb.142.3.635 (1998).
- 48 Tam, A., Schmidt, W. K. & Michaelis, S. The multispanning membrane protein Ste24 catalyzes CAAX proteolysis and NH₂-terminal processing of the yeast a-factor precursor. *Journal of Biological Chemistry* **276**, 46798-46806, doi:10.1074/jbc.M106150200 (2001).

- 49 FujimuraKamada, K., Nouvet, F. J. & Michaelis, S. A novel membrane-associated metalloprotease, Ste24, is required for the first step of NH₂-terminal processing of the yeast a-factor precursor. *Journal of Cell Biology* **136**, 271-285, doi:10.1083/jcb.136.2.271 (1997).
- 50 Sinensky, M. Recent advances in the study of prenylated proteins. *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids* **1484**, 93-106, doi:10.1016/s1388-1981(00)00009-3 (2000).
- 51 Trueblood, C. E. *et al.* The CaaX proteases, Afc1p and Rce1p, have overlapping but distinct substrate specificities. *Molecular and Cellular Biology* **20**, 4381-4392, doi:10.1128/mcb.20.12.4381-4392.2000 (2000).
- 52 Dittmer, T. & Misteli, T. The lamin protein family. *Genome Biology* **12**, doi:10.1186/gb-2011-12-5-222 (2011).
- 53 Worman, H. J., Fong, L. G., Muchir, A. & Young, S. G. Laminopathies and the long strange trip from basic cell biology to therapy. *Journal of Clinical Investigation* **119**, 1825-1836, doi:10.1172/jci37679 (2009).
- 54 Spear, E. D. *et al.* ZMPSTE24 missense mutations that cause progeroid diseases decrease prelamin A cleavage activity and/or protein stability. *Disease Models & Mechanisms* **11**, doi:10.1242/dmm.033670 (2018).
- 55 Barrowman, J. & Michaelis, S. ZMPSTE24, an integral membrane zinc metalloprotease with a connection to progeroid disorders. *Biological Chemistry* **390**, 761-773, doi:10.1515/bc.2009.080 (2009).
- 56 Court, H., Hahne, K., Philips, M. R. & Hrycyna, C. A. Biochemical and Biological Functions of Isoprenylcysteine Carboxyl Methyltransferase. *Enzymes, Vol 30: Protein Prenylation, Pt B* **30**, 71-90, doi:10.1016/b978-0-12-415922-8.00004-5 (2011).
- 57 Quigley, A. *et al.* The Structural Basis of ZMPSTE24-Dependent Laminopathies. *Science* **339**, 1604-1607, doi:10.1126/science.1231513 (2013).
- 58 Pryor, E. E. *et al.* Structure of the Integral Membrane Protein CAAX Protease Ste24. *Science* **339**, 1600-+, doi:10.1126/science.1232048 (2013).
- 59 Dai, Q. *et al.* Mammalian prenylcysteine carboxyl methyltransferase is in the endoplasmic reticulum. *Journal of Biological Chemistry* **273**, 15030-15034, doi:10.1074/jbc.273.24.15030 (1998).
- 60 Romano, J. D., Schmidt, W. K. & Michaelis, S. The *Saccharomyces cerevisiae* prenylcysteine carboxyl methyltransferase Ste14 is in the endoplasmic reticulum membrane. *Molecular Biology of the Cell* **9**, 2231-2247, doi:10.1091/mbc.9.8.2231 (1998).
- 61 Casey, P. J. Protein lipidation in cell signaling. *Science* **268**, 221-225, doi:10.1126/science.7716512 (1995).

- 62 Anderson, J. L., Henriksen, B. S., Gibbs, R. A. & Hrycyna, C. A. The isoprenoid substrate specificity of isoprenylcysteine carboxylmethyltransferase - Development of novel inhibitors. *Journal of Biological Chemistry* **280**, 29454-29461, doi:10.1074/jbc.M504982200 (2005).
- 63 Yang, J. *et al.* Mechanism of Isoprenylcysteine Carboxyl Methylation from the Crystal Structure of the Integral Membrane Methyltransferase ICMT. *Molecular Cell* **44**, 997-1004, doi:10.1016/j.molcel.2011.10.020 (2011).
- 64 Anderson, J. L. & Hrycyna, C. A. Structure and Function of Isoprenylcysteine Carboxylmethyltransferase (Icmt): A Key Enzyme in CaaX Processing. *Enzymes: Protein Methyltransferases, Vol 24* **24**, 245-272 (2006).
- 65 Chen, Z., Otto, J. C., Bergo, M. O., Young, S. G. & Casey, P. J. The C-terminal polylysine region and methylation of K-Ras are critical for the interaction between K-Ras and microtubules. *Journal of Biological Chemistry* **275**, 41251-41257, doi:10.1074/jbc.M006687200 (2000).
- 66 Maske, C. P. *et al.* A carboxyl-terminal interaction of lamin B1 is dependent on the CAAAX endoprotease Rce1 and carboxymethylation. *Journal of Cell Biology* **162**, 1223-1232, doi:10.1083/jcb.200303113 (2003).
- 67 Cushman, I. & Casey, P. J. Role of Isoprenylcysteine Carboxylmethyltransferase-catalyzed Methylation in Rho Function and Migration. *Journal of Biological Chemistry* **284**, 27964-27973, doi:10.1074/jbc.M109.025296 (2009).
- 68 Sapperstein, S., Berkower, C. & Michaelis, S. Nucleotide-sequence of the yeast Ste14 gene, which encodes farnesylcysteine carboxyl methyltransferase, and demonstration of its essential role in a-factor export. *Molecular and Cellular Biology* **14**, 1438-1449, doi:10.1128/mcb.14.2.1438 (1994).
- 69 Clarke, S., Vogel, J. P., Deschenes, R. J. & Stock, J. Posttranslational modification of the Ha-ras oncogene protein: evidence for a third class of protein carboxyl methyltransferases. **85**, 4643-4647 (1988).
- 70 Anderegg, R. J., Betz, R., Carr, S. A., Crabb, J. W. & Duntze, W. Structure of *Saccharomyces cerevisiae* mating hormone a-factor - Identification of s-farnesyl cysteine as a structural component. *Journal of Biological Chemistry* **263**, 18236-18240 (1988).
- 71 Marr, R. S., Blair, L. C. & Thorner, J. *Saccharomyces-cerevisiae*-Ste14 gene is required for COOH-terminal methylation of a-factor mating pheromone. *Journal of Biological Chemistry* **265**, 20057-20060 (1990).
- 72 Pillinger, M. H., Volker, C., Stock, J. B., Weissmann, G. & Philips, M. R. Characterization of a plasma membrane-associated prenylcysteine-directed alpha-carboxyl methyltransferase in human neutrophils. *Journal of Biological Chemistry* **269**, 1486-1492 (1994).

- 73 Stotzler, D., Betz, R. & Duntze, W. Stimulation of yeast mating hormone-activity by synthetic oligopeptides. *Journal of Bacteriology* **132**, 28-35 (1977).
- 74 Suazo, K. F., Schaber, C., Palsuledesai, C. C., John, A. R. O. & Distefano, M. D. Global proteomic analysis of prenylated proteins in *Plasmodium falciparum* using an alkyne-modified isoprenoid analogue. *Scientific Reports* **6**, doi:10.1038/srep38615 (2016).
- 75 Butler, K. V., Bohn, K., Hrycyna, C. A. & Jin, J. Non-substrate based, small molecule inhibitors of the human isoprenylcysteine carboxyl methyltransferase. *Medchemcomm* **7**, 1016-1021, doi:10.1039/c6md00130k (2016).
- 76 Eckert, G. P. *et al.* Regulation of the brain isoprenoids farnesyl- and geranylgeranylpyrophosphate is altered in male Alzheimer patients. *Neurobiology of Disease* **35**, 251-257, doi:10.1016/j.nbd.2009.05.005 (2009).
- 77 Li, H. *et al.* Protein Prenylation Constitutes an Endogenous Brake on Axonal Growth. *Cell Reports* **16**, 545-558, doi:10.1016/j.celrep.2016.06.013 (2016).
- 78 Ochocki, J. D. & Distefano, M. D. Prenyltransferase inhibitors: treating human ailments from cancer to parasitic infections. *Medchemcomm* **4**, 476-492, doi:10.1039/c2md20299a (2013).
- 79 Gordon, L. B., Rothman, F. G., Lopez-Otin, C. & Misteli, T. Progeria: A Paradigm for Translational Medicine. *Cell* **156**, 400-407, doi:10.1016/j.cell.2013.12.028 (2014).
- 80 Der, C. J., Krontiris, T. G. & Cooper, G. M. Transforming genes of human bladder and lung-carcinoma cell-lines are homologous to the ras genes of harvey and kirsten sarcoma-viruses. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **79**, 3637-3640, doi:10.1073/pnas.79.11.3637 (1982).
- 81 Bos, J. L. Ras oncogenes in human cancer - A review. *Cancer Research* **49**, 4682-4689 (1989).
- 82 Boulalas, I., Zaravinos, A., Karyotis, I., Delakas, D. & Spandidos, D. A. Activation of RAS Family Genes in Urothelial Carcinoma. *Journal of Urology* **181**, 2312-2319, doi:10.1016/j.juro.2009.01.011 (2009).
- 83 van Dijk, M., Bernsen, M. R. & Ruiter, D. J. Analysis of mutations in B-RAF, N-RAS, and H-RAS genes in the differential diagnosis of Spitz nevus and spitzoid melanoma. *American Journal of Surgical Pathology* **29**, 1145-1151, doi:10.1097/01.pas.0000157749.18591.9e (2005).
- 84 Hruban, R. H. *et al.* K-RAS oncogene activation in adenocarcinoma of the human pancreas - a study of 82 carcinomas using a combination of mutant-enriched polymerase chain-reaction analysis and allele-specific oligonucleotide hybridization. *American Journal of Pathology* **143**, 545-554 (1993).

- 85 Prior, I. A., Lewis, P. D. & Mattos, C. A Comprehensive Survey of Ras Mutations in Cancer. *Cancer Research* **72**, 2457-2467, doi:10.1158/0008-5472.can-11-2612 (2012).
- 86 Quinlan, M. P. & Settleman, J. Isoform-specific ras functions in development and cancer. *Future Oncology* **5**, 105-116, doi:10.2217/14796694.5.1.105 (2009).
- 87 Yang, S. H. *et al.* Severe hepatocellular disease in mice lacking one or both CaaX prenyltransferases. *Journal of Lipid Research* **53**, 77-86, doi:10.1194/jlr.M021220 (2012).
- 88 Bourne, H. R., Sanders, D. A. & McCormick, F. The GTPase superfamily - a conserved switch for diverse cell functions. *Nature* **348**, 125-132, doi:10.1038/348125a0 (1990).
- 89 Karnoub, A. E. & Weinberg, R. A. Ras oncogenes: split personalities. *Nature Reviews Molecular Cell Biology* **9**, 517-531, doi:10.1038/nrm2438 (2008).
- 90 Colicelli, J. Human RAS superfamily proteins and related GTPases. *Science's STKE : signal transduction knowledge environment* **2004**, RE13-RE13, doi:10.1126/stke.2502004re13 (2004).
- 91 Hancock, J. F., Cadwallader, K., Paterson, H. & Marshall, C. J. A CAAX or a CAAL motif and a 2nd signal are sufficient for plasma-membrane targeting of ras proteins. *Embo Journal* **10**, 4033-4039, doi:10.1002/j.1460-2075.1991.tb04979.x (1991).
- 92 Leon, J., Guerrero, I. & Pellicer, A. Differential expression of the ras gene family in mice. *Molecular and Cellular Biology* **7**, 1535-1540, doi:10.1128/mcb.7.4.1535 (1987).
- 93 Schlichting, I. *et al.* Time-resolved x-ray crystallographic study of the conformational change in ha-ras p21 protein on gtp hydrolysis. *Nature* **345**, 309-315, doi:10.1038/345309a0 (1990).
- 94 Herrmann, C. Ras-effector interactions: after one decade. *Current Opinion in Structural Biology* **13**, 122-129, doi:10.1016/s0959-440x(02)00007-6 (2003).
- 95 Field, J., Broek, D., Kataoka, T. & Wigler, M. Guanine-nucleotide activation of, and competition between, Ras proteins from *saccharomyces-cerevisiae*. *Molecular and Cellular Biology* **7**, 2128-2133, doi:10.1128/mcb.7.6.2128 (1987).
- 96 Vetter, I. R. & Wittinghofer, A. Signal transduction - The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299-1304, doi:10.1126/science.1062023 (2001).
- 97 Almoguera, C. *et al.* Most human carcinomas of the exocrine pancreas contain mutant C-K-Ras genes. *Cell* **53**, 549-554, doi:10.1016/0092-8674(88)90571-5 (1988).
- 98 Lohr, M., Kloppel, G., Maisonneuve, P., Lowenfels, A. B. & Luttges, J. Frequency of K-ras mutations in pancreatic intraductal neoplasias associated with pancreatic ductal adenocarcinoma and chronic pancreatitis: A meta-analysis. *Neoplasia* **7**, 17-23, doi:10.1593/neo.04445 (2005).

- 99 Fernandez, C. *et al.* From pancreatic intraepithelial neoplasia to cancer: a dramatic progression with K-ras status analysis. *Gastroenterologie Clinique Et Biologique* **29**, 465-468, doi:10.1016/s0399-8320(05)80818-8 (2005).
- 100 Rojas, J. M. & Santos, E. ras genes and human cancer: Different implications and different roles. *Current Genomics* **3**, 295-311, doi:10.2174/1389202023350381 (2002).
- 101 Cox, A. D. & Der, C. J. Ras history: the saga continues. *Small GTPases* **1**, 2-27, doi:10.4161/sgtp.1.1.12178 (2010).
- 102 Bos, J. L. Genetic mechanisms in tumor initiation and progression. 10. The ras gene family and human carcinogenesis. *Mutation Research* **195**, 255-271, doi:10.1016/0165-1110(88)90004-8 (1988).
- 103 Barbacid, M. RAS GENES. *Annual Review of Biochemistry* **56**, 779-827, doi:10.1146/annurev.bi.56.070187.004023 (1987).
- 104 Miller, M. S., and Miller, L. D. RAS mutations and oncogenesis: Not all RAS mutations are created equally. *Front. Genet.* 10.3389/fgene.2011.00100 (2012).
- 105 Anderson, J. L., Frase, H., Michaelis, S. & Hrycyna, C. A. Purification, functional reconstitution, and characterization of the *Saccharomyces cerevisiae* isoprenylcysteine carboxylmethyltransferase Ste14. *Journal of Biological Chemistry* **280**, 7336-7345, doi:10.1074/jbc.M410292200 (2005).
- 106 Romano, J. D. & Michaelis, S. Topological and mutational analysis of *Saccharomyces cerevisiae* Ste14, founding member of the isoprenylcysteine carboxyl methyltransferase family. *Molecular Biology of the Cell* **12**, 1957-1971, doi:10.1091/mbc.12.7.1957 (2001).
- 107 Ashby, M. N., Errada, P. R., Boyartchuk, V. L. & Rine, J. Isolation and dna-sequence of the ste14 gene encoding farnesyl cysteine - carboxyl methyltransferase. *Yeast* **9**, 907-913, doi:10.1002/yea.320090810 (1993).
- 108 Wright, L. P. *et al.* Topology of Mammalian Isoprenylcysteine Carboxyl Methyltransferase Determined in Live Cells with a Fluorescent Probe. *Molecular and Cellular Biology* **29**, 1826-1833, doi:10.1128/mcb.01719-08 (2009).
- 109 Logue, A. *Characterization of the S-Adenosyl-L-Methionine (SAM) Binding Domain of the Yeast Isoprenylcysteine Carboxyl Methyltransferase, Ste14* Ph.D. thesis, Purdue University, (2013).
- 110 Olsen, K. *Mechanism and Conformational Dynamics of the Yeast Isoprenylcysteine Carboxyl Methyltransferase* Ph.D. thesis, Purdue University, (2016).
- 111 Diver, M. M. & Long, S. B. Mutational Analysis of the Integral Membrane Methyltransferase Isoprenylcysteine Carboxyl Methyltransferase (ICMT) Reveals Potential Substrate Binding Sites. *Journal of Biological Chemistry* **289**, 26007-26020, doi:10.1074/jbc.M114.585125 (2014).

- 112 Senes, A., Gerstein, M. & Engelman, D. M. Statistical analysis of amino acid patterns in transmembrane helices: The GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *Journal of Molecular Biology* **296**, 921-936, doi:10.1006/jmbi.1999.3488 (2000).
- 113 Russ, W. P. & Engelman, D. M. TOXCAT: A measure of transmembrane helix association in a biological membrane. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 863-868, doi:10.1073/pnas.96.3.863 (1999).
- 114 Russ, W. P. & Engelman, D. M. The GxxxG motif: A framework for transmembrane helix-helix association. *Journal of Molecular Biology* **296**, 911-919, doi:10.1006/jmbi.1999.3489 (2000).
- 115 Gerber, D. & Shai, Y. In vivo detection of hetero-association of glycophorin-A and its mutants within the membrane. *Journal of Biological Chemistry* **276**, 31229-31232, doi:10.1074/jbc.M101889200 (2001).
- 116 MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. A transmembrane helix dimer: Structure and implications. *Science* **276**, 131-133, doi:10.1126/science.276.5309.131 (1997).
- 117 Bhatia, A., Schafer, H. J. & Hrycyna, C. A. Oligomerization of the human ABC transporter ABCG2: Evaluation of the native protein and chimeric dimers. *Biochemistry* **44**, 10893-10904, doi:10.1021/bi0503807 (2005).
- 118 Polgar, O. *et al.* Mutational analysis of ABCG2: Role of the GXXXG motif. *Biochemistry* **43**, 9448-9456, doi:10.1021/bi0497953 (2004).
- 119 Arselin, G. *et al.* The GxxxG motif of the transmembrane domain of subunit e is involved in the dimerization/oligomerization of the yeast ATP synthase complex in the mitochondrial membrane. *European Journal of Biochemistry* **270**, 1875-1884, doi:10.1046/j.1432-1033.2003.03557.x (2003).
- 120 Whittington, D. A. *et al.* Crystal structure of the dimeric extracellular domain of human carbonic anhydrase XII, a bitopic membrane protein overexpressed in certain cancer tumor cells. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 9545-9550, doi:10.1073/pnas.161301298 (2001).
- 121 McClain, M. S. *et al.* Essential role of a GXXXG motif for membrane channel formation by *Helicobacter pylori* vacuolating toxin. *Journal of Biological Chemistry* **278**, 12101-12108, doi:10.1074/jbc.M212595200 (2003).
- 122 Overton, M. C., Chinault, S. L. & Blumer, K. J. Oligomerization, biogenesis, and signaling is promoted by a glycophorin A-like dimerization motif in transmembrane domain 1 of a yeast G protein-coupled receptor. *Journal of Biological Chemistry* **278**, 49369-49377, doi:10.1074/jbc.M308654200 (2003).

- 123 Senes, A., Engel, D. E. & DeGrado, W. F. Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Current Opinion in Structural Biology* **14**, 465-479, doi:10.1016/j.sbi.2004.07.007 (2004).
- 124 Griggs, A. M., Hahne, K. & Hrycyna, C. A. Functional Oligomerization of the *Saccharomyces cerevisiae* Isoprenylcysteine Carboxyl Methyltransferase, Ste14. *Journal of Biological Chemistry* **285**, 13380-13387, doi:10.1074/jbc.M109.061366 (2010).
- 125 Shi, Y. Q. & Rando, R. R. Kinetic mechanism of isoprenylated protein methyltransferase. *Journal of Biological Chemistry* **267**, 9547-9551 (1992).
- 126 Baron, R. & Casey, P. J. Analysis of the kinetic mechanism of recombinant human isoprenylcysteine carboxylmethyltransferase (Icmt) *BMC Biochem* **5**, 1-12 (2004).
- 127 Volker, C. *et al.* Effects of farnesylcysteine analogs on protein carboxyl methylation and signal transduction. *Journal of Biological Chemistry* **266**, 21515-21522 (1991).
- 128 Hahne, K. *et al.* Evaluation of substrate and inhibitor binding to yeast and human isoprenylcysteine carboxyl methyltransferases (Icmts) using biotinylated benzophenone-containing photoaffinity probes. *Biochemical and Biophysical Research Communications* **423**, 98-103, doi:10.1016/j.bbrc.2012.05.089 (2012).
- 129 Desrosiers, R. R., Nguyen, Q. T. & Beliveau, R. The carboxyl methyltransferase modifying G proteins is a metalloenzyme. *Biochemical and Biophysical Research Communications* **261**, 790-797, doi:10.1006/bbrc.1999.0936 (1999).
- 130 Hodges, H. B. *et al.* Inhibition of membrane-associated methyltransferases by a cholesterol-based metal chelator. *Bioconjugate Chemistry* **16**, 490-493, doi:10.1021/bc050027d (2005).
- 131 Funk, A. L. *Biochemical Elucidation of the Isoprenoid Binding Site of the Yeast Isoprenylcysteine Carboxyl Methyltransferase, Ste14* Ph.D. thesis, Purdue University, (2017).
- 132 Diver, M. M., Pedi, L., Koide, A., Koide, S. & Long, S. B. Atomic structure of the eukaryotic intramembrane RAS methyltransferase ICMT. *Nature* **553**, 526-+, doi:10.1038/nature25439 (2018).
- 133 Volker, C., Lane, P., Kwee, C., Johnson, M. & Stock, J. A Single activity carboxyl methylates both farnesyl and geranylgeranyl cysteine residues. *Febs Letters* **295**, 189-194, doi:10.1016/0014-5793(91)81415-5 (1991).
- 134 Stephenson, R. C. & Clarke, S. Characterization of a rat-liver protein carboxyl methyltransferase involved in the maturation of proteins with the -CXXX C-terminal sequence motif. *Journal of Biological Chemistry* **267**, 13314-13319 (1992).
- 135 Tan, E. W., Perezsala, D., Canada, F. J. & Rando, R. R. Identifying the recognition unit for G-protein methylation. *Journal of Biological Chemistry* **266**, 10719-10722 (1991).

- 136 Gilbert, B. A., Tan, E. W., Perezsala, D. & Rando, R. R. Structure activity studies on the retinal rod outer segment isoprenylated protein methyltransferase. *Journal of the American Chemical Society* **114**, 3966-3973, doi:10.1021/ja00036a052 (1992).
- 137 Volker, C., Miller, R. A. & Stock, J. B. S Farnesylcysteine methyltransferase in bovine brain. *Methods (Orlando)* **1**, 283-287, doi:10.1016/s1046-2023(05)80329-9 (1990).
- 138 Rahib, L. *et al.* Projecting Cancer Incidence and Deaths to 2030: The Unexpected Burden of Thyroid, Liver, and Pancreas Cancers in the United States. *Cancer Research* **74**, 2913-2921, doi:10.1158/0008-5472.can-14-0155 (2014).
- 139 Fisher, G. H. *et al.* Induction and apoptotic regression of lung adenocarcinomas by regulation of a K-Ras transgene in the presence and absence of tumor suppressor genes. *Genes & Development* **15**, 3249-3262, doi:10.1101/gad.947701 (2001).
- 140 Ying, H. *et al.* Oncogenic Kras Maintains Pancreatic Tumors through Regulation of Anabolic Glucose Metabolism. *Cell* **149**, 656-670, doi:10.1016/j.cell.2012.01.058 (2012).
- 141 Singh, A. *et al.* A Gene Expression Signature Associated with "K-Ras Addiction" Reveals Regulators of EMT and Tumor Cell Survival. *Cancer Cell* **15**, 489-500, doi:10.1016/j.ccr.2009.03.022 (2009).
- 142 John, J. *et al.* Kinetics of interaction of nucleotides with nucleotide-free H-ras p21. *Biochemistry* **29**, 6058-6065, doi:10.1021/bi00477a025 (1990).
- 143 Wittinghofer, A. & Pai, E. F. The structure of ras protein - a model for a universal molecular switch. *Trends in Biochemical Sciences* **16**, 382-387, doi:10.1016/0968-0004(91)90156-p (1991).
- 144 Papke, B. & Der, C. J. Drugging RAS: Know the enemy. *Science* **355**, 1158-1163, doi:10.1126/science.aam7622 (2017).
- 145 Kohl, N. E. *et al.* Selective-inhibition of ras-dependent transformation by a farnesyltransferase inhibitor. *Science* **260**, 1934-1937, doi:10.1126/science.8316833 (1993).
- 146 Kohl, N. E. *et al.* Inhibition of farnesyltransferase induces regression of mammary and salivary carcinomas in ras transgenic mice. *Nature Medicine* **1**, 792-797, doi:10.1038/nm0895-792 (1995).
- 147 James, G. L. *et al.* Benzodiazepine peptidomimetics - Potent inhibitors of ras farnesylation in animal-cells. *Science* **260**, 1937-1942, doi:10.1126/science.8316834 (1993).
- 148 Venet, M., End, D. & Angibaud, P. Farnesyl protein transferase inhibitor ZARNESTRA (TM) R115777 - History of a discovery. *Current Topics in Medicinal Chemistry* **3**, 1095-1102, doi:10.2174/1568026033452050 (2003).

- 149 Van Cutsem, E. *et al.* Phase III trial of gemcitabine plus tipifarnib compared with gemcitabine plus placebo in advanced pancreatic cancer. *Journal of Clinical Oncology* **22**, 1430-1438, doi:10.1200/jco.2004.10.112 (2004).
- 150 Rao, S. *et al.* Phase III double-blind placebo-controlled study of farnesyl transferase inhibitor R115777 in patients with refractory advanced colorectal cancer. *Journal of Clinical Oncology* **22**, 3950-3957, doi:10.1200/jco.2004.10.037 (2004).
- 151 Blumenschein, G. *et al.* A randomized phase III trial comparing ionafarnib/carboplatin/paclitaxel versus carboplatin/paclitaxel (CP) in chemotherapy-naïve patients with advanced or metastatic non-small cell lung cancer (NSCLC). *Lung Cancer* **49**, S30-S30, doi:10.1016/s0169-5002(05)80215-6 (2005).
- 152 Harousseau, J.-L. *et al.* A randomized phase 3 study of tipifarnib compared with best supportive care, including hydroxyurea, in the treatment of newly diagnosed acute myeloid leukemia in patients 70 years or older. *Blood* **114**, 1166-1173, doi:10.1182/blood-2009-01-198093 (2009).
- 153 Whyte, D. B. *et al.* K- and N-Ras are geranylgeranylated in cells treated with farnesyl protein transferase inhibitors. *Journal of Biological Chemistry* **272**, 14459-14464, doi:10.1074/jbc.272.22.14459 (1997).
- 154 Rowell, C. A., Kowalczyk, J. J., Lewis, M. D. & Garcia, A. M. Direct demonstration of geranylgeranylation and farnesylation of Ki-Ras in vivo. *Journal of Biological Chemistry* **272**, 14093-14097, doi:10.1074/jbc.272.22.14093 (1997).
- 155 Ohkanda, J. *et al.* A Guanidyl-based Bivalent Peptidomimetic Inhibits K-Ras Prenylation and Association with c-Raf. *Chemistry (Weinheim an der Bergstrasse, Germany)*, doi:10.1002/chem.201903129 (2019).
- 156 Bergo, M. O. *et al.* Absence of the CAAX endoprotease Rce1: Effects on cell growth and transformation. *Molecular and Cellular Biology* **22**, 171-181, doi:10.1128/mcb.22.1.171-181.2002 (2002).
- 157 Ma, Y. T., Gilbert, B. A. & Rando, R. R. Inhibitors of the isoprenylated protein endoprotease. *Biochemistry* **32**, 5924-5924, doi:10.1021/bi00073a600 (1993).
- 158 Chen, Y. L., Ma, Y. T. & Rando, R. R. Solubilization, partial purification, and affinity labeling of the membrane-bound isoprenylated protein endoprotease. *Biochemistry* **35**, 3227-3237, doi:10.1021/bi952529s (1996).
- 159 Schlitzer, M., Winter-Vann, A. & Casey, P. J. Non-peptidic, non-prenylic inhibitors of the prenyl protein-specific protease Rce1. *Bioorganic & Medicinal Chemistry Letters* **11**, 425-427, doi:10.1016/s0960-894x(00)00685-5 (2001).
- 160 Mohammed, I. *et al.* 8-Hydroxyquinoline-based inhibitors of the Rce1 protease disrupt Ras membrane localization in human cells. *Bioorganic & Medicinal Chemistry* **24**, 160-178, doi:10.1016/j.bmc.2015.11.043 (2016).

- 161 Bergo, M. *et al.* Targeting ICMT ameliorates K-RAS-induced myeloproliferative disease and lung cancer. *Haematologica-the Hematology Journal* **92**, 345-346 (2007).
- 162 Bergo, M. O. *et al.* On the physiological importance of endoproteolysis of CAAX proteins - Heart-specific RCE1 knockout mice develop a lethal cardiomyopathy. *Journal of Biological Chemistry* **279**, 4729-4736, doi:10.1074/jbc.M310081200 (2004).
- 163 Bergo, M. O. *et al.* Zmpste24 deficiency in mice causes spontaneous bone fractures, muscle weakness, and a prelamin A processing defect. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 13049-13054, doi:10.1073/pnas.192460799 (2002).
- 164 Pendas, A. M. *et al.* Defective prelamin A processing and muscular and adipocyte alterations in Zmpste24 metalloproteinase-deficient mice. *Nature Genetics* **31**, 94-99, doi:10.1038/ng871 (2002).
- 165 Bergo, M. O. *et al.* Inactivation of Icmt inhibits transformation by oncogenic K-Ras and B-Raf. *Journal of Clinical Investigation* **113**, 539-550, doi:10.1172/jci200418829 (2004).
- 166 Bergo, M. O. *et al.* Isoprenylcysteine carboxyl methyltransferase deficiency in mice. *Journal of Biological Chemistry* **276**, 5841-5845, doi:10.1074/jbc.C000831200 (2001).
- 167 Lau, H. Y., Tang, J., Casey, P. J. & Wang, M. Isoprenylcysteine carboxylmethyltransferase is critical for malignant transformation and tumor maintenance by all RAS isoforms. *Oncogene* **36**, 3934-3942, doi:10.1038/onc.2016.508 (2017).
- 168 Wahlstrom, A. M. *et al.* Inactivating Icmt ameliorates K-RAS-induced myeloproliferative disease. *Blood* **112**, 1357-1365, doi:10.1182/blood-2007-06-094060 (2008).
- 169 Court, H. *et al.* Isoprenylcysteine carboxylmethyltransferase deficiency exacerbates KRAS-driven pancreatic neoplasia via Notch suppression. *Journal of Clinical Investigation* **123**, 4681-4694, doi:10.1172/bci65764 (2013).
- 170 Molony, L., NgSikorski, J., Hellberg, C. & Andersson, T. Inhibitors of farnesyl and geranylgeranyl methyltransferases prevent beta(2) integrin-induced actin polymerization without affecting beta(2) integrin-induced Ca²⁺ signaling in neutrophils. *Biochemical and Biophysical Research Communications* **223**, 612-617, doi:10.1006/bbrc.1996.0943 (1996).
- 171 Majmudar, J. D., Hahne, K., Hrycyna, C. A. & Gibbs, R. A. Probing the isoprenylcysteine carboxyl methyltransferase (Icmt) binding pocket: Sulfonamide modified farnesyl cysteine (SMFC) analogs as Icmt inhibitors. *Bioorganic & Medicinal Chemistry Letters* **21**, 2616-2620, doi:10.1016/j.bmcl.2011.01.078 (2011).
- 172 Bergman, J. A., Hahne, K., Song, J., Hrycyna, C. A. & Gibbs, R. A. S-Farnesyl-Thiopropionic Acid Triazoles as Potent Inhibitors of Isoprenylcysteine Carboxyl Methyltransferase. *Acs Medicinal Chemistry Letters* **3**, 15-19, doi:10.1021/ml200106d (2012).

- 173 Lau, H. Y. *et al.* An improved isoprenylcysteine carboxylmethyltransferase inhibitor induces cancer cell death and attenuates tumor growth in vivo. *Cancer Biology & Therapy* **15**, 1280-1291, doi:10.4161/cbt.29692 (2014).
- 174 Buchanan, M. S. *et al.* Spermatinamine, the first natural product inhibitor of isoprenylcysteine carboxyl methyltransferase, a new cancer target. *Bioorganic & Medicinal Chemistry Letters* **17**, 6860-6863, doi:10.1016/j.bmcl.2007.10.021 (2007).
- 175 Buchanan, M. S. *et al.* Aplysamine 6, an alkaloidal inhibitor of isoprenylcysteine carboxyl methyltransferase from the sponge *Pseudoceratina* sp. *Journal of Natural Products* **71**, 1066-1067, doi:10.1021/np0706623 (2008).
- 176 Buchanan, M. S. *et al.* Small-molecule inhibitors of the cancer target, isoprenylcysteine carboxyl methyltransferase, from *Hovea parvicalyx*. *Phytochemistry* **69**, 1886-1889, doi:10.1016/j.phytochem.2008.04.011 (2008).
- 177 Wang, M. *et al.* A small molecule inhibitor of isoprenylcysteine carboxymethyltransferase induces autophagic cell death in PC3 prostate cancer cells. *Journal of Biological Chemistry* **283**, 18678-18684, doi:10.1074/jbc.M801855200 (2008).
- 178 Winter-Vann, A. M. *et al.* A small-molecule inhibitor of isoprenylcysteine carboxyl methyltransferase with antitumor activity in cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 4336-4341, doi:10.1073/pnas.0408107102 (2005).
- 179 Go, M.-L. *et al.* Amino Derivatives of Indole As Potent Inhibitors of Isoprenyleysteine Carboxyl Methyltransferase. *Journal of Medicinal Chemistry* **53**, 6838-6850, doi:10.1021/jm1002843 (2010).
- 180 Tsimberidou, A. M. *et al.* Phase I clinical trial of hepatic arterial infusion of paclitaxel in patients with advanced cancer and dominant liver involvement. *Cancer Chemotherapy and Pharmacology* **68**, 247-253, doi:10.1007/s00280-010-1482-y (2011).
- 181 Porta, E. O. J. *et al.* Repositioning Salirasib as a new antimalarial agent. *Med. Chem. Commun.*, doi:10.1039/c9md00298g (2019).
- 182 Sievers F., Wilm A., Dineen D., Gibson T. J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J. D. & Higgins D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
- 183 Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. and Barton, G. J. Jalview Version 2 - a multiple sequence alignment editor and analysis workbench Bioinformatics doi: 10.1093/bioinformatics/btp033, (2009).

CHAPTER 2. OPTIMIZATION OF EXPRESSION AND PURIFICATION OF HIS-STE14 UTILIZING THE BACULOVIRUS EXPRESSION SYSTEM

2.1 Introduction

To better understand the structure and function of His-Ste14, we choose to use various biochemical and biophysical techniques because we needed an efficient way to express the enzyme. Techniques used later in this dissertation like electron paramagnetic resonance (EPR) spectroscopy (chapter 3) and single particle cryo-electron microscopy (cryo-EM) (chapter 4), require high concentrations of pure and homogeneous protein. Thus, it was essential to test different expression systems to produce large yields of His-Ste14. We expressed His₁₀-myc₃N-Ste14 (His-Ste14) in yeast, *Saccharomyces cerevisiae* (*S. cerevisiae*), and two insect cell lines, *Spondoptera frugiperda* (Sf9s) and *Trichoplusia ni* (High Fives). Yeast is advantageous because it is less expensive and easy to handle.^{1,2} On the other hand, insect cells are more costly and more sensitive to their environment, yet they can produce higher yields of protein.

Additionally, membrane proteins are challenging to work with and difficult to extract and purify from crude membrane preparations, therefore purification protocols required optimization to retain structural integrity and function of the protein. Detergents are amphiphilic molecules that auto-assemble into micelles above the critical micelle concentration (CMC). Above the CMC, the detergent solubilizes membrane proteins forming a hydrophobic lipid bilayer around the protein, creating protein-detergent complexes of a particular size.^{3,4} Purification of membrane proteins in detergent allows the protein to become water soluble. However, even if the protein is solubilized, it cannot always be isolated from the membrane in a stable state or in its correct native structure. Thus, it becomes imperative to find which detergent best stabilizes the protein through a detergent screening process. To accomplish this goal, it was necessary to understand the physiochemical characteristics of different detergents. For example, membrane proteins are usually solubilized with mild, nonionic, or zwitterionic detergents rather than ionic detergents, which can cause denaturation. Anderson *et al.* performed an extensive screen on His-Ste14 with a panel of detergents and found *n*-dodecyl- β -D-maltopyranoside (DDM) had optimal solubilization and retention of His-Ste14 activity.⁵ Although a variety of nonionic detergents (DDM, TritonX-100, TritonX-114, TritonX-305, TritonX-605, Tween 80, Brij 35, Brij 58, *n*-octyl- β -D-glucopyranoside

(OG), *n*-octyl- β -D-thioglucopyranoside (OTG)) and several zwitterionic detergents (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPSO)) were used to solubilize His-Ste14 in this previous study, we purified His-Ste14 in additional detergents more commonly used to purify membrane proteins for biophysical techniques. Anderson *et al.* tested a variety of TritonX and Tween, yet preparations of these industrial grade detergents are often heterogeneous in the composition of their polyoxyethylene chains.⁶ Thus, TritonX and Tween detergents may not always be sufficient for protein stabilization and future technique like EPR spectroscopy, cryo-EM, and X-ray crystallography. Some of the more common detergents used for membrane protein purification and further functional and structural studies are the alkyl glycosides.⁷ Therefore, we solubilized His-Ste14 with other alkyl glycosides including decyl maltose neopentyl glycol (DMNG) and lauryl maltose neopentyl glycol (LMNG). A complete summary of detergents used in this study can be found in Table 3.1. Finally, buffer components and pH can also have an effect on the purification of proteins.⁸ Hence, several buffer systems at pH values 7.2 and 6.0 were used to purify His-Ste14 from crude membranes. In addition, three types of buffers were also used to purify His-Ste14: (1) a phosphate buffer (1x EQ buffer), (2) a HEPES buffer, and (3) a Tris-HCl buffer only at pH 7.2. Together, this data was used to determine the best conditions to purify membrane protein, His-Ste14.

2.2 Materials and Methods

2.2.1 Materials

Yeast parent strain (SM1188) was a gift from Dr. Susan Michaelis (Johns Hopkins School of Medicine, MD). Yeast strain 2 μ URA3 P_{PGK}-His₁₀-myc₃N-STE14 (CH2704) was previously produced by Hrycyna lab members. Sf9 cells, High Five cells, and ES 921 media were purchased from Expression Systems (Davis, CA). The monoclonal α -myc IgG, the polyclonal goat- α -rabbit IgG HRP conjugate and the polyclonal goat- α -mouse IgG HRP conjugate were purchased from Invitrogen Life Sciences (Carlsbad, CA). GeneJET Plasmid Miniprep Kit and Enhanced chemiluminescence (ECL) and pFastBac vectors were purchased from ThermoFisher (Waltham, MA). BigDye® Terminator v3.1 was obtained from Applied Biosystems (Foster City, CA) through the Purdue University Genomics Facility. *S*-adenosyl-L-methionine was purchased from Sigma Aldrich (St. Louis, MO). GeneGnome XRQ was purchased from Syngene (Fredrick, MD). 4-(2-

aminoethyl)-benzenesulfonylfluoride hydrochloride (AEBSF), aprotinin, and dithiothreitol (DTT) were purchased from Gold Biotechnology (St. Louis, MO). Tetraethylene glycol monododecyl ether (C₈E₄), heptaethylene glycol monododecyl ether (C₁₂E₇), *n*-Dodecyl- β -D-maltopyranoside (DDM), decyl maltose neopentyl glycol (DMNG), and lauryl maltose neopentyl glycol (LMNG) were purchased from Anatrace (Maumee, OH). The detergents maltose neopentyl glycol-3 (MNG-3) and Lauryldimethylamine oxide (LDAO) were gifts from Dr. Nicholas Noinaj (Purdue University, IN). The substrate *N*-acetyl-*S*-farnesyl-*L*-cysteine (AFC) was purchased from Enzo Life Sciences (Farmingdale, NY). ¹⁴C-labeled *S*-adenosyl-*L*-methionine ([¹⁴C]-SAM) was purchased from Perkin Elmer (Waltham, MA). Nitrocellulose membrane (0.22 μ m) and Superdex S200 column were purchased from GE Healthcare Life Sciences (Pittsburg, PA). Talon cobalt metal affinity resin was purchased from Clontech (Mountain View, CA). Amicon Ultra 30,000 MWCO concentrators and centrifugal filter units (0.22 μ m) GV Durapore® were purchased from Millipore Co. (Billerica, MA). Poly-Prep® chromatography columns were purchased from Bio-Rad Laboratories (Hercules, CA). Biosafe II scintillation fluid was purchased from Research Products International (Mount Prospect, IL). All other materials and reagents were purchased from Fisher Scientific (Hampton, NH).

2.2.2 Yeast Strains, Transformation, and Yeast Crude Membrane Preparations

The yeast *STE14* plasmid was transformed into the Δ *ste14* yeast strain SM1188, as previously described.⁹ The wild-type *STE14* plasmid has the genotype: 2 μ URA3 P_{PGK}-His₁₀-myc₃N-STE14. SM1188 was grown on yeast peptone dextrose (YPD) agarose plates and was cultured at 220rpm in YPD media overnight at 30°C. The next day, the culture was cut back 1:3 with media and placed back in the 30°C shaking incubator. After an additional 2 hours, 500 μ L of the culture was centrifuged and the supernatant was removed. The pellet was resuspended with single stranded salmon testes DNA (0.1mg). 1 μ g of each plasmid DNA of interest was added, along with autoclaved plate mix (40.5% polyethylene glycol, 0.1M lithium acetate, 10mM Tris-HCl, pH 7.5, 1mM ethylenediaminetetraacetic acid (EDTA), pH 8), and 50mM dithiothreitol (DTT). The resulting mixture incubated overnight at room temperature followed by a 10 min heat shock at 42°C. 100 μ L was plated onto synthetic complete medium without uracil (SC-URA) plates and allowed to grow for 48 hours at 30°C.

The transformed yeast cell cultures were grown in SC-URA to mid-log phase (OD_{600} 3-5), pelleted by centrifugation at 4000 xg and stored at -80°C . Preparation of yeast crude membranes was completed as previously described with minor changes.⁵ Briefly, the yeast cell pellets overexpressing His-Ste14 were resuspended in lysis buffer (300 mM sorbitol, 100mM NaCl, 5mM $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$, 10mM Tris-HCl, pH 7.5, and 2mM AEBSF) and was placed on ice to swell for 15 mins. Pellets underwent rapid freeze-thaw cycles twice in liquid N_2 and a 25°C water bath, respectively. Resuspended cell pellets were passed through a French press at 12,000 p.s.i. twice. The lysate was centrifuged at 500 xg for 10 mins at 4°C to remove any cell debris. The supernatant was collected, and the remaining membrane fraction was isolated through ultracentrifugation at 100,000 xg for 1 hour at 4°C . After centrifugation, the supernatant was removed and the pellet was resuspended in 10 mM Tris-HCl, pH 7.5, aliquoted, and flash frozen in liquid N_2 . All aliquots were stored at -80°C until further experimentation. Total protein concentration was determined using the Bradford protein assay.¹⁰

2.2.3 Expression and Membrane Preparation of His-Ste14 from Insect Cell Lines

A cDNA encoding *S. cerevisiae* Ste14 was subcloned into the pFastBac-Dual vector with the addition of ten histidine residues (His_{10} -), and three *myc* tags with the sequence 5' EQKLISEEDL 3' (*myc3*-). The gene and vector sequence were confirmed with bidirectional dye-terminator sequencing (BigDye® Terminator v3.1). Protein was expressed in baculovirus-infected Sf9 or High Five cells and harvested from ES 921 media with 1x pencillin-streptomycin at 48 h post-infection. Pelleted cells were flash frozen for membrane preparation. Briefly, the insect cell pellets were thawed and resuspended in ~20mL of lysis buffer per 20g of cell pellet harvested (50 mM mannitol, 50 mM Tris-HCl, pH 7.5, 1% aprotinin, 1 mM AEBSF and 2mM DTT) and were rocked for 15 min at 4°C . Typically, 20g of cell pellet is equivalent to 750mL of cells harvested. The resuspended pellets were lysed via dounce homogenization and isolated through ultracentrifugation at 100,000 x g for 1 hour at 4°C . Following centrifugation, the membrane pellet was resuspended in buffer (300 mM mannitol, 50 mM Tris-HCl, pH 7.5, 1% aprotinin, 1 mM AEBSF and 1mM DTT, 10% glycerol), aliquoted, flash frozen in liquid nitrogen and stored at -80°C . Total protein concentration was determined using the Bradford protein assay.¹⁰

2.2.4 *In vitro* Methyltransferase Vapor Diffusion Assay of Yeast and Insect Cell Crude Membranes

An *in vitro* methyltransferase vapor diffusion assay was performed as previously described.^{5,11,12} Briefly, reactions were prepared in a 100mM Tris-HCl, pH 7.5 buffer, containing crude membranes overexpressing His-Ste14 (CH2704) or His-Ste14 isolated from insect cell lines (5µg), AFC (200µM), and [¹⁴C]-SAM (20µM). The negative control reaction contained a yeast deletion strain of *ste14* (CH2714, 5µg) and all other components described above. Samples were incubated for 30 mins at 30°C and then were stopped with 50 µL of 1 M NaOH/1% SDS (v/v). Each reaction (100µL) was spotted onto a filter and placed in the neck of a scintillation vial filled with Biosafe II scintillation fluid (10mL) and capped at room temperature. After three hours, the filters are removed and the vials were placed in a Packard Tri-Carb 1600CA liquid scintillation counter in order to detect the [¹⁴C]-methanol that diffused from filters into the scintillation fluid. All specific activities were characterized as pmol of methyl groups from [¹⁴C]-SAM transferred to the AFC substrate per minute of assay per mg of His-Ste14 (CH2704) or His-Ste14 isolated from insect cell lines. These activities were calculated after subtracting background counts from the average of the negative control samples. Each experiment had three reactions per sample and all experiments were performed in triplicate.

2.2.5 SDS-PAGE Staining and Immunoblot Analysis

Crude membrane extracts overexpressing His-Ste14 (CH2704 or insect cell lines) (1µg) or purified His-Ste14 (CH2704 or insect cell lines) (0.1µg for immunoblot and 1.0µg for stained gels) were separated on a 10% SDS-PAGE. Gels were either incubated with Coomassie stain (0.25% (w/v) Coomassie Brilliant Blue R-250, 80% methanol and 20% acetic acid) or silver stain containing silver nitrate. Gels stained with Coomassie incubated overnight. The next day, gels were incubated with Destaining Solution (30% methanol, 10% acetic acid) followed by rinses with ddH₂O. Once bands were visualized, the gel was incubated in Casting Solution (5% glycerol, 20% ethanol) and preserved in cellophane. Silver staining was performed as previously described.¹⁴ Briefly, gels were incubated with Fixing Solution (40% ethanol, 8% acetic acid) overnight followed by a 10 min incubation with 50% Fixing Solution and ddH₂O (20% ethanol, 4% acetic acid). Then, gels were incubated in ddH₂O for 10mins, followed by a 10min incubation with 0.025mM dithiothreitol (DTT). Next, gels were incubated with a 0.1% silver nitrate solution for

10mins, followed by an incubation in Developing Solution (50% formaldehyde, 1% (w/v) sodium carbonate) for several seconds until bands are visible. A 5% acetic acid solution stopped the development of bands and the gels were incubated in a 3% glycerol solution before incubating in Casting Solution (5% glycerol, 20% ethanol) and preserved in cellophane.

For immunoblot analysis, gels were transferred to a nitrocellulose membrane (0.22 μ m) and the nitrocellulose membrane was blocked in 20% (w/v) non-fat dry milk in phosphate-buffered saline (2.70 mM KCl, 137 mM NaCl, 4.00 mM Na₂HPO₄, 1.80 mM KH₂PO₄, pH 7.4) with 0.05% (v/v) Tween-20 (PBST) at room temperature for 3 hours. After three washes with PBST, the membranes were incubated with α -myc (1:10,000) antibody in 5% (w/v) non-fat dry milk in PBST at 4°C overnight. After an additional three washes with PBST, the membranes were incubated with goat α -mouse IgG-HRP (1:2,500) antibody in 5% (w/v) non-fat dry milk in PBST at room temperature for 1 hour. The membrane was washed three more times with PBST and visualized with enhanced chemiluminescence (ECL) on a GeneGnome XRQ.

2.2.6 Purification of His-Ste14

Purification of His-Ste14 was performed as previously described.⁵ All detergents and percentages (w/v) utilized in this study can be found in Table 2.2. Briefly, the crude membranes overexpressing the protein of interest were solubilized in a solution of 0.02M imidazole, detergent (Table 2.2), and 1xEQ buffer (50mM Na₂HPO₄, 300 mM NaCl, pH 6.0 or 7.2) with the addition of 10% glycerol and 2mM AEBSF. Only purification preparations with DDM was at pH 6.0, all other detergents were at pH 7.2. For the solubilization reaction, crude membranes were at a final concentration of 5 mg/mL. The reaction was rocked at 4°C for 1 hr. The resulting mixture was centrifuged at 100,000 xg for 45 mins. The soluble fraction was incubated for 1 hour at 4°C with Talon® cobalt metal affinity beads equilibrated in 1xEQ buffer with the addition of 10% glycerol and 2mM AEBSF. 2mL of resin was used for every 50-75mg of His-Ste14 in crude membranes. After the protein was bound to the resin, the solution was centrifugated at 350 xg for 2 mins at 4°C and the remaining buffer was removed via aspiration. Next, the beads underwent four wash steps: 1xEQ buffer was added to the resin and was allowed to rock at 4°C for 10 mins followed by centrifugation at 350 x g for 2 min before aspiration of excess buffer. The resin was washed twice with buffer A (1xEQ buffer with the addition of detergent [Table 2.2] and 40mM imidazole), once with buffer B (1xEQ buffer with the addition of detergent [Table 2.2], 500mM KCl, and 40mM

imidazole), and once with buffer C (1xEQ buffer with the addition of detergent [Table 2.2], 500mM KCl, and 40mM imidazole). After the final wash, buffer C was not aspirated, but instead and loaded onto a Poly-Prep® chromatography column. His-Ste14 was eluted from the cobalt resin with elution buffer (detergent [Table 2.2], 1M imidazole, and 2mM AEBSF). Collected protein was concentrated at 5000 xg in a 30,000 MWCO concentrator at 4°C until the final volume was ~1mL. Concentrated protein was filtered in centrifugal filter units (0.22 µm) GV Durapore® by centrifuging at 12,000 xg for 2 mins at 4°C. Once filtered, the His-Ste14 mutants were applied to a Superdex S200 column pre-equilibrated with filtered 1xEQ buffer (with the addition of detergent (Table 2.2) and 2mM DTT) at 4°C and spectra was collected at 280nm. Fractions containing purified protein were pooled, concentrated to a volume under 500µL in a new 30,000 MWCO concentrator, and aliquoted. Samples were flash frozen in liquid nitrogen and stored at -80°C until further experimentation. Pure protein concentration was determined using an Amido Black protein assay.¹³ For experiments utilizing 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer and Tris-HCl buffer, all other conditions are the same as stated above except for the substitution of 1xEQ buffer at pH 7.2.

2.3 Results and Discussion

2.3.1 Expression in Yeast vs Insect Cells Lines

His-Ste14 was expressed in yeast, *Saccharomyces cerevisiae* (*S. cerevisiae*), and two insect cell lines (Sf9s and High Fives). Crude membranes were isolated and equal concentrations of total protein was run on an SDS-PAGE and band intensity was quantified using ImageJ (Figure 2.1). Crude membranes isolated from *S. cerevisiae* had the lowest expression of His-Ste14, followed by High Five cells, and Sf9 cells. Sf9 cells consistently had between 4 and 5-fold higher His-Ste14 crude membrane expression than yeast and 2-fold higher than High Five cells (Figure 2.1B). All cell lines produced catalytically active protein (data not shown). Due to the high expression levels of His-Ste14 in Sf9 cells, this expression system was used for all other experiments in this chapter. Immunoblot analysis was performed in triplicate.

The quantification method of His-Ste14 expression in various organisms we used could be improved. Since some of the immunoblot bands appeared blown out or at their maximum capacity of illumination, it is necessary to repeat the immunoblot analysis at various concentrations of crude

membranes. Thus, we can determine the ratios of expression of Ste14 in yeast compared to insect cells more accurately.

2.3.2 Optimization of Size Exclusion Chromatography of His-Ste14

Since large quantities of His-Ste14 could be produced, we needed to optimize the purification process, including the type of buffer and pH. First, three buffers were assessed for the purification of His-Ste14 via affinity chromatography and size exclusion chromatography (SEC): a phosphate buffer (1xEQ, 50mM Na₂HPO₄, 300mM NaCl, pH 7.2), a HEPES buffer (20mM HEPES, 50mM NaCl, pH 7.2), and a Tris-HCl buffer (25mM Tris-HCl, 150mM NaCl, pH 7.2) (Figure 2.2A). The purification process was identical apart from the buffer composition; DDM was the detergent used for solubilization (Table 2.2, Figure 2.2B). In the past, phosphate buffer was most commonly used to purify His-Ste14, however, this buffer is unfavorable for some biophysical techniques.¹⁵ Thus, HEPES and Tris-HCl were selected for potential purification buffers. Specifically, HEPES was a buffer of interest because it was used to successfully purify active His-Ste24.^{15,16} It is important to note that His-Ste14 acts as a dimer and therefore we aim to purify the enzyme in that oligomeric state.²¹

The phosphate buffer (Figure 2.2, blue) eluted His-Ste14 earlier than the other buffers at ~7-10mL. The sharp peak at 7mL appears to be very close to the void volume due to samples aggregating. The smaller peak at ~14-15mL is predicted to be the monomeric form of His-Ste14. The HEPES buffer produced a much broader peak compared to the phosphate buffer, eluted protein is between ~9-14mL (Figure 2.2, green). Similarly, the Tris-HCl buffer (Figure 2.2, purple) has several very broad peaks, similar to that of HEPES. This could indicate a slight break down in the protein, since all other buffer components remained the same. Although HEPES and Tris-HCl buffers reduce the aggregation of His-Ste14, the SEC peaks in these buffers are extremely broad and do not result in a homogenous sample.^{17,18} It is important to note that we cannot determine the molecular weights of the SEC peaks with standards because His-Ste14 is purified in detergent, which can alter its elution volume compared to soluble proteins. Multi-angle light scattering (MALS) was used to determine molecular weight of these peaks (see Chapter 5 and Publications).

pH was also an important factor for the purification of His-Ste14. All previous purification had been performed at a biological pH of 7.2. However, biological pH is not always the optimal pH when working with an *in vitro* system, especially with detergent involvement.¹⁹ Altering pH

of the buffer solution for a membrane protein purification can lower the amount of aggregation in the sample. During these experiments, His-Ste14 was purified in 1xEQ buffer and DDM detergent, pH 6.0. Under these conditions, there was a dramatic change in the SEC chromatograph (Figure 2.3A). Only one sharp peak eluted at ~13mL, which is far from the void volume. Fractions were collected and concentrated from the peak of interest and the methyltransferase activity of the sample was tested (Figure 2.3B). His-Ste14 purified at pH 6.0 had $96.5\% \pm 9.9$ activity compared to His-Ste14 purified at pH 7.2, performed in triplicate (\pm S.D.). Immunoblots and Coomassie stained SDS-PAGE confirmed normal expression levels and purity compared to purified His-Ste14 pH 7.2, respectively (Figure 2.3C & D). Due to these results, all size exclusion chromatography (SEC), multi-angle light scattering (MALS), and small angle light scattering (SAXS) experiments performed in the “Appendix” and “Publication” sections of this dissertation, were performed at pH 6.0.

2.3.3 Detergent Screening of Pure His-Ste14

Previously, DDM was the detergent most commonly used to purify His-Ste14 due to its mild nature and its lower CMC, making it easier to perform buffer exchanges by dialysis or gel filtration (Table 2.1).⁷ However, as previously mentioned, DDM forms very large micelles resulting in large protein-detergent complexes, which is unfavorable for most biophysical experiments.⁷ Mild detergents are defined as having a neutral charge (rather than zwitterionic or charged), a large head group, and a long alkyl tail (12 carbons is considered long, while 8 is considered short).⁷ We solubilized His-Ste14 with six different detergents, then evaluated their protein expression levels, purity, and methyltransferase activity. LMNG and C₁₂E₇ were the only other detergents tested that successfully purified His-Ste14 stably and actively (Figure 2.5, Figure 2.6, respectively). All experiments discussed in section 2.3.3 were purified in 1xEQ buffer at pH 7.2. It is important to note that these experiments were performed before we found that pH 6.0 is more optimal to purify His-Ste14.

LMNG was used to purify His-Ste14 as described in the methods via SEC. Two major peaks were seen when evaluating the 280nm spectra; one peak from ~8-14mL and one at ~15-16mL (Figure 2.4). LMNG is an amphipathic molecule that has been found to have higher membrane protein stabilization efficacy than DDM.²⁰ Interestingly the LMNG molecule is comprised of two branched DDM head groups, a particularly rigid structure (Figure 2.4B). The

larger of the eluted peaks was the majority of the His-Ste14 sample, while the smaller had a much small quantity of the protein. Immunoblots and silver stained SDS-PAGE confirmed normal expression levels and purity compared to His-Ste14 purified in DDM, respectively (Figure 2.4C & D). We predict that the peak at ~15-16mL is monomeric His-Ste14, where the peak from ~8-14mL is the dimer. His-Ste14 does not seem to be aggregating because there is no protein in the void volume. Although solubilization with LMNG results in a broad SEC peak, it may be worth investigating the use of this detergent in 1xEQ buffer at pH 6.0. LMNG may be a useful detergent for any future structural analysis like X-ray crystallography and cryo-EM.

Although C₁₂E₇ typically stabilizes β -barrel transmembrane proteins, the linear, neutral detergent successfully extracted His-Ste14 from the membrane followed by purification with affinity chromatography and SEC (Figure 2.5A). Two major peaks were identified with SEC, one eluted at ~8-12mL and the second eluted at ~14-15mL. The chromatograph appears identical to His-Ste14 purified with DDM in 1xEQ buffer at pH 7.2, which was previously found to be the most stable conditions to purify His-Ste14 (Figure 2.3).⁵ Similarly, we believe that the sharp peak at ~8mL was near the void volume, which could be due to sample aggregation. The smaller peak starting at ~14mL was predicted to be the monomeric form of His-Ste14. C₁₂E₇ is one of the longest detergents synthesized in the polyoxyethylene family (Figure 2.5B). The longer chain and neutral charge of this detergent results in milder characteristics, making C₁₂E₇ a candidate for His-Ste14 purification. Immunoblots and Coomassie stained SDS-PAGE confirmed normal expression levels and purity compared to His-Ste14 purified in DDM, respectively (Figure 2.5C & D). The sample predicted to be monomeric had an additional band during immunoblot analysis, which could be due to the protein being less stable as a monomer compared to its native dimeric state.²¹

DMNG was unsuccessful for the active purification of His-Ste14. DMNG has similar structural and biochemical characteristics with LMNG, except for shorter hydrophobic alkyl chains (Figure 2.6A). His-Ste14 was successfully extracted from the membrane during solubilization due to its similar expression levels and purity compared to His-Ste14 purified in DDM (Figure 2.6B & C). However, since the activity of His-Ste14 in DMNG was not conserved, the protein was never purified via SEC (Figure 2.9). Interestingly, Diver *et al.* (2018) utilized DMNG in a Tris-HCl buffer (pH 7.5) to purify beetle Icmt for X-ray crystallography.²² Potentially, the increased protein size of beetle Icmt (the addition of two alpha helices) could cause a stability preference to another detergent.

C₈E₄ is one of the smallest detergents in the polyoxyethylene family (Figure 2.7A). C₈E₄ has been shown to successfully purify and crystalize many proteins including β -barrel outer membrane proteins from Gram negative bacteria as well as other membrane proteins.^{23,24} However, due to its shorter carbon chain, it is more harsh than C₁₂E₇ and has more similar properties to the detergent n-octyl- β -D-glucopyranoside (OG). Anderson *et al.* already discovered that OG was ineffective for purifying His-Ste14.⁵ This could account for the slight degradation of His-Ste14 visualized with immunoblot analysis and Coomassie staining of SDS-PAGE, as well as the inactivity of His-Ste14 in C₈E₄ experiments (Figure 2.7B & C, Figure 2.9).

It is not surprising that LDAO, the only zwitterionic detergent tested, was not successful in actively purifying His-Ste14 (Figure 2.7A). It is a relatively harsh detergent and only ~20% of membrane proteins are resistant to denaturing effects of this detergent.²⁵ In addition, LDAO is more harsh than OG, which has already been shown to be ineffective for purifying His-Ste14.⁵ Although immunoblotting and Coomassie stained SDS-PAGE show equal expression and purity to that of His-Ste14 purified in DDM, the protein did not remain activity (Figure 2.8, Figure 2.9).

2.4 Future Directions

Now that we have established an efficient method to produce and purify His-Ste14, future experiments would benefit from expression His-Ste14 from the baculovirus expression system. The most advantageous mutant to create would be the His-Ste14-TA baculovirus followed by other His-Ste14-TA single cysteine mutants needed for experiments discussed in chapter 3.

Additionally, it would be advantageous to purify His-Ste14 in HEPES (20mM HEPES, 50mM NaCl), and Tris-HCl buffer (25mM Tris-HCl, 150mM NaCl) at pH 6.0. Since we found the best chromatograph spectra of pure His-Ste14 at pH 6.0 in 1x EQ buffer, it would be advantageous to purify His-Ste14 at pH 6.0 in HEPES or Tris-HCl as well. Also, it would be advantageous to observe the effects of pH 6.0 buffer on each of the detergents tested above. All of these experiments could be performed to further optimize the purification of His-Ste14 and His-Ste14 mutants in future experiments.

Table 2.1 Characteristics of detergents used in this study^{3,26,27}

| Detergent | CMC (mM) | CMC (wt%) | Micelle Size (kDa) | Characteristics |
|---|----------|-----------|--------------------|-----------------|
| <i>N</i>-Dodecyl-β-D-maltopyranoside (DDM) | 0.17 | 0.0087% | 65-70 | Non-ionic |
| Lauryl maltose neopentyl glycol (LMNG) | 0.010 | 0.0010% | 91 | Non-ionic |
| Heptaethylene glycol monododecyl ether (C₁₂E₇) | 0.090 | 0.0048% | -----* | Non-ionic |
| Decyl maltose neopentyl glycol (DMNG) | 0.024 | 0.0020% | -----* | Non-ionic |
| Lauryldimethylamine-<i>N</i>-oxide (LDAO) | 1.0-2.0 | 0.023% | 21.5 | Zwitterionic |
| Tetraethylene glycol monododecyl ether (C₈E₄) | 8.0 | 0.10% | -----* | Non-ionic |

*Micelle size has not been determined for this detergent

Table 2.2 Percent (w/v) of detergents used for purification.

| Detergent | Solubilization % (w/v) | Buffer A % (w/v) | Buffer B % (w/v) | Buffer C % (w/v) | Elution % (w/v) | FPLC % (w/v) |
|---|---------------------------|---------------------|---------------------|---------------------|--------------------|-----------------|
| <i>N</i>-Dodecyl-β-D-maltopyranoside (DDM) | 1% | 1% | 1% | 0.1% | 0.1% | 0.05% |
| Lauryl maltose neopentyl glycol (LMNG) | 1% | 1% | 1% | 0.1% | 0.1% | 0.05% |
| Heptaethylene glycol monododecyl ether (C₁₂E₇) | 1% | 1% | 1% | 0.1% | 0.1% | 0.04% |
| Decyl maltose neopentyl glycol (DMNG) | 1% | 1% | 1% | 0.1% | 0.1% | N/A* |
| Lauryldimethylamine-<i>N</i>-oxide (LDAO) | 1% | 0.1% | 0.1% | 0.05% | 0.05% | N/A* |
| Tetraethylene glycol monododecyl ether (C₈E₄) | 1% | 1% | 1% | 0.6% | 0.6% | N/A* |

*Solubilization and purification of His-Ste14 with designated detergents did not result in active protein after elution from the Talon® cobalt metal affinity resin and was therefore never analyzed via SEC.

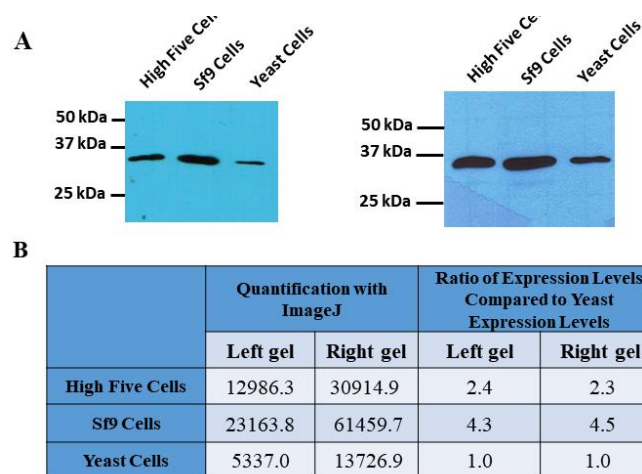


Figure 2.1 Expression of His-Ste14 extracted from yeast and insect cell lines by immunoblotting. (A) Protein samples (1 μ g) were run on a 10% SDS-PAGE followed by immunoblot analysis with α -myc (1:10,000) and goat α -mouse IgG-HRP (1:2,500). Immunoblots were visualized with ECL. Two examples are shown. (B) ImageJ 1.x was used to quantify the intensity of each band and was used to create a ratio between each expression system compared to the expression levels in yeast.²⁸

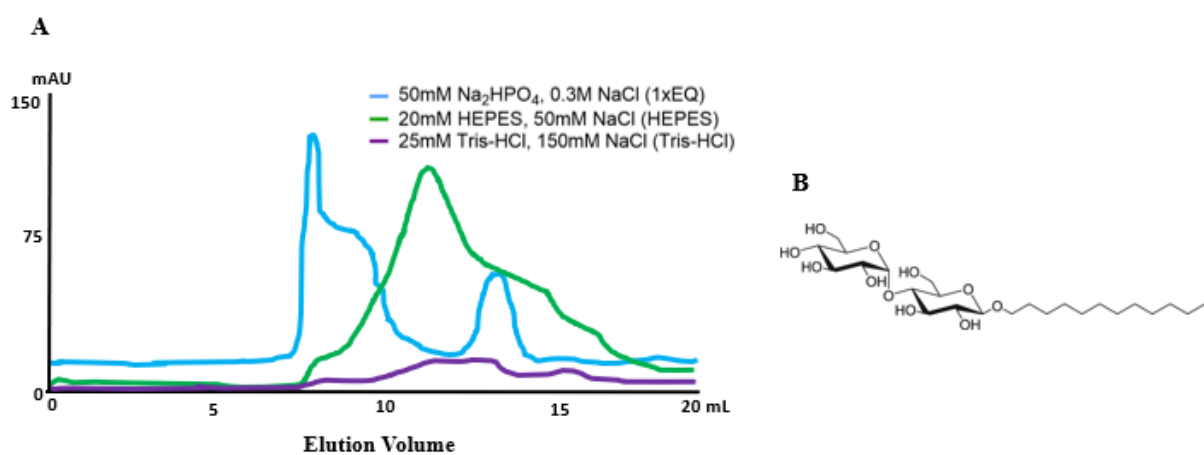


Figure 2.2 Purification of His-Ste14 purified in various buffers.

(A) Size exclusion chromatography (SEC) analysis of His-Ste14 purified in DDM at pH 7.2 in phosphate buffer (1xEQ, 50mM Na₂HPO₄, 300mM NaCl), HEPES buffer (20mM HEPES, 50mM NaCl), and Tris buffer (25mM Tris-HCl, 150mM NaCl), depicted in blue, green, and purple, respectively. (B) Structure of single DDM molecule.

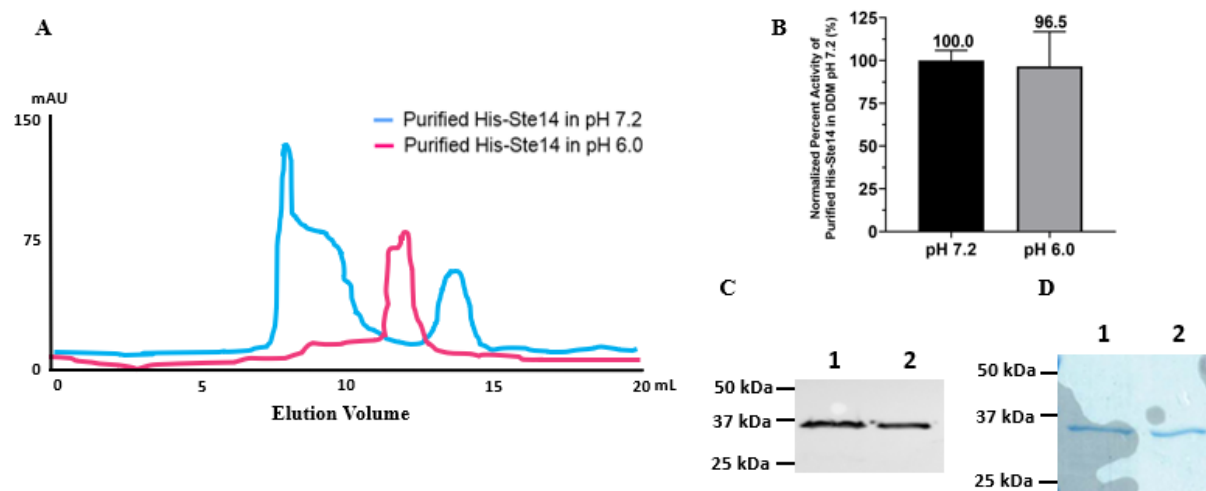


Figure 2.3 Purification of His-Ste14 with DDM detergent at pH 7.2 and 6.0. (A) Size exclusion chromatography (SEC) analysis of His-Ste14 purified in 1xEQ buffer and DDM at pH 7.2 (blue) or pH 6.0 (pink). (B) *In vitro* vapor diffusion methyltransferase activity assay of His-Ste14 after SEC purification at pH 7.2 and 6.0. Each experiment had three samples per condition and each experiment was performed in triplicate \pm S.D. (C) Expression levels of purified His-Ste14 indicated by immunoblotting: protein samples (0.1 μ g) were run on a 10% SDS-PAGE followed by immunoblot analysis with α -myc (1:10,000) and goat α -mouse IgG-HRP (1:2,500). Immunoblots were visualized with ECL. (D) Determination of purity of protein samples via Coomassie on 10% SDS-PAGE. (C & D) Samples in each lane were as followed: Lane 1 – Purified His-Ste14 in DDM pH 7 and Lane 2 – Purified His-Ste14 in DDM pH 6.

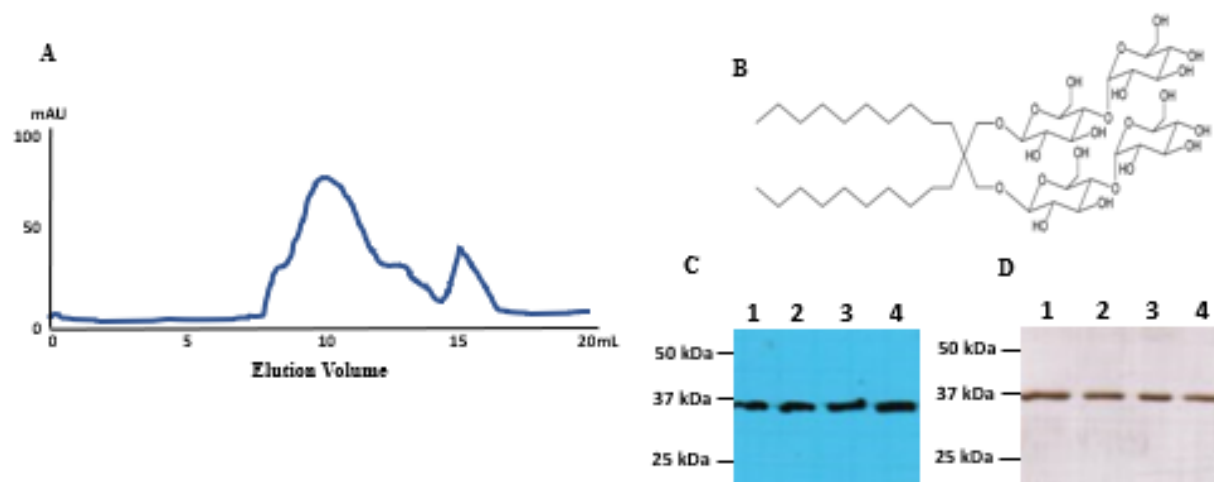


Figure 2.4 Purification of His-Ste14 with LMNG detergent at pH 7.2.

(A) Size exclusion chromatography (SEC) analysis of His-Ste14 solubilized in LMNG. (B) Structure of single LMNG molecule, demonstrating that the structure is comprised of two DDM molecules. (C) Expression levels of purified His-Ste14 indicated by immunoblotting: protein samples (0.1 μ g) were run on a 10% SDS-PAGE followed by immunoblot analysis with α -myc (1:10,000) and goat α -mouse IgG-HRP (1:2,500). Immunoblots were visualized with ECL. (D) Determination of purity of protein samples via silver staining on 10% SDS-PAGE. (C & D) Samples in each lane were as followed: Lane 1 – Purified His-Ste14 solubilized in DDM (positive control), Lane 2 – Purified His-Ste14 solubilized in LMNG after elution from Talon® cobalt metal affinity resin, Lane 3 - Purified His-Ste14 in LMNG, peak that eluted at ~8-14mL, and Lane 4 - Purified His-Ste14 solubilized in LMNG, peak that eluted at ~15-16mL.

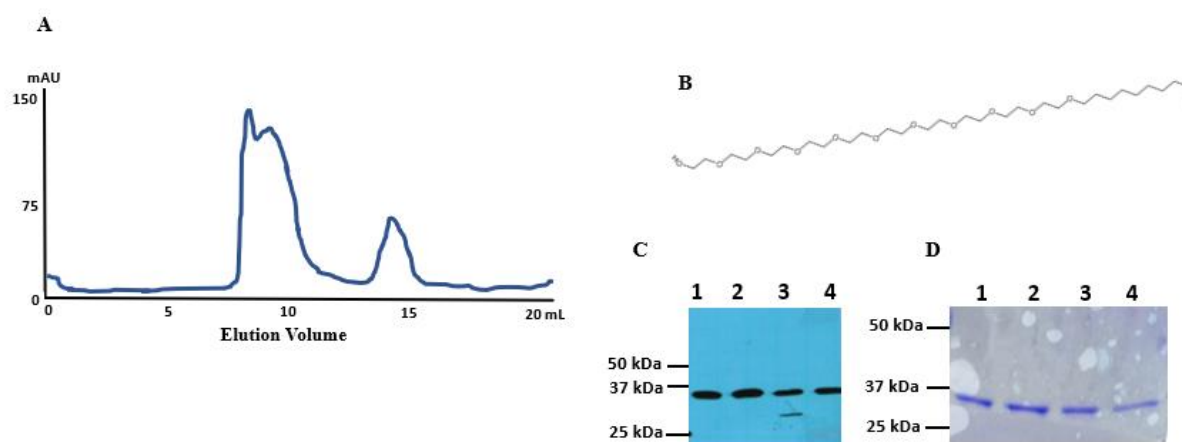


Figure 2.5 Purification of His-Ste14 with C₁₂E₇ detergent at pH 7.2.

(A) Size exclusion chromatography (SEC) analysis of His-Ste14 solubilized in C₁₂E₇. (B) Structure of single C₁₂E₇ molecule. (C) Expression levels of purified His-Ste14 indicated by immunoblotting: protein samples (0.1 μg) were run on a 10% SDS-PAGE followed by immunoblot analysis with α-myc (1:10,000) and goat α-mouse IgG-HRP (1:2,500). Immunoblots were visualized with ECL. (D) Determination of purity of protein samples via Coomassie staining on 10% SDS-PAGE. (C & D) Samples in each lane were as followed: Lane 1 – Purified His-Ste14 solubilized in DDM (positive control), Lane 2 – Purified His-Ste14 solubilized in C₁₂E₇ after elution from Talon® cobalt metal affinity resin, Lane 3 - Purified His-Ste14 in C₁₂E₇, peak that eluted from ~14-15mL, and Lane 4 - Purified His-Ste14 solubilized in LMNG, peak that eluted from ~8-12mL.

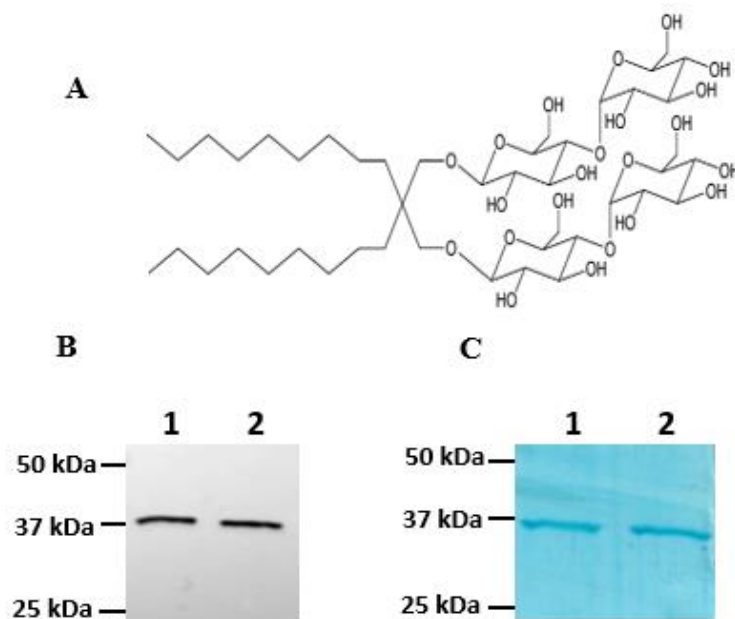


Figure 2.6 Purification of His-Ste14 with DMNG detergent at pH 7.2.

A) Structure of single DMNG molecule. (B) Expression levels of purified His-Ste14 indicated by immunoblotting: protein samples (0.1 μ g) were run on a 10% SDS-PAGE followed by immunoblot analysis with α -myc (1:10,000) and goat α -mouse IgG-HRP (1:2,500). Immunoblots were visualized with ECL. (C) Determination of purity of protein samples via Coomassie staining on 10% SDS-PAGE. (B & C) Samples in each lane were as followed: Lane 1 – Purified His-Ste14 solubilized in DDM (positive control) and Lane 2 – Purified His-Ste14 solubilized in DMNG after elution from Talon® cobalt metal affinity resin.

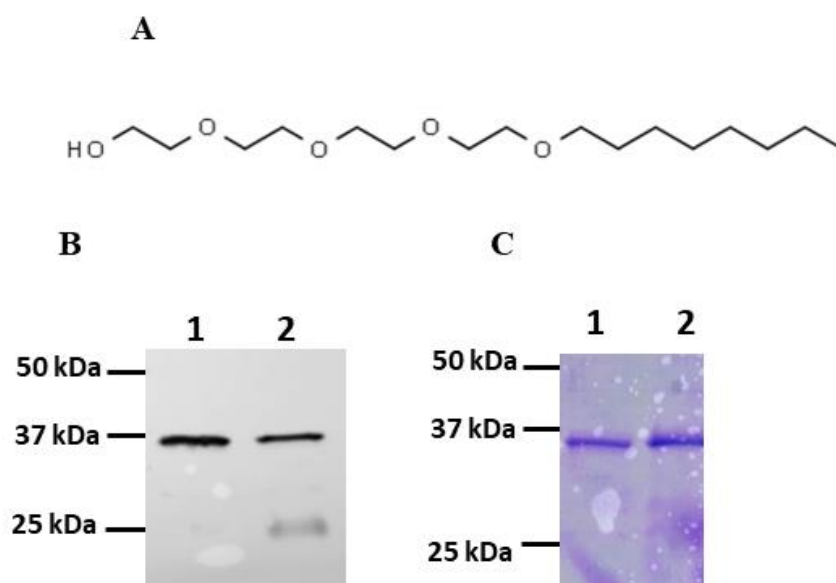


Figure 2.7 Purification of His-Ste14 with C₈E₄ detergent at pH 7.2.

A) Structure of single C₈E₄ molecule. (B) Expression levels of purified His-Ste14 indicated by immunoblotting: protein samples (0.1 μg) were run on a 10% SDS-PAGE followed by immunoblot analysis with α-myc (1:10,000) and goat α-mouse IgG-HRP (1:2,500). Immunoblots were visualized with ECL. (C) Determination of purity of protein samples via Coomassie staining on 10% SDS-PAGE. (B & C) Samples in each lane were as followed: Lane 1 – Purified His-Ste14 solubilized in DDM (positive control) and Lane 2 – Purified His-Ste14 solubilized in C₈E₄ after elution from Talon® cobalt metal affinity resin.

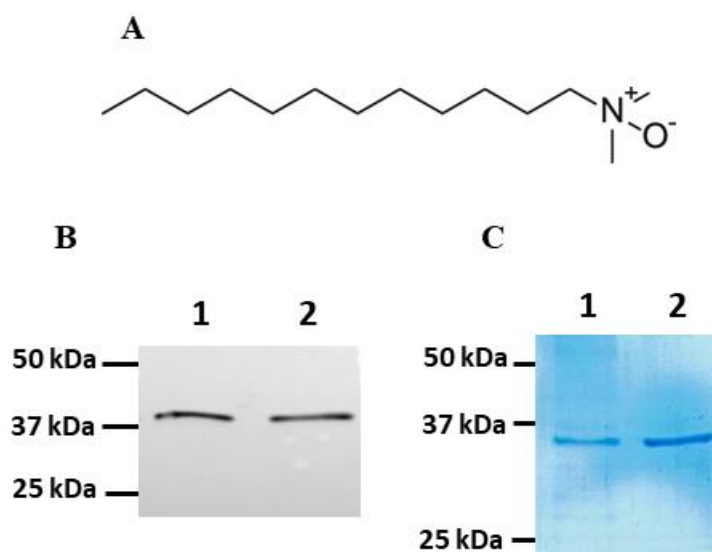


Figure 2.8 Purification of His-Ste14 with LDAO detergent at pH 7.2.

A) Structure of single LDAO molecule. (B) Expression levels of purified His-Ste14 indicated by immunoblotting: protein samples (0.1 μ g) were run on a 10% SDS-PAGE followed by immunoblot analysis with α -myc (1:10,000) and goat α -mouse IgG-HRP (1:2,500). Immunoblots were visualized with ECL. (C) Determination of purity of protein samples via Coomassie staining on 10% SDS-PAGE. (B & C) Samples in each lane were as followed: Lane 1 – Purified His-Ste14 solubilized in DDM (positive control) and Lane 2 – Purified His-Ste14 solubilized in LDAO after elution from Talon® cobalt metal affinity resin

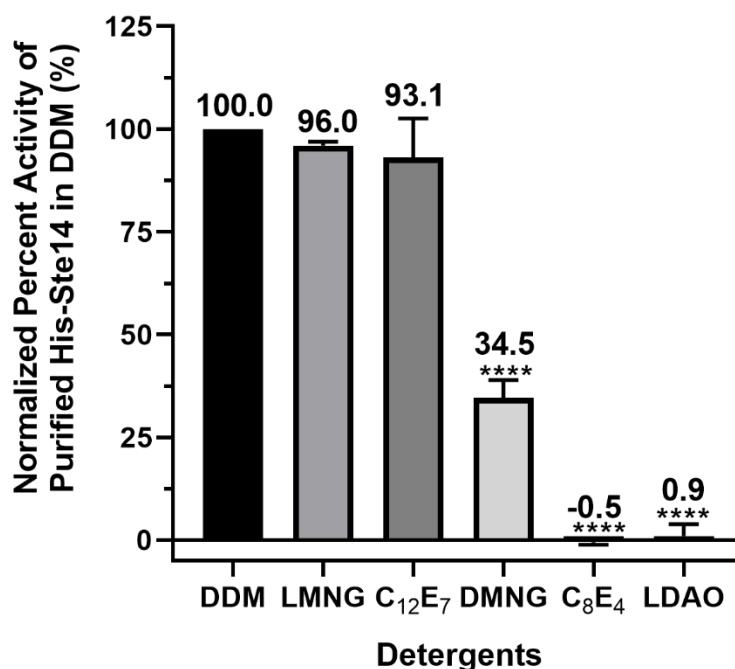


Figure 2.9 *In vitro* vapor diffusion methyltransferase activity assay of His-Ste14 solubilized in various detergents at pH 7.2.

All specific activities were characterized as pmol of methyl groups from [¹⁴C]-SAM transferred to the AFC substrate per minute of assay per mg of purified His-Ste14. These activities were calculated after subtracting out background counts from the average of the negative control samples (ddH₂O added to buffer in place of pure protein). Percent activity compared to His-Ste14 purified in DDM was calculated for each purification experiment with a different detergent. Each experiment was performed in triplicate \pm S.D. (****, $p \leq 0.0001$; based on one-way ANOVA followed by Dunnett's multiple comparisons test vs. His-Ste14 purified in DDM).

2.5 References

- 1 Hrycyna, C. A., Sapperstein, S. K., Clarke, S. & Michaelis, S. The *Saccharomyces cerevisiae* Ste14 gene encodes a methyltransferase that mediates C-terminal methylation of a-factor and Ras proteins. *Embo Journal* **10**, 1699-1709, doi:10.1002/j.1460-2075.1991.tb07694.x (1991).
- 2 Krull, I. & Rathore, A. Popular cell expression systems in the biotechnology industry. *Lc Gc North America* **28**, 800-809 (2010).
- 3 le Maire, M., Champeil, P. & Moller, J. V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochimica Et Biophysica Acta-Biomembranes* **1508**, 86-111, doi:10.1016/s0304-4157(00)00010-1 (2000).
- 4 Breyton, C. *et al.* Assemblies of lauryl maltose neopentyl glycol (LMNG) and LMNG-solubilized membrane proteins. *Biochimica Et Biophysica Acta-Biomembranes* **1861**, 939-957, doi:10.1016/j.bbamem.2019.02.003 (2019).
- 5 Anderson, J. L., Frase, H., Michaelis, S. & Hrycyna, C. A. Purification, functional reconstitution, and characterization of the *Saccharomyces cerevisiae* isoprenylcysteine carboxymethyltransferase Ste14p. *Journal of Biological Chemistry* **280**, 7336-7345, doi:10.1074/jbc.M410292200 (2005).
- 6 Anatrache Products, LLC. Detergents and their uses in membrane protein science. (2019).
- 7 Prive, G. G. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397, doi:10.1016/j.ymeth.2007.01.007 (2007).
- 8 Boivin, S., Kozak, S. & Meijers, R. Optimization of protein purification and characterization using Thermofluor screens. *Protein Expression and Purification* **91**, 192-206, doi:10.1016/j.pep.2013.08.002 (2013).
- 9 Elble, R. A Simple and efficient procedure for transformation of yeasts. *Biotechniques* **13**, 18-20 (1992).
- 10 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254 (1976).
- 11 Hrycyna, C. A. & Clarke, S. Farnesyl cysteine C-terminal methyltransferase activity is dependent upon the ste14 gene-product in *Saccharomyces-cerevisiae*. *Molecular and Cellular Biology* **10**, 5071-5076, doi:10.1128/mcb.10.10.5071 (1990).

- 12 Anderson, J. L., Henriksen, B. S., Gibbs, R. A. & Hrycyna, C. A. The isoprenoid substrate specificity of isoprenylcysteine carboxylmethyltransferase - Development of novel inhibitors. *Journal of Biological Chemistry* **280**, 29454-29461, doi:10.1074/jbc.M504982200 (2005).
- 13 Schaffne.W & Weissman.C. Rapid, sensitive, and specific method for determination of protein in dilute-solution. *Analytical Biochemistry* **56**, 502-514, doi:10.1016/0003-2697(73)90217-0 (1973).
- 14 Rabilloud, T. Mechanisms of protein silver staining in polyacrylamide gels - A 10-year synthesis. *Electrophoresis* **11**, 785-794, doi:10.1002/elps.1150111003 (1990).
- 15 Scarff, C. A., Fuller, M. J. G., Thompson, R. F. & Iadaza, M. G. Variations on negative stain electron microscopy methods: Tools for tackling challenging systems. *Jove-Journal of Visualized Experiments*, doi:10.3791/57199 (2018).
- 16 Pryor, E. E. *et al.* Structure of the integral membrane protein CAAX protease Ste24p. *Science* **339**, 1600, doi:10.1126/science.1232048 (2013).
- 17 Hunte, C., Von Jagow, G. & Schagger, H. *Membrane protein purification and crystallization*. 2nd edn, 316 (2003).
- 18 Von Jagow, G. & Revzin, A. *A practical guide to membrane protein purification*. 1 edn, Vol. 2 (2013).
- 19 GE Healthcare, Superdex High-performance Columns. Data file 18-1163-79 AD (2007).
- 20 Cho, K. H. *et al.* Maltose neopentyl glycol-3 (MNG-3) analogues for membrane protein study. *Analyst* **140**, 3157-3163, doi:10.1039/c5an00240k (2015).
- 21 Griggs, A. M., Hahne, K. & Hrycyna, C. A. Functional oligomerization of the *Saccharomyces cerevisiae* Isoprenylcysteine Carboxyl Methyltransferase, Ste14p. *Journal of Biological Chemistry* **285**, 13380-13387, doi:10.1074/jbc.M109.061366 (2010).
- 22 Diver, M. M., Pedi, L., Koide, A., Koide, S. & Long, S. B. Atomic structure of the eukaryotic intramembrane RAS methyltransferase ICMT. *Nature* **553**, 526-+, doi:10.1038/nature25439 (2018).
- 23 Charbonnier, F., Kohler, T., Pechere, J. C. & Ducruix, A. Overexpression, refolding, and purification of the histidine-tagged outer membrane efflux protein OprM of *Pseudomonas aeruginosa*. *Protein Expression and Purification* **23**, 121-127, doi:10.1006/prep.2001.1473 (2001).
- 24 Rouse, S. L. *et al.* Purification, crystallization and characterization of the *Pseudomonas* outer membrane protein FapF, a functional amyloid transporter. *Acta Crystallographica Section F-Structural Biology Communications* **72**, 892-896, doi:10.1107/s2053230x16017921 (2016).

- 25 *International Tables for Crystallography*. Vol. F (Dordrecht, 2001).
- 26 Yang, Z. *et al.* Membrane protein stability can be compromised by detergent interactions with the extramembranous soluble domains. *Protein Science* **23**, 769-789, doi:10.1002/pro.2460 (2014).
- 27 Breibeck, J. & Rompel, A. Successful amphiphiles as the key to crystallization of membrane proteins: Bridging theory and practice. *Biochimica Et Biophysica Acta-General Subjects* **1863**, 437-455, doi:10.1016/j.bbagen.2018.11.004 (2019).
- 28 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675, doi:10.1038/nmeth.2089 (2012).

CHAPTER 3. ASSESSMENT OF STE14 CO-FACTOR BINDING DYNAMICS USING ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY

3.1 Introduction

To date, few molecular details are known about the mechanism of Ste14 except that it must accommodate chemically diverse methyl donor and acceptor molecules: the hydrophilic co-factor *S*-adenosyl-*L*-methionine (SAM) and a lipophilic isoprenylated CaaX protein substrate, respectively. Ultimately, it is important to understand in mechanistic detail how such an enzyme mediates catalysis at the membrane/cytosol interface. The goal of this research is to understand the structure-function relationships of co-factor and substrate recognition by the Ste14 enzyme and to elucidate the structural dynamics of the SAM binding and *S*-adenosyl-*L*-homocysteine (SAH) release mechanism.

As mammalian Icmts have proven difficult to actively purify, we used Ste14 as our model enzyme. Ste14 was discovered in *Saccharomyces cerevisiae* and shares 63% similarity and 41% identity with human Icmt (hIcmt).^{1,2} Additionally, hIcmt and Ste14 are believed to function under the same mechanism because hIcmt expressed in yeast complements a *ste14Δ* deletion.² Currently, the two crystal structures of Icmts from *Methanosarcina acetivorans* (Ma-Icmt) and *Tribolium castaneum* (beetle Icmt) predict SAM binding to the portion of the protein made up by the C-terminal region of the protein.^{3,4} Specifically, in Ste14, this binding is proposed to occur between residues 115 and 239, mostly in the cytosolic regions. As there is no access port for the co-factor to enter in the folded structure, we hypothesize that conformational changes must occur between residues 115 and 239 in order for SAM to bind, exchange its methyl group to produce SAH, and complete the reaction.^{3,4} Based on the beetle Icmt crystal structure model from Diver *et al.* (2018), Loop 4 and the C-terminal tail of Ste14 may act as a hinge region that can open and close to allow SAM to enter or to let SAH be released back into the cytosol.³ However, this mechanism has yet to be proven because of the limitations of crystal structure models. During X-ray crystallography, the protein is frozen in one conformation, producing only static images.

Another common technique used to study protein dynamics is nuclear magnetic resonance (NMR) spectroscopy, which measures the interactions of nuclei in a magnetic field. However, NMR structural studies on membrane proteins are challenging with micelle complexes that cause

increasing spectral linewidth.⁵ Thus, we propose to use site-directed spin labeling (SDSL) and electron paramagnetic resonance (EPR) spectroscopy to evaluate structural rearrangements of the C-terminal and cytosolic loops upon co-factor binding. Additionally, EPR has better sensitivity and the sample does not need to be as concentrated than NMR.⁶

EPR spectroscopy is used to understand protein dynamics and the distances of their conformational changes. EPR spectroscopy measures the absorption of microwave radiation corresponding to the energy splitting of unpaired electrons in a magnetic field. The simplest EPR system, which was performed in this chapter, consists of a single unpaired electron spin residing in a molecular orbital placed in a fixed microwave frequency, but a varying magnetic field. The EPR transition occurs when the energy separation between the two electron spin states matches the constant microwave energy, which is known as resonance.^{6,7} Unfortunately, most biological systems do not have unpaired electrons. Thus, site-directed spin labeling (SDSL) is used to introduce a cysteine residue into a cysteine-less protein.⁸⁻¹⁰ Then, a nitroxide probe is introduced into a biological sample by forming a covalent disulfide bond with the sulfhydryl group of the newly introduced cysteine. The degree of mobility of the spin label in the magnetic field, which can result from side chain rotations, local backbone fluctuations and protein tumbling, reflects the environment of the spin label side chain.^{10,11} Sharp peaks are indicative of free, unrestricted spin label, whereas broad peaks are characteristic of an immobilized or buried probe.¹² In this way, SDSL-EPR can be used to understand if residues of Ste14 are free moving or buried inside the protein before and after SAM binding.

To assess the dynamics of His-Ste14, the nitroxide spin label probe (1-oxy-2,2,5,5-tetramethylpyrroline-3-methyl)-methanethiosulfonate) or MTSL was attached to a single cysteine residue positioned in defined regions within Loop 4 and the C-terminus. Mutants were incorporated into cytosolic amino acids because residues within the intermembrane region would be inaccessible to MTSL. The three endogenous cysteine residues (C99, C121, and C126) were mutated to alanine, resulting in His-Ste14 “Triple A”. These three residues caused little reduction to the methyltransferase activity levels of wild-type Ste14 despite being located within the membrane in transmembrane (TM) domains 3 and 4.² Site-directed mutagenesis was then used to introduce cysteine residues at points of interest within the cytosolic loops of Ste14 followed by their purification and linkage with MTSL. The pure labeled protein was tested for activity and monitored by EPR for SAM-induced changes in mobility of the spin labelled side chain. EPR

spectra that varied indicated that a large conformation change was occurring at the site of the cysteine mutant, while no changes in the EPR spectra indicated that the cysteine underwent minimal or no changes in its environment. Two mutations were assessed with this technique, His-V212C-TA and His-K231C-TA. Thus, we determined a successful technique to better understand the mechanism of Ste14 and the IcmT family.

3.2 Materials and Methods

3.2.1 Materials

All oligonucleotides were designed by Hrycyna Lab members and purchased from Integrated DNA Technologies (IDT) (Coralville, IA). GoTaq® Green Polymerase Master Mix, LigaFast™ Rapid DNA Ligation System, and Wizard® SV Gel and PCR Clean-up System were purchased from Promega (Madison, WI). All restriction enzymes and calf intestinal alkaline phosphatase (CIP) were purchased from New England Biolabs (Beverly, MA). DH5α chemically competent *E. coli* cells, the monoclonal α-myc IgG, the polyclonal goat-α-rabbit IgG HRP conjugate and the polyclonal goat-α-mouse IgG HRP conjugate were purchased from Invitrogen Life Sciences (Carlsbad, CA). GeneJET Plasmid Miniprep Kit and Enhanced chemiluminescence (ECL) were purchased from ThermoFisher (Waltham, MA). BigDye® Terminator v3.1 was obtained from Applied Biosystems (Foster City, CA) through the Purdue University Genomics Facility. Single stranded salmon testes DNA, 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO), and S-adenosyl-L-methionine were purchased from Sigma Aldrich (St. Louis, MO). Yeast parent strains SM1058 and SM1188 and the α-Ste14 antibody were gifts from Dr. Susan Michaelis (Johns Hopkins School of Medicine, MD). GeneGnome XRQ was purchased from Syngene (Fredrick, MD). 4-(2-aminoethyl)-benzenesulfonylfluoride hydrochloride (AEBSF) and dithiothreitol (DTT) were purchased from Gold Biotechnology (St. Louis, MO). N-Dodecyl-β-D-maltopyranoside (DDM) was purchased from Anatrace (Maumee, OH). The substrate N-acetyl-S-farnesyl-L-cysteine (AFC) was purchased from Enzo Life Sciences (Farmingdale, NY). Radioactively labelled S-adenosyl-L-methionine ([¹⁴C]-SAM) was purchased from Perkin Elmer (Waltham, MA). Nitrocellulose membrane (0.22 μm) and Superdex S200 column were purchased from GE Healthcare Life Sciences (Pittsburg, PA). Talon® cobalt metal affinity resin was purchased from Clontech (Mountain View, CA). Amicon Ultra 30,000 MWCO concentrators and centrifugal filter units (0.22 μm) GV Durapore® were purchased from Millipore Co. (Billerica, MA). Poly-Prep®

chromatography columns were purchased from Bio-Rad Laboratories (Hercules, CA). 1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl-methanethiosulfonate (MTSL) was purchased from Cayman Chemical Company (Ann Arbor, MI). Biosafe II scintillation fluid was purchased from Research Products International (Mount Prospect, IL). All other materials and reagents were purchased from Fisher Scientific (Hampton, NH).

3.2.2 Cloning

The pCHH₁₀m3N-STE14-TA plasmid was created by mutating the three endogenous cysteine (C99, C121, C126) into alanine. Thus, this plasmid was termed Ste14 “Triple A” (TA). Each His₁₀-myc₃-Ste14-TA mutant (Table 3.1) used in the following experiments were generated using site-directed mutagenesis via overlap extension polymerase chain reaction (PCR). Briefly, forward and reverse oligonucleotides were designed to insert a single point mutation into *STE14*. The previously created wild-type DNA plasmid template, pCHH₁₀m3N-STE14, was used for two-step PCR along with mutant primers and primers flanking *STE14* termed SVEC.² DNA was purified with Wizard® SV Gel and PCR Clean-up System and results were confirmed by a 1% agarose gel stained with ethidium bromide via electrophoresis. PCR products were combined and underwent a second round of PCR, creating the final, chimeric DNA products. Again, the DNA was purified with Wizard® SV Gel and PCR Clean-up System and results were confirmed by gel electrophoresis. The final PCR product and DNA template (pCHH₁₀m3N) were digested with restriction enzymes XmaI and SacII in order to have complimentary ends. Additionally, calf intestinal alkaline phosphatase (CIP) was added to the vector digest to catalyze the dephosphorylation of the 5' and 3' ends of the DNA to prevent the religation of linearized plasmid DNA. The expression vector contained the *URA3* gene and a phosphoglycerate kinase promoter (P_{PGK}) in order for uracil production as a selection marker and for protein overexpression, respectively.² Both digestion reactions were confirmed on an agarose gel, excised, and purified with Wizard® SV Gel and PCR Clean-up System. The products were ligated together using the LigaFast™ Rapid DNA Ligation System and transformed into DH5α chemically competent *E. coli* cells by following a standard heat shock transformation protocol (NEB). Cells were plated onto Luria-Bertani (LB) agar plates with ampicillin (100 µg/mL, LB-AMP) and grown at 37°C for 16-18 hours. Colonies that grew successfully were grown in tubes shaking at 220 RPM containing LB-AMP media at 37°C for 16-18 hours. Cells were pelleted and DNA was isolated with the

GeneJET Plasmid Miniprep Kit. The DNA sequence of each miniprep was confirmed with bidirectional dye-terminator sequencing (BigDye® Terminator v3.1) Results were confirmed using ApE – A plasmid Editor software.

3.2.3 Yeast Strains, Transformation, and Crude Membrane Preparations

STE14 plasmids (Table 3.2) were transformed into the $\Delta ste14$ yeast strain SM1188, as previously described.¹³ All *STE14* plasmids have the genotype: 2 μ URA3 P_{PGK}-His₁₀-myc₃N-STE14-mutant. SM1188 was grown on yeast peptone dextrose (YPD) agarose plates and then each yeast strain was cultured at 220rpm in YPD overnight at 30°C. The next day, the culture was diluted 1:3 with YPD media and placed back in the 30°C shaking incubator. After an additional 2 hours, 500 μ L of the culture was centrifuged and the supernatant was removed for every transformation reaction. The pellets were resuspended with single stranded salmon testes DNA (0.1mg). 1 μ g of each plasmid DNA of interest was added, along with autoclaved plate mix (40.5% polyethylene glycol, 0.1M lithium acetate, 10mM Tris-HCl, pH 7.5, 1mM ethylenediaminetetraacetic acid (EDTA), pH 8), and dithiothreitol (DTT). The resulting mixture incubated overnight at room temperature followed by a 10 min heat shock at 42°C. 100 μ L of each reaction was plated onto SC-URA plates and allowed to grow for 48 hours at 30°C.

Transformed yeast cells cultures were grown in SC-URA to mid-log phase (OD₆₀₀ 3-5), pelleted by centrifugation at 4000 xg and stored at -80°C. Preparation of crude membranes was completed as previously described with minor changes.² Briefly, the yeast cell pellets overexpressing His-Ste14 or mutants were resuspended in lysis buffer (300 mM sorbitol, 100mM NaCl, 5mM MgCl₂.6H₂O, 10mM Tris-HCl, pH 7.5, and 2mM AEBSF) and was placed on ice to swell for 15 mins. Pellets underwent rapid freeze-thaw cycles twice in liquid N₂ and a 25°C water bath, respectively. Thawed pellets were passed through a French press at 12,000 p.s.i, twice. The lysate was centrifuged at 500 xg for 10 mins at 4°C to remove any cell debris. The supernatant was collected, and the remaining membrane fraction was isolated through ultracentrifugation at 100,000 xg for 1 hour at 4°C. After centrifugation, the supernatant was removed and the pellet was resuspended in 10 mM Tris-HCl, pH 7.5, aliquoted, and flash frozen in liquid N₂. All aliquots were stored at -80°C until further experimentation. Total protein concentration was determined using the Bradford protein assay.¹⁴

3.2.4 *In vitro* Methyltransferase Vapor Diffusion Assay of Crude Membranes

An *in vitro* methyltransferase vapor diffusion assay was performed as previously described.^{2,15,16} Briefly, reactions were prepared in a 100mM Tris-HCl, pH 7.5 buffer, containing crude membranes overexpressing His-Ste14-TA (CH2803) or His-Ste14-TA mutants (5μg), AFC (200μM), and [¹⁴C]-SAM (20μM). The negative control reaction contained a yeast deletion strain of *ste14Δ* (CH2714, 5μg), and all other components just described. Samples were incubated for 30 mins at 30°C and then were stopped with 50 μL of 1 M NaOH/1% SDS (v/v). Each reaction (100μL) was spotted onto a filter and placed in the neck of a scintillation vial filled with Biosafe II scintillation fluid (10mL) and capped at room temperature. After three hours, the filters were removed, and the vials were placed in a Packard Tri-Carb 1600CA liquid scintillation counter in order to detect the [¹⁴C]-methanol that diffused from filters into the scintillation fluid. All specific activities were characterized as pmol of methyl groups from [¹⁴C]-SAM transferred to the AFC substrate per minute of assay per mg of His-Ste14-TA (CH2803) or His-Ste14 mutant protein. These activities were calculated after subtracting background counts from the average of the negative control samples. Each experiment had three reactions per sample and all experiments were performed in triplicate.

3.2.5 Purification of His-Ste14 and His-Ste14 Mutants

Purification of His-Ste14-TA or His-Ste14-TA single cysteine mutants was performed as previously described.² Briefly, the crude membranes overexpressing the protein of interest were solubilized in a solution of 0.02M imidazole, 1% (w/v) *n*-dodecyl-β-D-maltopyranoside (DDM), and 1xEQ buffer (50mM Na₂HPO₄, 0.3M NaCl, pH 6.0) with the addition of 10% glycerol and 2mM AEBSF. For the solubilization reaction, crude membranes were at a final concentration of 5 mg/mL. The reaction was rocked at 4°C for 1 hour. The resulting mixture was centrifuged at 100,000 xg for 45 mins. The soluble fraction was incubated for 1 hour at 4°C with Talon® cobalt metal affinity beads equilibrated in 1xEQ buffer (pH 6.0) with the addition of 10% glycerol and 2mM AEBSF. After the protein was incorporated with the resin, the solution was centrifugated at 350 xg for 2 mins at 4°C and the remaining buffer was removed via aspiration. Next, the beads underwent four wash steps: buffer was added to the resin and was allowed to rock at 4°C for 10 mins followed by centrifugation at 350 x g for 2 min before aspiration of excess buffer. The resin was washed twice with buffer A (1xEQ buffer, pH 6.0, with the addition of 1% DDM (w/v) and

40mM imidazole), once with buffer B (1xEQ buffer, pH 6.0, with the addition of 1% DDM (w/v), 500mM KCl, and 40mM imidazole), and once with buffer C (1xEQ buffer, pH 6.0, with the addition of 0.1% DDM (w/v), 500mM KCl, and 40mM imidazole). After the final wash, buffer C was not aspirated off and loaded onto a Poly-Prep® chromatography column. His-Ste14-TA or His-Ste14-TA mutants were eluted from the cobalt resin with elution buffer, pH 6.0 (0.1% DDM, 1M imidazole, and 2mM AEBSF). Collected protein was concentrated at 5000 xg in a 30,000 MWCO concentrator for ~40 mins at 4°C. The concentrated protein was aliquoted, flash frozen, and stored at -80°C until further experimentation. Pure protein concentration was determined using an Amido Black protein assay.¹⁷

3.2.6 Purification of His-Ste14 and His-Ste14 Mutants with the Addition of MTSL Spin Label

Purification of His-Ste14-TA single cysteine mutants were performed as previously described.² Briefly, the crude membranes previously isolated were solubilized in a solution of 0.02M imidazole, 1% (w/v) *n*-dodecyl- β -D-maltopyranoside (DDM), and 1xEQ buffer (50mM Na₂HPO₄, 0.3M NaCl, pH 6.0) with the addition of 10% glycerol and 2mM AEBSF. For the solubilization reaction, crude membranes were at a final concentration of 5 mg/mL. The reaction was rocked at 4°C for 1 hour. The resulting mixture was centrifuged at 100,000 xg for 45 mins. The soluble fraction was incubated for 1 hour at 4°C with Talon® cobalt metal affinity beads equilibrated in 1xEQ buffer (pH 6.0) with the addition of 10% glycerol and 2mM AEBSF. After the protein was incorporated with the resin, the solution was centrifuged at 350 xg for 2 mins at 4°C and the remaining buffer was removed via aspiration. Next, the beads underwent four wash steps: buffer was added to the resin and was allowed to rock at 4°C for 10 mins followed by centrifugation at 350 x g for 2 min before aspiration of excess buffer. The resin was washed twice with buffer A (1xEQ buffer, pH 6.0, with the addition of 1% DDM (w/v) and 40mM imidazole), once with buffer B (1xEQ buffer, pH 6.0, with the addition of 1% DDM (w/v), 500mM KCl, and 40mM imidazole), and once with buffer C (1xEQ buffer, pH 6.0, with the addition of 0.1% DDM (w/v), 500mM KCl, and 40mM imidazole). After the final wash, dithiothreitol (DTT) was added at a concentration of three times the estimated protein concentration for 30 mins at 4°C. DTT is necessary to reduce the single cysteine residue in order to successfully attach the (1-Oxyl-2,2,5,5-tetramethyl Δ^3 -pyrroline-3-methyl) Methanesulfonylthioate (MTSL) spin label. MTSL was added

at a ten-fold higher concentration than the estimated protein concentration and incubated for 4 hours at 4°C. Then, the solution underwent centrifugation at 350 x g for 2 min and loaded onto a Poly-Prep® chromatography column. The resin was washed with 50 bed volumes of 1xEQ buffer (pH 6.0) with the addition of 0.1% DDM and 40mM imidazole. Then, the His-Ste14-TA mutant was eluted from the cobalt resin with elution buffer, pH 6.0 (0.1% DDM, 1M imidazole, and 2mM AEBSF). Collected protein was concentrated at 5000 xg in a 30,000 MWCO concentrator at 4°C until the volume was under 700µL. A buffer exchange was performed by adding 1mL of new 1x EQ buffer, pH 6.0 (20mM imidazole, 0.01% DDM, and 2mM AEBSF) followed by centrifugation at 5000 xg in the same 30,000 MWCO concentrator at 4°C until the volume was under 1mL. Overall, the buffer exchange reduced the concentration of imidazole in the buffer containing MTSL-incorporated protein. Concentrated protein was filtered in centrifugal filter units (0.22 µm) GV Durapore® by centrifuging at 12,000 xg for 2 mins at 4°C. Once filtered, the His-Ste14-TA mutants were applied to a Superdex S200 column pre-equilibrated with filtered 1xEQ buffer, pH 6.0 (with the addition of 0.05% DDM and 2mM DTT) at 4°C and spectra was collected at 280nm. Fractions containing purified protein were pooled, concentrated in a new 30,000 MWCO concentrator, and aliquoted. Samples were flash frozen in liquid nitrogen and stored at -80°C until further experimentation. Pure protein concentration was determined using an Amido Black protein assay.¹⁷

3.2.7 SDS-PAGE Staining and Immunoblot Analysis

Crude membrane extracts overexpressing His-Ste14-TA (CH2803) (1µg), His-Ste14-TA mutants (1µg), or purified His-Ste14-TA mutants (0.1µg for immunoblot and 1.0µg for Coomassie stain) were eluted on a 10% SDS-PAGE. Gels were either incubated with Coomassie stain (0.25% (w/v) Coomassie Brilliant Blue R-250, 80% methanol and 20% acetic acid) or transferred to a nitrocellulose membrane (0.22 µm). For immunoblotting, the nitrocellulose membrane was blocked in 20% (w/v) non-fat dry milk in phosphate-buffered saline (2.70 mM KCl, 137 mM NaCl, 4.00 mM Na₂HPO₄, 1.80 mM KH₂PO₄, pH 7.4) with 0.05% (v/v) Tween-20 (PBST) at room temperature for 3 hours. After three washes with PBST, the membranes were incubated with α-myc (1:10,000) antibody in 5% (w/v) non-fat dry milk in PBST at 4°C overnight. After an additional three washes with PBST, the membranes were incubated with goat α-mouse IgG-HRP (1:2,500) antibody in 5% (w/v) non-fat dry milk in PBST at room temperature for 1 hour. The

membranes were washed three more times with PBST and visualized with enhanced chemiluminescence (ECL) on a GeneGnome XRQ. Gels stained with Coomassie (0.25% (w/v) Coomassie Brilliant Blue R-250, 80% methanol and 20% acetic acid) were incubated overnight. The next day, the gels were incubated with Destaining Solution (30% methanol, 10% acetic acid) followed by rinses with ddH₂O. Once bands were visualized, the gels were incubated in Casting Solution (5% glycerol, 20% ethanol) and preserved in cellophane.

3.2.8 *In vitro* Methyltransferase Vapor Diffusion Assay of Pure His-Ste14 and Mutants

His-Ste14-TA single cysteine mutants purified with and without MTSL were tested for methyltransferase activity levels with the same protocol. Briefly, the purified protein was diluted 1:5 in 1xEQ buffer with 0.05% DDM. Purified protein was combined with AFC (100μM) and [¹⁴C]-SAM (20μM) in 1xEQ buffer (pH 6.0) with 0.05% DDM. The negative control reaction contained ddH₂O instead of purified protein. Samples were incubated for 30 mins at 30°C and then were stopped with 50 μL of 1 M NaOH/1% SDS (v/v). Each reaction (100μL) was spotted onto a filter and placed in the neck of a scintillation vial filled with Biosafe II scintillation fluid (10mL) and capped at room temperature. After three hours, the filters are removed, and the vials were placed in a Packard Tri-Carb 1600CA liquid scintillation counter in order to detect the [¹⁴C]-methanol that diffused from filters into the scintillation fluid. All specific activities were characterized as pmol of methyl groups from [¹⁴C]-SAM transferred to the AFC substrate per minute of assay per mg of purified His-Ste14-TA mutant protein. These activities were calculated after subtracting background counts from the average of the negative control samples. Each experiment had three reactions per sample and all experiments were performed in triplicate.

3.2.9 EPR Spectroscopy Conditions

Purified His-Ste14-TA single cysteine mutations (100μM) were inserted into glass capillary tubes (30μL). The plugged capillary tube was placed into a Wilmad thin wall quartz EPR sample tube and placed into the ER4119HS cavity of the Bruker EMX EPR spectrometer. Continuous-wave (CW) spectra were obtained at 20mW power, 100 kHz modulation frequency, 5 G field modulation amplitude, 100 G sweep width, 100 receiver gain and ten scans at room temperature using the WinEPR Acquisition software. The same procedure was repeated with the addition of 5μL SAM (5mM) to 25μL purified His-Ste14-TA single cysteine mutation (100μM).

3.2.10 TEMPO Standard Curve

2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO) was diluted to 15 μ M, 50 μ M, 100 μ M, and 150 μ M in the 1xEQ buffer (with the addition of 0.05% DDM) and were inserted into glass capillary tubes (30 μ L). The plugged capillary tube was placed into a Wilmad thin wall quartz EPR sample tube and placed into the ER4119HS cavity of the Bruker EMX EPR spectrometer. Continuous-wave (CW) spectra were obtained at 20mW power, 100 kHz modulation frequency, 5 G field modulation amplitude, 100 G sweep width, 100 receiver gain and ten scans at room temperature using the WinEPR Acquisition software. A standard curve was generated by calculating the double integral (DI) of each spectrum versus the concentration of the TEMPO and was used to determine the concentration of MTSL in each sample (Figure 3.8). The DI was determined with the WinEPR Acquisition software. The concentration of protein in each sample (determined by the Amido Black assay) was divided by the concentration of MTSL of that sample from the standard curve to determine the efficiency of protein labeling with spin label probe.

3.3 Results and Discussion

Our EPR spectroscopy experiments required active protein with a single cysteine residue in the cytosolic portion of the protein. Therefore, a library of all 46 non-conserved residues of His-Ste14-TA were mutated to cysteine using primers in Table 3.1 and transformed into *S. cerevisiae* (Table 3.2). Only non-conserved residues had been mutated since conserved residues are essential for enzymatic activity.^{18,19} Crude membranes were extracted from yeast, assessed for protein expression levels, and methyltransferase activity assays were performed in triplicate (Figure 3.1, Table 3.3). All mutants had relatively similar expression levels compared to His-Ste14-TA except His-K154C-TA and His-I210C-TA. Both mutants had activity levels less than 10% compared to His-Ste14-TA, indicating that these mutants were detrimental for protein expression and folding. His-Ste14-TA mutant crude membranes that retained activity levels of $\geq 50\%$ compared to His-Ste14-TA crude membranes were solubilized in DDM and purified (Figure 3.2, Table 3.3). Cysteine mutants that had a marked decrease in enzymatic activity were not pursued further. Nineteen mutants in crude membranes maintained high enough activity levels to proceed to the purification step: His-Ste14-Y134C-TA, His-Ste14-T135C-TA, His-Ste14-I138C-TA, His-Ste14-H141C-TA, His-Ste14-H145C-TA, His-Ste14-V160C-TA, His-Ste14-T164C-TA, His-Ste14-S170C-TA, His-Ste14-S207C-TA, His-Ste14-R211C-TA, His-Ste14-V212C-TA, His-Ste14-

K215C-TA, His-Ste14-A223C-TA, His-Ste14-E227C-TA, His-Ste14-N230C-TA, His-Ste14-K231C-TA, His-Ste14-G233C-TA, Ste14-V234C-TA, and Ste14-I238C-TA. Interestingly, His-H141C-TA and His-S222C-TA both had percent activities greater than 100% of His-Ste14-TA.

All nineteen His-Ste14-TA mutants that retained activity were purified and evaluated for purity and protein expression (Figure 3.3, Figure 3.4). Nine purified His-Ste14-TA mutants retained activity levels of $\geq 50\%$ compared to purified His-Ste14-TA. Thus, our primary candidates for EPR spectroscopy: His-Ste14-T135C-TA, His-Ste14-I138C-TA, His-Ste14-H141C-TA, His-Ste14-T164C-TA, His-Ste14-V212C-TA, His-Ste14-K215C-TA, His-Ste14-A223C-TA, His-Ste14-G233C-TA, and His-Ste14-V234C-TA (Figure 3.5, Table 3.4). Heat maps of His-Ste14-TA cysteine mutants in crude membranes and purified were used to select the first residues to label with MTSL (Figure 3.6).

His-Ste14-V212C-TA and His-Ste14-K231C-TA were the first mutants selected to be purified with MTSL. We chose mutants in the C-terminal tail of His-Ste14-TA because Loop 4 is predicted to act as a hinge and we wanted to understand the lesser studied structural changes in the C-terminus.³ Both mutants were pure and active after the purification process (data not shown). EPR spectra for each mutant was collected in the presence or absence of saturation concentrations of SAM (5 mM) (Figure 3.7, Figure 3.8). When SAM was added to the His-Ste14-V212C-TA mutant, small changes in EPR spectra could be detected, especially in the wings of the spectra. The mobility of the probe at V212C may be lowered slightly during SAM binding, causing a small broadening of the peaks. EPR spectroscopy has high sensitivity to structural changes in protein, thus, this spectrum could be indicative of the movement of one alpha helix. In contrast, when SAM was added to the His-K231C-TA sample, the EPR spectral line became noticeably broader compared to the sample without the cofactor. Broadening of EPR spectra is indicative of the MTSL probe becoming buried and tumbling rate of the lone electron slowing. We predict that when SAM is not bound, K231 is in the cytosol, not acting with the active site. Upon addition of SAM, the enzyme must undergo a conformational change where K231 is now in closer proximity to itself or SAM. More EPR spectroscopy data must be collected in order to construct an informed model of SAM binding to Ste14.

The efficiency of labeling was determined by comparing the concentration of label bound to the protein (calculated by taking the double integral of the spectrum) to a standard curve of TEMPO (Figure 3.9). The last point of the double integral is equal to the area of the EPR absorption

and has a linear dependence with respect to the modulation amplitude and concentration of sample.²⁰ TEMPO is identical to MTSL except it does not contain the thiosulfonate ester functional groups necessary to form a covalent bond with a cysteine residue. At concentrations of TEMPO ranging between 15 – 150 μ M, EPR spectra were taken and the double integral of each spectra were taken. A linear regression line was determined when the concentration of the TEMPO was plotted against the double integral. The same value (the double integral) was determined for both His-V212C-TA and His-K231C-TA spectra and the concentration of spin label was calculated, ~75 μ M and ~55 μ M, respectively. Then the actual concentration of the protein in each sample was divided by the amount of spin-label, therefore determining the labeling efficiency of the MTSL probe to the introduced cysteine residue. His-V212C-TA had an average of 40.5% labeling efficiency, while His-K231C-TA had an average of 67.5% labeling efficiency. Although the labeling to -His-V212C-TA was quiet low, this could be due the its location in the C-terminal loop. The mutation at V212 could be too close to the membrane and partially obstructed by the detergent micelle surrounding the protein, making it difficult for the MTSL spin-label to bind. On the other hand, the K231C mutation in the C-terminus of Ste14 is not at the end of any transmembrane domains and was most likely not being affected by detergent molecules. Thus, it is not surprising that His-K231C-TA has a higher average binding efficiency than His-V212C-TA.

3.4 Future Directions

The next step for this project is to attach MTSL to other Ste14-TA single cysteine mutants that retained activity similar to His-Ste14-TA and evaluate its EPR spectra with and without SAM. Now that the purification process and MTSL incorporation process has been proven successful, further data collection is required. Table 3.4 summarizes which purified His-Ste14-TA single cysteine mutants retained methyltransferase activity and would be optimal to evaluate next via EPR spectroscopy. Based on the binding efficiency of His-V212C-TA and His-K231C-TA, it would be advantageous to select a mutation further from the membrane spanning regions of the enzyme to ensure higher binding efficiency with MTSL and a great chance of environmental changes.

Additionally, a library of double cysteine mutations in Loop 4 and the C-terminus could be created for double electron-electron resonance (DEER) spectroscopy. In combination with site-directed mutagenesis, DEER is a powerful biophysical technique that can obtain distance

measurements up to 80Å between the two cysteine residues by measuring the dipolar coupling between two unpaired electron spins.^{21,22} This is one of the only biophysical methods that can detect a large conformation change in an active protein, as well as, give quantifiable information about the distance between specific regions. Based on Yang *et al.*, residues S146 and V160 in His-Ste14 are predicted to be only 6.4Å apart, however conformational changes are predicted to occur during SAM binding.⁴ As discussed earlier, it is hypothesized that Loop 4 unfolds or acts as a hinge to allow SAM access to the binding site. We could examine the conformational dynamics by making distance measurements of cysteine mutants located in Loop 4 and/or the C-terminal loop upon SAM binding to better understand the importance of this region for function and its role in the dynamics of the catalytic process.

Table 3.1 Oligonucleotides used in this study to generate *STE14* mutants

| Name | Oligonucleotide (5' – 3') |
|-----------|--------------------------------|
| SVEC FWD* | CAAGGGGGTGGTTTAGTTAG |
| SVEC REV* | CAACTGTTGGGAAAGGCGATC |
| Y134C FWD | CTGGGACAATGTACCAGAACTATTGCTATG |
| T135C FWD | GGACAATATTGCAGAACTATTG |
| T137C FWD | CAATATAACCAGATGTATTGCTATG |
| I138C FWD | CAATATAACCAGAACTTGTGCTATGCATAC |
| A139C FWD | CCAGAACTATTTGTATGCATACTGCGGGA |
| H141C FWD | CTATTGCTATGTGTACTGCGGG |
| T142C FWD | GCTATGCATTGTGCGGGACATTC |
| G144C FWD | CATACTGCGTGTTCATTCTTTTC |
| H145C FWD | CATACTGCGGGATGTTCTTTTCTC |
| S146C FWD | GCGGGACATTGCTTTTCTCATATTG |
| S148C FWD | CATTCCTTTTGTTCATATTGTG |
| I150C FWD | CTTTTCTCATTGTGTGAAGACC |
| K152C FWD | CATATTGTGTGTACCAAGAAG |
| T153C FWD | CATATTGTGAAGTGCAAGAAGGAG |
| K154C FWD | GTGAAGACCTGTAAGGAGTCC |
| E156C FWD | GACCAAGAAGTGTTCCGATCATG |
| S157C FWD | CCAAGAAGGAGTGCGATCATGTTTTAGTG |
| D158C FWD | GAAGGAGTCCTGTCATGTTTTAG |
| V160C FWD | CCGATCATTGTTTAGTG |
| K163C FWD | GTTTTAGTGTGTACTGGC |
| T164C FWD | GTGAAATGTGGCGTTTAC |
| S168C FWD | GGCGTTTACTGCTGGTCAAG |
| W169C FWD | GTTTACTCCTGCTCAAGACATCC |
| S170C FWD | GTTTACTCCTGGTGTAGACATCCA |
| S174C FWD | GACATCCATGTTATCTTG |
| S207C FWD | GGAAATTTTTCTGTGATAGAATTCG |
| D208C FWD | GGAAATTTTTTCAGTTGTAGAATTCGTGTG |
| I210C FWD | CAGTGATAGATGTCGTGTGGAAG |
| R211C FWD | GATAGAATTTGTGTGGAAGAAAAATATTTG |
| V212C FWD | GATAGAATTCGTTGCGAAGAAAAATATTTG |
| K215C FWD | GTGGAAGAATGTTATTTGATAG |
| Y216C FWD | GAAGAAAAATGTTTGATAGAG |
| E219C FWD | GAAAAATATTTGATATGCTTCTTTAGCGCG |
| S222C FWD | GAGTTCCTTTGCGCGGAATACATA |
| A223C FWD | GTTCTTTAGCTGCGAATA |
| E224C FWD | CTTTAGCGCGTGCTACATAGAATAC |

Table 3.1 continued

| | |
|-----------|---------------------------------|
| I226C FWD | GCGCGGAATACTGCGAATACAAG |
| E227C FWD | GCGGAATACATATGTTACAAG |
| N230C FWD | CCAACCTTGCACTTGTATTC |
| K231C FWD | TACAAGAATTGCGTTGGTG |
| V232C FWD | CAAGAATAAGTGTGGTGTCGG |
| K231C FWD | TACAAGAATTGCGTTGGT |
| G233C FWD | GAATAAGGTTTGTGTCGGAATA |
| V234C FWD | GAATAAGGTTGGTTGCGGAATACC |
| I236C FWD | GTTGGTGTCGGATGCCCTTTTATATAACCGC |
| F238C FWD | GTCGGAATACCTTGTATATAACCGCGGTGG |
| I239C FWD | GGAATACCTTTTGTGCTAACCGCGGTGG |

*SVEC primers align within the pCHH10m3N vector. SVEC FWD aligned with the forward 200 base pairs upstream of STE14 and SVEC REV aligned with the 182 base pairs downstream of STE14.

Table 3.2 *S. cerevisiae* plasmids used in this study

| Name | Genotype |
|---------------------|---|
| SM1058 ¹ | <i>Mata trp1 leu2 ura3 his4 can1</i> |
| SM1188 ¹ | <i>Ste14-4::TRP1</i> , isogenic to SM1058 |
| CH2704 ² | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14 |
| CH2714 | 2 μ URA3 P _{PGK} |
| CH2803 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA |
| CH3152 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-Y134C |
| CH3150 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-T135C |
| CH3151 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-T137C |
| CH3153 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-I138C |
| CH3146 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-A139C |
| CH3108 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-H141C |
| CH3109 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-T142C |
| CH3103 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-G144C |
| CH3104 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-H145C |
| CH3105 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-S146C |
| CH3106 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-S148C |
| CH3092 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-I150C |
| CH3163 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-K152C |
| CH3093 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-T153C |
| CH3149 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-K154C |
| CH3164 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-E156C |
| CH2804 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-S157C |
| CH3165 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-D158C |
| CH3065 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-V160C |
| CH3166 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-K163C |
| CH3066 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-T164C |
| CH3110 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-S168C |
| CH3111 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-W169C |
| CH3167 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-S174C |
| CH3107 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-S207C |
| CH3147 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-D208C |
| CH3270 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-I210C |
| CH3145 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-R211C |
| CH3154 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-V212C |
| CH3168 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-K215C |
| CH3271 | 2 μ URA3 P _{PGK} -His ₁₀ -myc ₃ N-STE14-TA-Y216C |

Table 3.2 continued

| | |
|--------|---|
| CH3144 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-E219C |
| CH3272 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-S222C |
| CH3067 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-A223C |
| CH3273 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-E224C |
| CH3155 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-I226C |
| CH3274 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-E227C |
| CH3077 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-N230C |
| CH3094 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-K231C |
| CH3275 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-V232C |
| CH3068 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-G233C |
| CH3276 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-V234C |
| CH3277 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-I236C |
| CH3278 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-F238C |
| CH3279 | 2 μ URA3 P _{PGK} -His ₁₀ - <i>myc</i> ₃ N-STE14-TA-I239C |

All strains are isogenic to SM1188 and has the following characteristics located at the N-terminus of the gene: PGK-phosphoglycerate kinase promoter (P_{PGK}), ten histidine residues (His₁₀-), and three *myc* tags with the sequence 5' EQKLISEEDL 3' (*myc*₃-).

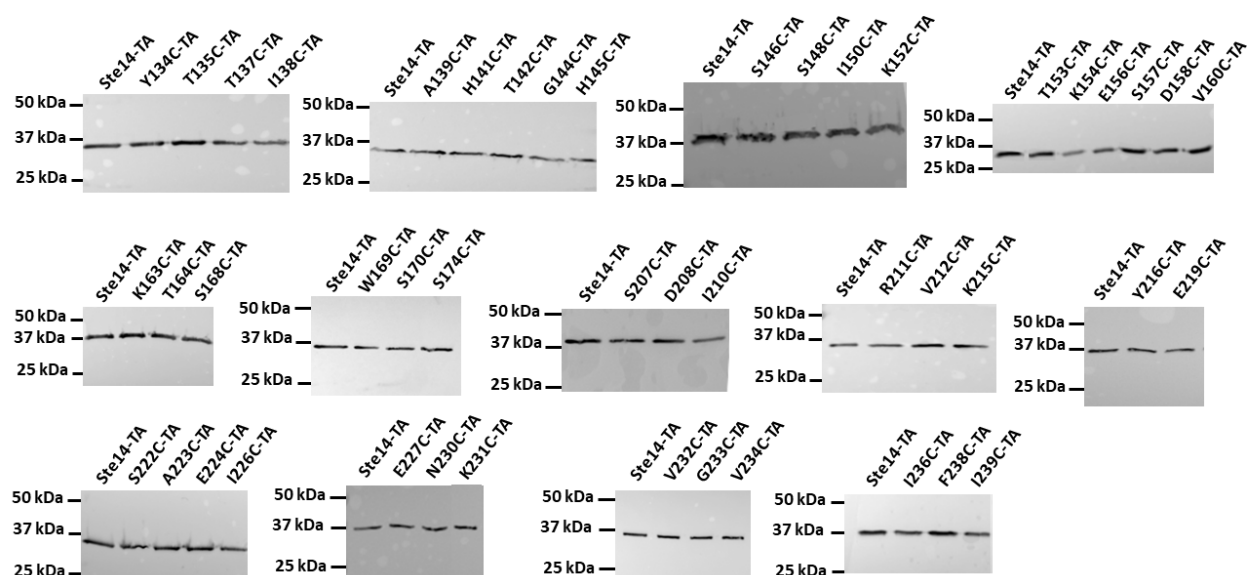


Figure 3.1 Expression levels of His-Ste14-TA and single cysteine mutants. Indicated by immunoblotting, expression levels from all mutants were similar to that of His-Ste14-TA. Protein samples (1 μ g) were run on a 10% SDS-PAGE followed by immunoblot analysis with α -myc (1:10,000) and goat α -mouse IgG-HRP (1:2,500). Immunoblots were visualized with ECL. His-Ste14-TA from crude membranes was the positive control located in the first lane of each gel.

Table 3.3 Specific methyltransferase activities of His-Ste14-TA mutants from crude membranes tested in this study

| His-Ste14 mutant | Activity (% of His-Ste14-TA) |
|--------------------|------------------------------|
| Triple A (TA) | 100 |
| His-Ste14-Y134C-TA | 59.6 \pm 8.3 |
| His-Ste14-T135C-TA | 91.7 \pm 14.9 |
| His-Ste14-T137C-TA | 18.0 \pm 6.2 |
| His-Ste14-I138C-TA | 77.6 \pm 7.8 |
| His-Ste14-A139C-TA | 2.9 \pm 3.0 |
| His-Ste14-H141C-TA | 125.2 \pm 17.0 |
| His-Ste14-T142C-TA | 8.4 \pm 6.0 |
| His-Ste14-G144C-TA | 5.4 \pm 4.3 |
| His-Ste14-H145C-TA | 44.2 \pm 1.4 |
| His-Ste14-S146C-TA | 38.8 \pm 6.2 |
| His-Ste14-S148C-TA | 3.6 \pm 3.6 |
| His-Ste14-I150C-TA | 14.3 \pm 1.8 |
| His-Ste14-K152C-TA | 14.7 \pm 1.9 |
| His-Ste14-T153C-TA | 39.6 \pm 8.6 |
| His-Ste14-K154C-TA | 5.9 \pm 2.1 |
| His-Ste14-E156C-TA | 3.3 \pm 3.6 |
| His-Ste14-S157C-TA | 29.6 \pm 6.0 |
| His-Ste14-D158C-TA | 24.7 \pm 1.1 |
| His-Ste14-V160C-TA | 58.6 \pm 10.1 |
| His-Ste14-K163C-TA | 28.7 \pm 2.6 |
| His-Ste14-T164C-TA | 76.3 \pm 7.9 |
| His-Ste14-S168C-TA | 15.4 \pm 4.5 |
| His-Ste14-W169C-TA | 12.4 \pm 7.1 |
| His-Ste14-S170C-TA | 42.6 \pm 5.4 |
| His-Ste14-S174C-TA | 10.4 \pm 0.9 |
| His-Ste14-S207C-TA | 63.9 \pm 7.9 |
| His-Ste14-D208C-TA | 11.0 \pm 7.7 |
| His-Ste14-I210C-TA | 0.0 \pm 0.0 |
| His-Ste14-R211C-TA | 49.7 \pm 5.9 |
| His-Ste14-V212C-TA | 51.9 \pm 6.5 |
| His-Ste14-K215C-TA | 80.0 \pm 1.0 |
| His-Ste14-Y216C-TA | 17.5 \pm 3.1 |
| His-Ste14-E219C-TA | 27.3 \pm 5.4 |
| His-Ste14-S222C-TA | 29.2 \pm 2.2 |
| His-Ste14-A223C-TA | 89.3 \pm 15.9 |
| His-Ste14-E224C-TA | 0.9 \pm 0.1 |
| His-Ste14-I226C-TA | 23.9 \pm 6.6 |

Table 3.3 continued

| | |
|--------------------|-----------------|
| His-Ste14-E227C-TA | 112.7 ± 5.2 |
| His-Ste14-N230C-TA | 70.9 ± 6.2 |
| His-Ste14-K231C-TA | 46.7 ± 11.6 |
| His-Ste14-V232C-TA | 13.1 ± 6.0 |
| His-Ste14-G233C-TA | 79.1 ± 8.4 |
| His-Ste14-V234C-TA | 62.1 ± 10.9 |
| His-Ste14-I236C-TA | 11.0 ± 8.1 |
| His-Ste14-F238C-TA | 7.5 ± 1.8 |
| His-Ste14-I239C-TA | 40.8 ± 2.8 |

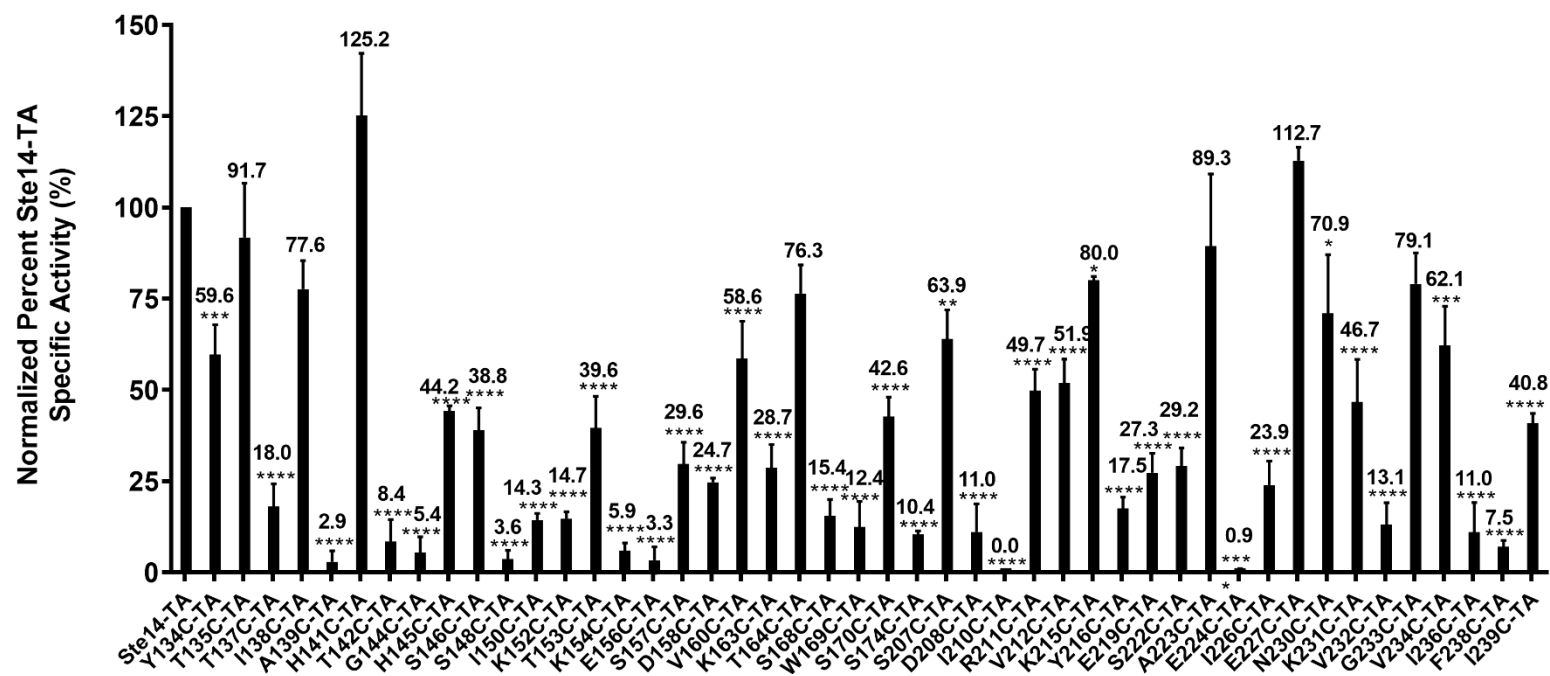


Figure 3.2 In vitro vapor diffusion methyltransferase activity assay of 46 non-conserved cytosolic His-Ste14-TA single cysteine mutants in crude membranes.

All specific activities were characterized as pmol of methyl groups from [14 C]-SAM transferred to the AFC substrate per minute of assay per mg of His-Ste14-TA (CH2803) or His-Ste14 mutant protein. These activities were calculated after subtracting the background counts from the average of the negative control samples. Percent activity compared to His-Ste14-TA was calculated for each mutant. Each experiment had three reactions per sample and all experiments were performed in triplicate \pm S.D. (****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.1$; based on one-way ANOVA followed by Dunnett's multiple comparisons test vs. His-Ste14).

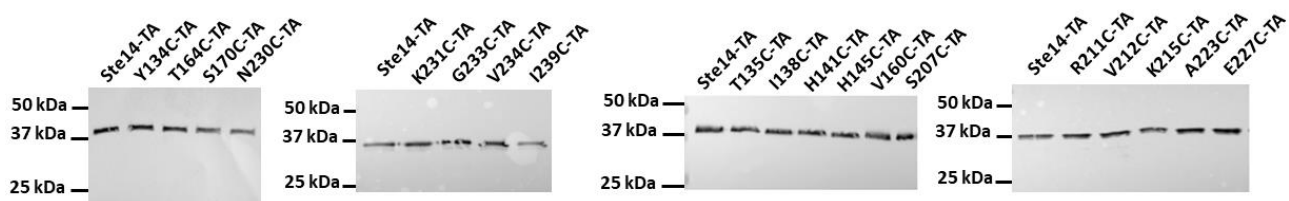


Figure 3.3 Expression levels and purity of purified His-Ste14-TA and single cysteine mutants. Indicated by immunoblotting, expression levels from all mutants were similar to that of purified His-Ste14-TA. Protein samples (0.1 μ g) were run on a 10% SDS-PAGE followed by immunoblot analysis with α -myc (1:10,000) and goat α -mouse IgG-HRP (1:2,500). Immunoblots were visualized with ECL. Purified His-Ste14-TA was the positive control located in the first lane of each gel.

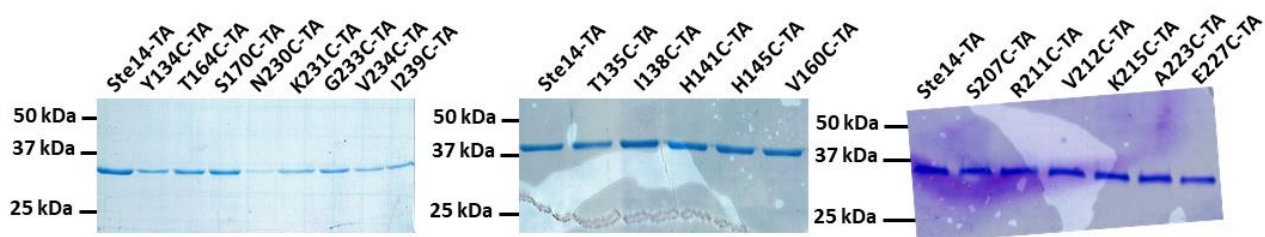


Figure 3.4 Purity of His-Ste14-TA and single cysteine mutants. Pure protein samples (2 μ g) were run on a 10% SDS-PAGE and visualized with Coomassie (0.25% (w/v) Coomassie Brilliant Blue R-250, 80% methanol and 20% acetic acid). Purified His-Ste14-TA was the positive control located in the first lane of each gel.

Table 3.4 Specific methyltransferase activities of purified His-Ste14-TA mutants tested in this study

| His-Ste14 mutant | Activity (% of His-Ste14-TA) |
|--------------------|------------------------------|
| Triple A (TA) | 100 |
| His-Ste14-Y134C-TA | 2.3 \pm 0.4 |
| His-Ste14-T135C-TA | 60.4 \pm 1.5 |
| His-Ste14-I138C-TA | 59.7 \pm 3.6 |
| His-Ste14-H141C-TA | 217.6 \pm 0.8 |
| His-Ste14-H145C-TA | 2.9 \pm 0.2 |
| His-Ste14-V160C-TA | 11.4 \pm 0.2 |
| His-Ste14-T164C-TA | 90.0 \pm 14.1 |
| His-Ste14-S170C-TA | 28.0 \pm 2.5 |
| His-Ste14-S207C-TA | 12.3 \pm 0.2 |
| His-Ste14-R211C-TA | 41.4 \pm 2.5 |
| His-Ste14-V212C-TA | 117.4 \pm 2.5 |
| His-Ste14-K215C-TA | 88.6 \pm 4.1 |
| His-Ste14-A223C-TA | 104.3 \pm 1.8 |
| His-Ste14-E227C-TA | 48.8 \pm 1.6 |
| His-Ste14-N230C-TA | 26.7 \pm 1.6 |
| His-Ste14-K231C-TA | 62.4 \pm 1.8 |
| His-Ste14-G233C-TA | 3.0 \pm 0.3 |
| His-Ste14-V234C-TA | 1.3 \pm 0.4 |
| His-Ste14-I239C-TA | 7.7 \pm 0.6 |

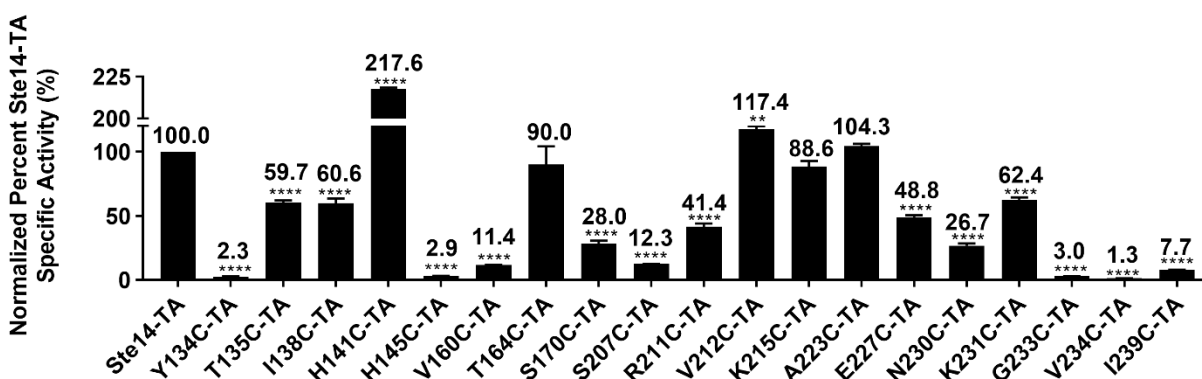
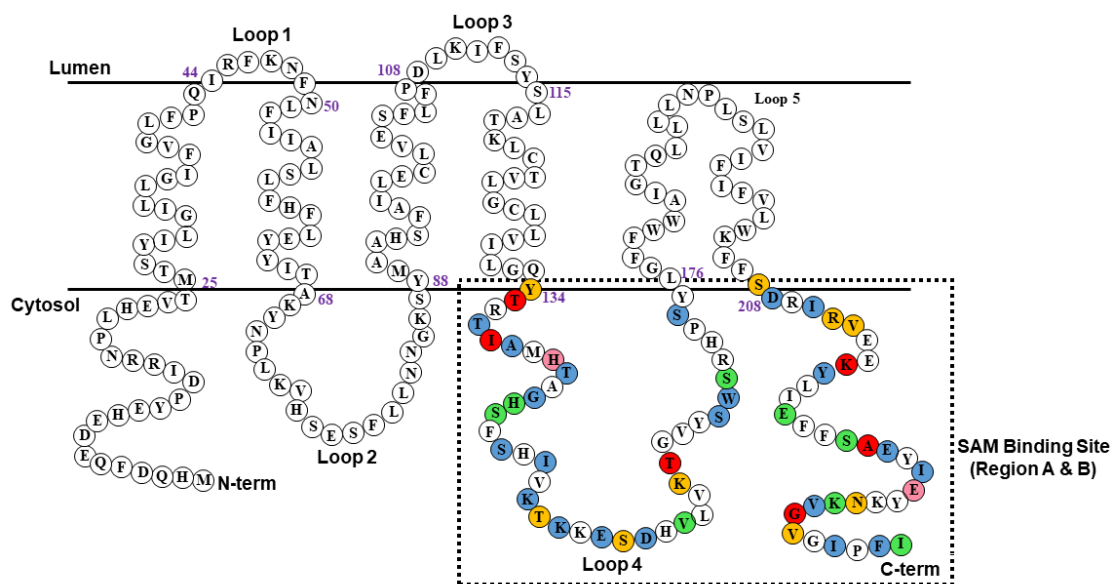


Figure 3.5 *In vitro* vapor diffusion methyltransferase activity assay of nineteen purified non-conserved cytosolic His-Ste14-TA single cysteine mutants.

All specific activities were characterized as pmol of methyl groups from [^{14}C]-SAM transferred to the AFC substrate per minute of assay per mg of purified His-Ste14-TA (CH2803) or purified His-Ste14 mutant protein. These activities were calculated after subtracting the background counts from the average of the negative control samples. Percent activity compared to purified His-Ste14-TA was calculated for each mutant. Each experiment had three reactions per sample and all experiments were performed in triplicate \pm S.D. (****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; based on one-way ANOVA followed by Dunnett's multiple comparisons test vs. His-Ste14).

A



B

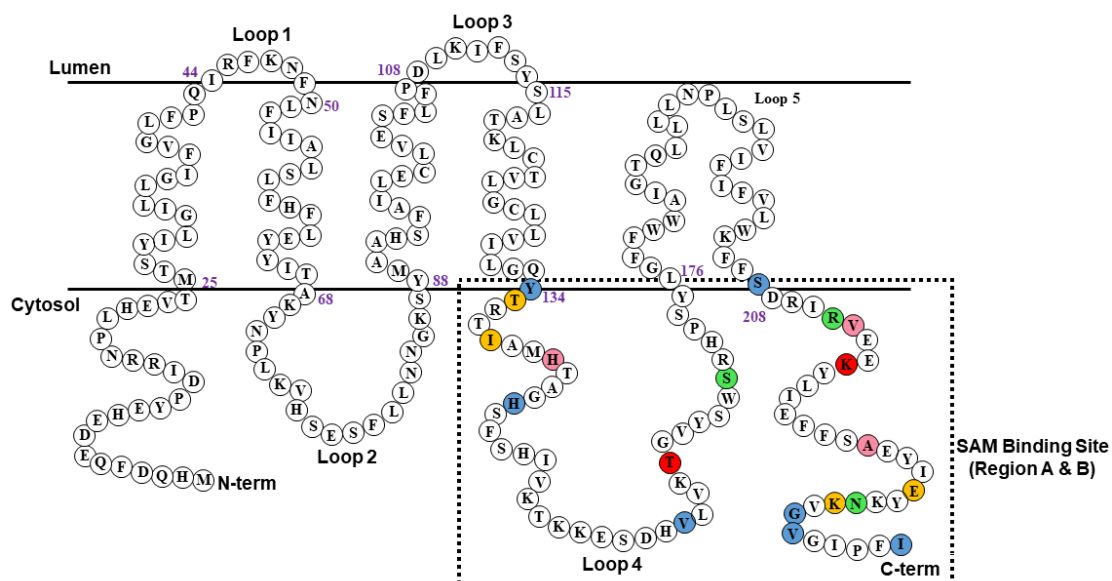


Figure 3.6 Heat map of single cysteine mutants in crude membranes (A) and purified in DDM (B) of all non-conserved residues of His-Ste14-TA in Loop 4 and the C-terminus. Mutants that retained $\geq 50\%$ methyltransferase activity levels were candidate for EPR spectroscopy. Specific activities of His-Ste14-TA mutants were calculated as a percent compared to His-Ste14-TA. $[^{14}\text{C}]$ -SAM was used as the methyl donor and AFC was the methyl acceptor. Color code: blue- 0-25% His-Ste14-TA; green- 25-50% His-Ste14-TA; yellow- 50-75% His-Ste14-TA; red- 75-100% His-Ste14-TA; pink- over 100% His-Ste14-TA.

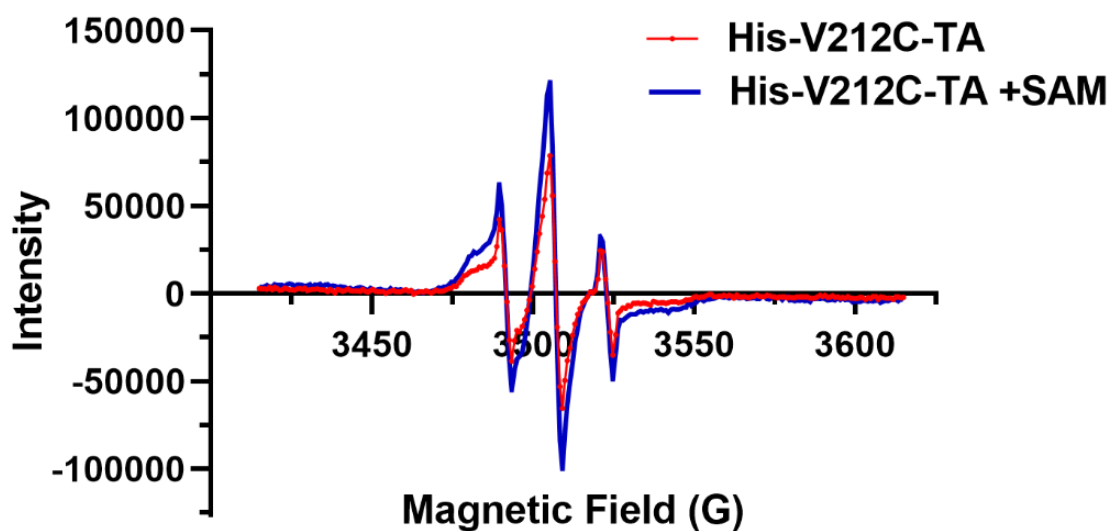


Figure 3.7 EPR spectral analysis of purified His-Ste14-V212C-TA. Purified His-Ste14-V212C-TA (red, 75 μ M) in 1xEQ buffer (50mM Na₂HPO₄, 0.3M NaCl, 2mM DTT, 0.05% DDM, pH 6.0) was added to glass capillary tube and inserted into a narrow quartz EPR tube. Spectra were collected as described in the methods section 3.2.9 at room temperature. Another sample of purified His-Ste14-V212C-TA (red, 75 μ M) in 1xEQ buffer (50mM Na₂HPO₄, 0.3M NaCl, 2mM DTT, 0.05% DDM, pH 6.0) was incubated with 5mM SAM followed by immediate data collection identical to the enzyme without substrate.

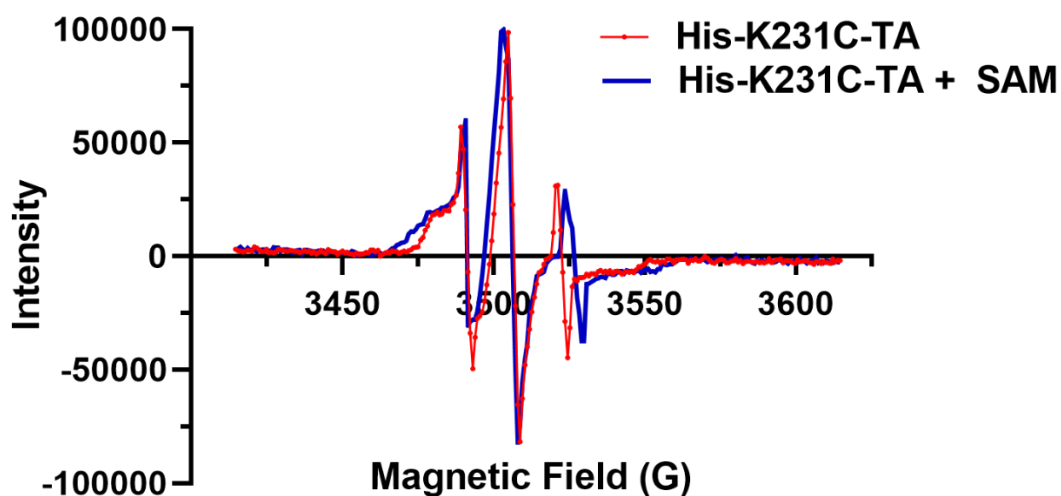


Figure 3.8 EPR spectral analysis of purified His-Ste14-K231C-TA. Purified His-Ste14-K231C-TA (red, 55 μ M) in 1xEQ buffer (50mM Na₂HPO₄, 0.3M NaCl, 2mM DTT, 0.05% DDM, pH 6.0) was added to glass capillary tube and inserted into a narrow quartz EPR tube. Spectra were collected as described in the methods section 3.2.9 at room temperature. Another sample of purified His-Ste14-K231C-TA (red, 55 μ M) in 1xEQ buffer (50mM Na₂HPO₄, 0.3M NaCl, 2mM DTT, 0.05% DDM, pH 6.0) was incubated with 5mM SAM followed by immediate data collection identical to the enzyme without substrate.

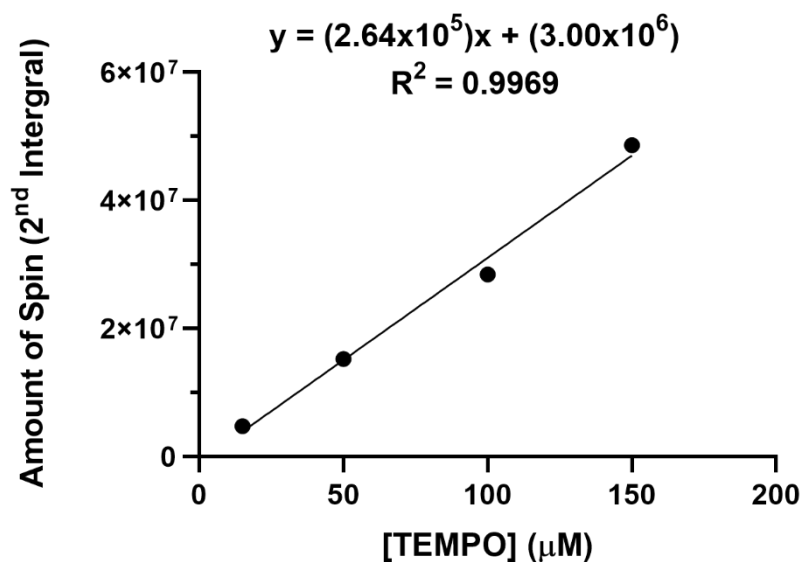


Figure 3.9 TEMPO standard curve.

Free TEMPO at varying concentrations in 1x EQ buffer (50mM Na₂HPO₄, 0.3M NaCl, pH 6.0 with the addition of 2mM DTT and 0.05% DDM) was added to glass capillary tubes and inserted into a narrow quartz EPR tube. Spectra were collected as described in the methods section 3.2.10 at room temperature. The difference in the second integral of the EPR data was determined for each concentration of TEMPO and a best-fit linear regression line was determined.

3.5 References

- 1 Hrycyna, C. A., Sapperstein, S. K., Clarke, S. & Michaelis, S. The *Saccharomyces cerevisiae* Ste14 gene encodes a methyltransferase that mediates C-terminal methylation of a-factor and Ras proteins. *Embo Journal* **10**, 1699-1709, doi:10.1002/j.1460-2075.1991.tb07694.x (1991).
- 2 Anderson, J. L., Frase, H., Michaelis, S. & Hrycyna, C. A. Purification, functional reconstitution, and characterization of the *Saccharomyces cerevisiae* isoprenylcysteine carboxylmethyltransferase Ste14p. *Journal of Biological Chemistry* **280**, 7336-7345, doi:10.1074/jbc.M410292200 (2005).
- 3 Diver, M. M., Pedi, L., Koide, A., Koide, S. & Long, S. B. Atomic structure of the eukaryotic intramembrane RAS methyltransferase ICMT. *Nature* **553**, 526-+, doi:10.1038/nature25439 (2018).
- 4 Yang, J. *et al.* Mechanism of Isoprenylcysteine carboxyl methylation from the crystal structure of the integral membrane methyltransferase ICMT. *Molecular Cell* **44**, 997-1004, doi:10.1016/j.molcel.2011.10.020 (2011).
- 5 Torres, J., Stevens, T. J. & Samso, M. Membrane proteins: the 'Wild West' of structural biology. *Trends in Biochemical Sciences* **28**, 137-144, doi:10.1016/s0968-0004(03)00026-4 (2003).
- 6 Weil, J. A. & Bolton, J. R. *Electron paramagnetic resonance: Elementary theory and practical applications*. (Wiley-Interscience, 2007).
- 7 Sahu, I. D. & Lorigan, G. A. Biophysical EPR studies applied to membrane proteins. *Journal of physical chemistry & biophysics* **5** (2015).
- 8 Altenbach, C., Marti, T., Khorana, H. G. & Hubbell, W. L. Transmembrane protein - structure - Spin labeling of bacteriorhodopsin mutants. *Science* **248**, 1088-1092, doi:10.1126/science.2160734 (1990).
- 9 Hubbell, W. L. & Altenbach, C. Investigation of structure and dynamics in membrane-proteins using site-directed spin-labeling. *Current Opinion in Structural Biology* **4**, 566-573, doi:10.1016/s0959-440x(94)90219-4 (1994).
- 10 Hubbell, W. L., Cafiso, D. S. & Altenbach, C. Identifying conformational changes with site-directed spin labeling. *Nature Structural Biology* **7**, 735-739, doi:10.1038/78956 (2000).
- 11 Columbus, L. & Hubbell, W. L. A new spin on protein dynamics. *Trends in Biochemical Sciences* **27**, 288-295, doi:10.1016/s0968-0004(02)02095-9 (2002).

- 12 Biswas, R., Kuhne, H., Brudvig, G. W. & Gopalan, V. Use of EPR spectroscopy to study macromolecular structure and function. *Science Progress* **84**, 45-68, doi:10.3184/003685001783239050 (2001).
- 13 Elble, R. A Simple and efficient procedure for transformation of yeasts. *Biotechniques* **13**, 18-20 (1992).
- 14 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254 (1976).
- 15 Hrycyna, C. A. & Clarke, S. Farnesyl cysteine C-terminal methyltransferase activity is dependent upon the Ste14 gene-product in *Saccharomyces-cerevisiae*. *Molecular and Cellular Biology* **10**, 5071-5076, doi:10.1128/mcb.10.10.5071 (1990).
- 16 Anderson, J. L., Henriksen, B. S., Gibbs, R. A. & Hrycyna, C. A. The isoprenoid substrate specificity of isoprenylcysteine carboxylmethyltransferase - Development of novel inhibitors. *Journal of Biological Chemistry* **280**, 29454-29461, doi:10.1074/jbc.M504982200 (2005).
- 17 Schaffne.W & Weissman.C. Rapid, sensitive, and specific method for determination of protein in dilute-solution. *Analytical Biochemistry* **56**, 502-514, doi:10.1016/0003-2697(73)90217-0 (1973).
- 18 Olsen, K. *Mechanism and conformational dynamics of the yeast isoprenylcysteine carboxyl methyltransferase* Ph.D. thesis, Purdue University, (2016).
- 19 Logue, A. *Characterization of the S-adenosyl-L-methionine (SAM) binding domain of the yeast isoprenylcysteine carboxyl methyltransferase, Ste14p* Ph.D. thesis, Purdue University, (2013).
- 20 Altenbach, C., Lopez, C. J., Hideg, K. & Hubbell, W. L. Exploring structure, dynamics, and topology of nitroxide spin-labeled proteins using continuous-wave electron paramagnetic resonance spectroscopy. *Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions, Pt B* **564**, 59-100, doi:10.1016/bs.mie.2015.08.006 (2015).
- 21 Borbat, P. P., McHaourab, H. & Freed, J. H. Protein structure determination using long-distance constraints from double-quantum coherence ESR: Study of T4 Lysozyme. *Biophysical Journal* **82**, 360A-360A (2002).
- 22 Jeschke, G. & Polyhach, Y. Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Physical Chemistry Chemical Physics* **9**, 1895-1910, doi:10.1039/b614920k (2007).

CHAPTER 4. PRELIMINARY STRUCTURAL ANALYSIS OF HIS-STE14 UTILIZING VARIOUS BIOPHYSICAL TECHNIQUES

4.1 Introduction

The Icmt family is the only known class of membrane-associated methyltransferases and does not contain any of the four motifs characteristically found in soluble methyltransferases. As a result, the Icmt family members make up the class VI methyltransferases.¹ Interestingly, in a primarily hydrophobic environment, Icmt has the ability to accommodate both a hydrophilic co-factor, *S*-adenosyl-L-methionine (SAM), as a methyl donor and a lipophilic methyl acceptor. There is a gap in the knowledge about the structure and mechanism of Icmt activity. Therefore, we study the yeast homolog, *Saccharomyces cerevisiae* Icmt, Ste14, to bridge these gaps. Our model system, His₁₀-myc₃N-Ste14 (His-Ste14), has been altered to include ten histidine residues (His₁₀-), and three myc tags (myc₃-) at the N-terminus of the enzyme in order to simplify the purification process and immunoblot analysis. Although there are two current crystal structures of Icmt from *Methanosarcina acetivorans* (Ma-Icmt) and *Tribolium castaneum* (beetle Icmt), the structure of Ste14 is still unclear and could provide more insights about this unique class of methyltransferase.^{1,2}

In order to gain structural insight of this enzyme, we pursued initial studies, including negative stain electron microscopy (EM) and single particle electron microscopy (cryo-EM). However, to obtain useful information from these techniques, it was advantageous to incorporate His-Ste14 into nanodiscs rather than detergents to improve stability of the protein and quality of structures.^{3,4} Nanodiscs consist of an aggregation of 120-160 lipids arranged in a bilayer held together by amphipathic proteins known as membrane scaffolding proteins (MSPs), and closely resembles a detergent micelle in structure.⁵⁻⁸ Specifically, we used MSP 1E3D1, which has a height of 4.5-5.6nm and a diameter of 9.7-12.9nm when self-assembled into a helical protein belt because Ste14 is predicted to have a diameter of ~9.5-10nm in its native dimeric state.^{7,9,10} During the protein purification process, the protein of interest is first solubilized using a mild detergent. Then, lipids and nanodiscs are incubated with the protein and the detergent is removed, leaving the targeted membrane protein enclosed within the nanodisc assembly.⁷ Nanodiscs are especially advantageous because they mimic a native-state condition and allow greater protein stability.⁹ Overall, this method is ideal for the extraction of transmembrane proteins from crude membranes.

However, this method requires an optimized purification protocol, specifically which detergent is used for solubilization in combination with nanodiscs. The evaluation of purified His-Ste14 in various detergents was provided in Chapter 2.

Once in a stable, native-like state, purified His-Ste14 was observed through negative stain EM, followed by cryo-EM. Negative stain EM is an advantageous technique because negative stain reconstructions can determine up to $\sim 18\text{\AA}$ resolution structures of the purified protein and act as a screening process to verify sample quality before proceeding to cryo-EM for a high resolution structure.¹¹⁻¹⁴

With single particle analysis, a molecular model of a protein (including membrane proteins) can be constructed at atomic resolution by using rapid freezing such that the sample is stabilized in vitreous ice.¹⁵⁻¹⁷ Additionally, compared to negative stain EM, cryo-EM offers the advantage of the images being less likely to contain stain artifacts and often yields a higher resolution model. Previously, only 3D constructions of larger proteins and complexes could be determined with this technique. Yet, recently the structures of smaller proteins have also been modelled, including a homodimer composed of two 30kDa proteins.^{18,19} Other small proteins successfully reconstructed with cryo-EM also include isocitrate dehydrogenase and hemoglobin at 93kDa and 64kDa, respectively.^{19,20}

4.2 Materials and Methods

4.2.1 Materials

Sf9 cells, High Five cells, and ES 921 media were purchased from Expression Systems (Davis, CA). The monoclonal α -myc IgG and the polyclonal goat- α -rabbit IgG HRP conjugate were purchased from Invitrogen Life Sciences (Carlsbad, CA). Enhanced chemiluminescence (ECL), pFastBac vectors, FEI Tecnai G2 20 - 200kV filament scanning microscope, Titan 80-300 kV environmental transmission electron microscope, and FEI Talos F200C transmission electron microscope were purchased from ThermoFisher (Waltham, MA). BigDye® Terminator v3.1 was obtained from Applied Biosystems (Foster City, CA) through the Purdue University Genomics Facility. GeneGnome XRQ was purchased from Syngene (Fredrick, MD). 4-(2-aminoethyl)-benzenesulfonylfluoride hydrochloride (AEBSF), aprotinin, and dithiothreitol (DTT) were purchased from Gold Biotechnology (St. Louis, MO). *E. coli* polar lipid extract was purchased from Avanti (Alabaster, AL). Membrane scaffold protein 1E3D1 was purchased from Sigma

Aldrich (St. Louis, MO). The substrate *N*-acetyl-*S*-farnesyl-*L*-cysteine (AFC) was purchased from Enzo Life Sciences (Farmingdale, NY). ¹⁴C-labelled *S*-adenosyl-*L*-methionine ([¹⁴C]-SAM) was purchased from Perkin Elmer (Waltham, MA). *N*-Dodecyl-β-D-maltopyranoside (DDM) was purchased from Anatrace (Maumee, OH). Nitrocellulose membrane (0.22 μm) and Superdex S200 column were purchased from GE Healthcare Life Sciences (Pittsburg, PA). Talon cobalt metal affinity resin was purchased from Clontech (Mountain View, CA). Amicon Ultra 30,000 MWCO concentrators and centrifugal filter units (0.22 μm) GV Durapore® were purchased from Millipore Co. (Billerica, MA). Poly-Prep® chromatography columns were purchased from Bio-Rad Laboratories (Hercules, CA). Biosafe II scintillation fluid and Bio-Beads™ SM-2 resin were purchased from Research Products International (Mount Prospect, IL). Carbon coated 400-mesh copper grids and uranyl formate were purchased from Electron Microscopy Sciences (Hatfield, PA). The US1000 2K x 2K CCD camera was a product from Gatan (Pleasanton, CA). All other materials and reagents were purchased from Fisher Scientific (Hampton, NH).

4.2.2 Expression and Membrane Preparation of His-Ste14 from Insect Cell Lines

cDNA encoding *S. cerevisiae* Ste14 was subcloned into the pFastBac-Dual vector with the addition of ten histidine residues (His₁₀-), and three *myc* tags with the sequence 5' EQKLISEEDL 3' (*myc3*-). The gene and vector sequence were confirmed with bidirectional dye-terminator sequencing (BigDye® Terminator v3.1). Protein was expressed in baculovirus-infected Sf9 or High Five cells and harvested from ES 921 media with 1x penicillin-streptomycin at 48 h post-infection. Pelleted cells were flash frozen for membrane preparation. Briefly, the insect cell pellets were thawed and resuspended in ~20mL of lysis buffer per 20g of cell pellet harvested (50 mM mannitol, 50 mM Tris-HCl, pH 7.5, 1% aprotinin, 1 mM AEBSF and 2mM DTT) and were rocked for 15 min at 4°C. Typically, 20g of cell pellet is equivalent to 750mL of cells harvested. The resuspended pellets were lysed via dounce homogenization and isolated through ultracentrifugation at 100,000 x g for 1 hour at 4°C. Following centrifugation, the membrane pellet was resuspended in buffer (300 mM mannitol, 50 mM Tris-HCl, pH 7.5, 1% aprotinin, 1 mM AEBSF and 1mM DTT, 10% glycerol), aliquoted, flash frozen in liquid nitrogen and stored at -80°C. Total protein concentration was determined using the Bradford protein assay.²¹

4.2.3 *In vitro* Methyltransferase Vapor Diffusion Assay of Yeast and Insect Cell Crude Membranes

An *in vitro* methyltransferase vapor diffusion assay was performed as previously described.²²⁻²⁴ Briefly, reactions were prepared in a 100mM Tris-HCl, pH 7.5 buffer, containing crude membranes overexpressing His-Ste14 (CH2704) or His-Ste14 isolated from insect cell lines (5µg), AFC (200µM), and [¹⁴C]-SAM (20µM). The negative control reaction contained a yeast deletion strain of *ste14Δ* (CH2714, 5µg), and all other components just described. Samples were incubated for 30 mins at 30°C and then were stopped with 50 µL of 1 M NaOH/1% SDS (v/v). Each reaction (100µL) was spotted onto a filter and placed in the neck of a scintillation vial filled with Biosafe II scintillation fluid (10mL) and capped at room temperature. After three hours, the filters were removed, and the vials were placed in a Packard Tri-Carb 1600CA liquid scintillation counter in order to detect the [¹⁴C]-methanol that diffused from filters into the scintillation fluid. All specific activities were characterized as pmol of methyl groups from [¹⁴C]-SAM transferred to the AFC substrate per minute of assay per mg of His-Ste14-TA (CH2803) or His-Ste14 mutant protein. These activities were calculated after subtracting background counts from the average of the negative control samples. Each experiment had three reactions per sample and all experiments were performed in triplicate.

4.2.4 SDS-PAGE Staining and Immunoblot Analysis

Crude membrane extracts overexpressing His-Ste14 from Sf9 cells (1µg) or purified His-Ste14 from Sf9 cells (0.1µg for immunoblot and 1.0µg for Coomassie stain) were separated on a 10% SDS-PAGE. Gels incubated with Coomassie stain (0.25% (w/v) Coomassie Brilliant Blue R-250, 80% methanol and 20% acetic acid) overnight, then incubated with Destaining Solution (30% methanol, 10% acetic acid) until all excess dye was removed. Once bands were visualized, the gel was rinsed in ddH₂O, and then incubated in Casting Solution (5% glycerol, 20% ethanol) and preserved in cellophane. For immunoblot analysis, gels were transferred to a nitrocellulose membrane (0.22 µm) and the nitrocellulose membrane was blocked in 20% (w/v) non-fat dry milk in phosphate-buffered saline (2.70 mM KCl, 137 mM NaCl, 4.00 mM Na₂HPO₄, 1.80 mM KH₂PO₄, pH 7.4) with 0.05% (v/v) Tween-20 (PBST) at room temperature for 3 hours. After three washes with PBST, the membranes were incubated with α-myc (1:10,000) antibody in 5% (w/v) non-fat dry milk in PBST at 4°C overnight. After an additional three washes with PBST, the

membranes were incubated with goat α -mouse IgG-HRP (1:2,500) antibody in 5% (w/v) non-fat dry milk in PBST at room temperature for 1 hour. The membrane was washed three more times with PBST and visualized with enhanced chemiluminescence (ECL) on a GeneGnome XRQ.

4.2.5 Purification of His-Ste14

Purification of His-Ste14 was performed as previously described.²² Briefly, the crude membranes overexpressing the protein of interest were solubilized in a solution of 0.02M imidazole, 1% DDM, and (2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer, pH 6.0 with the addition of 10% glycerol and 2mM AEBSF. For the solubilization reaction, crude membranes were at a final concentration of 5 mg/mL. The reaction was rocked at 4°C for 1 hr. The resulting mixture was centrifuged at 100,000 xg for 45 mins. The soluble fraction was incubated for 1 hour at 4°C with Talon® cobalt metal affinity beads equilibrated in HEPES buffer with the addition of 10% glycerol and 2mM AEBSF. After the protein was incorporated with the resin, the solution was centrifuged at 350 xg for 2 mins at 4°C and the remaining buffer was removed via aspiration. Next, the beads underwent four wash steps: HEPES buffer (pH 6.0) was added to the resin and was allowed to rock at 4°C for 10 mins followed by centrifugation at 350 xg for 2 min before aspiration of excess buffer. The resin was washed twice with buffer A (HEPES buffer, pH 6.0, with the addition of 1% DDM and 40mM imidazole), once with buffer B (HEPES buffer, pH 6.0, with the addition of 1% DDM, 500mM KCl, and 40mM imidazole), and once with buffer C (HEPES buffer, pH 6.0 with the addition of 0.1% DDM, 500mM KCl, and 40mM imidazole). After the final wash, buffer C was not aspirated, but instead loaded onto a Poly-Prep® chromatography column. His-Ste14 was eluted from the cobalt resin with elution buffer, pH 6.0 (0.1% DDM, 1M imidazole, and 2mM AEBSF).

Collected protein was concentrated at 5000 xg in a 30,000 MWCO concentrator at 4°C until the final volume was ~1mL. Concentrated protein was filtered in centrifugal filter units (0.22 μ m) GV Durapore® by centrifuging at 12,000 xg for 2 mins at 4°C. Once filtered, the His-Ste14 sample was applied to a Superdex S200 column pre-equilibrated with filtered Tris-HCl buffer (25mM Tris-HCl, a 50mM NaCl, pH 6.0) with the addition of 0.05% DDM and 2mM DTT at 4°C and spectra was collected at 280nm. Fractions containing purified protein were pooled, concentrated to a volume under 300 μ L in a new 30,000 MWCO concentrator and immediately incorporated into nanodiscs.

4.2.6 Incorporation of His-Ste14 into Nanodiscs

E. coli polar lipid extract (15mM) was dissolved in Tris-HCl buffer (1.5% DDM, 100 μ M NaCl, and 20mM Tris-HCl pH 6.0). Lipids were added to purify His-Ste14 (~300 μ M, section 4.2.5) at a 1:1 ratio and solubilized together for 10 minutes on ice. During this time, Bio-Beads™ SM-2 resin was washed three times with Tris-HCl buffer (1.5% DDM, 100 μ M NaCl, and 20mM Tris-HCl pH 6.0) at a 1:1 ratio. MSP 1E3D1 nanodiscs were added to the lipid, protein mixture at a 1:2:2 ratio respectively in Tris-HCl buffer (1.5% DDM, 100 μ M NaCl, and 20mM Tris-HCl pH 6.0). Bio-Beads™ SM-2 resin was added to the solution in order to remove excess detergent from the sample. The solution was rocked overnight at 4°C. Next, Talon® cobalt metal affinity beads were washed three times with Tris-HCl buffer (1.5% DDM, 100 μ M NaCl, and 20mM Tris-HCl pH 6.0). The His-Ste14, nanodisc, and lipid solution was removed from the Bio-Beads™ SM-2 resin and added to the cobalt resin for 1 hour at 4°C. The resin was washed three times with Tris-HCl buffer (1.5% DDM, 100 μ M NaCl, and 20mM Tris-HCl pH 6.0) by centrifugation at 8000 xg. His-Ste14 was eluted off the resin with Tris-HCl buffer containing 400mM imidazole. All eluted protein was pooled and concentrated at 5000 xg in a 30,000 MWCO concentrator at 4°C until the final volume was ~300 μ L. Concentrated protein was filtered in centrifugal filter units (0.22 μ m) GV Durapore® by centrifuging at 12,000 xg for 2 mins at 4°C. Once filtered, the His-Ste14 sample was applied to a Superdex S200 column pre-equilibrated with filtered Tris-HCl buffer (25mM Tris-HCl, 50mM NaCl, pH 6.0) with the addition of 0.05% DDM and 2mM DTT at 4°C and spectra were collected at 280nm. Again, fractions containing purified protein were pooled, concentrated to a volume under 300 μ L in a new 30,000 MWCO concentrator. Samples were either diluted to concentration needed for negative stain EM, cryo-EM, or flash frozen in liquid nitrogen and stored at -80°C. Pure protein concentration was determined using an Amido Black protein assay.²⁵

4.2.7 Negative Stain EM Sample Preparation and Imaging

Purified His-Ste14 (3 μ L at 0.05 mg/mL) was applied to glow-discharged carbon coated 400-mesh copper grids and stained with 0.75% uranyl formate.²⁶ Samples were imaged with a FEI Tecnai G2 T20, 200 kV LaB6 filament scanning transmission electron microscope (TEM) operated at 200 kV. Micrographs were collected with a US1000 2K x 2K CCD camera, using a 1 second

exposure time, and an intended defocus range of -1.19 to -1.6 μm . The magnification was 71,000, resulting in a sampling of 1.56 $\text{\AA}/\text{pixel}$ at the protein level.

4.2.8 Cryo-EM Sample Preparation and Imaging

Purified His-Ste14 (3 μL at 0.8 mg/mL) was applied to glow-discharged carbon coated 400-mesh copper grids and vitrified using a Vitrobot.²⁶ Samples were imaged with a FEI Talos F200C transmission electron microscope (TEM) with Ceta 4K x 4K camera and K2 summit direct electron detector for initial screening. Titan 80-300 kV environmental transmission electron microscope (TEM) operated at 300 kV was used to obtain higher resolution images. Micrographs were collected using a Falcon 3EC direct electron detector, 3.2 second exposure time, and an intended defocus range of 1.5 to 2.5 μm . The magnification was 81,000, resulting in a sampling of 0.54 $\text{\AA}/\text{pixel}$ at the protein level and dose of 53.67 $\text{e}^-/\text{\AA}^2$.

4.2.9 Introduction of His-Ste14 into Nanodiscs

Initial negative stain EM images of purified His-Ste14 in DDM were taken with unsuccessful results (data not shown). 1xEQ and HEPES buffer, pH 6.0 were used for the preliminary analysis. Both buffers resulted in low protein homogeneity and little contrast between sample and carbon background. Therefore, a new approach was needed to obtain structural analysis of His-Ste14. As discussed earlier, the nanodisc system allows for the stabilization of membrane proteins in a stable and monodisperse state.²⁷ Thus, we successfully incorporated His-Ste14 into MSP 1E3D1 nanodiscs (Figure 4.1). Size exclusion chromatography (SEC) was used to purify His-Ste14 and remove any excess nanodiscs in the buffer solution after solubilization (Figure 4.1A). Fractions collected between ~12mL and 17mL were identified as His-Ste14, while ~25mL to ~27mL was empty nanodiscs (Figure 4.1B-D). After SEC fractions were pooled and concentrated, the methyltransferase activity was examined compared to previously purified His-Ste14 in DDM without being incorporated in nanodiscs (Figure 4.1B). Compared to the positive control, His-Ste14 in nanodiscs retained 88.3% activity. There was no statistical difference in the methyltransferase activity levels based on one-way ANOVA followed by Dunnett's multiple comparisons test vs. His-Ste14 control. The small percent of activity that was lost could be due to the overnight incubation with the nanodiscs at 4°C. His-Ste14 is unstable, even at cool temperature, and may have broken down during that necessary protocol step. Immunoblot and Coomassie stain

analysis confirmed His-Ste14 incorporation into nanodiscs (Figure 4.1 C & D). Nanodiscs do not have *myc*-tags and will not appear on any immunoblots but will appear in Coomassie stains. Wash steps during the purification process indicated nanodiscs were present on the Coomassie stain (~30kDa) but were not present on the immunoblot. During the wash steps, excess nanodiscs were successfully being removed from the buffer. In the Coomassie, the fractions from SEC exhibited two bands, His-Ste14 (~35kDa) and nanodiscs (~30kDa), indicating that His-Ste14 is incorporated into the nanodisc system. While in the immunoblot, only one band we visualized at ~35kDa, indicating His-Ste14 presence. Both gels had two controls: (1) a purified His-Ste14 control that was purified in DDM and (2) empty nanodiscs. The control bands align with the nanodiscs and pure His-Ste14 in our final sample, confirming our results.

4.2.10 Negative Staining with His-Ste14 Incorporated into Nanodiscs

In negative stain EM, His-Ste14 incorporated into nanodiscs was adsorbed onto the surface of an EM grid that had undergone glow-discharge. The top layer of the grid was comprised of a thin layer of carbon, a hydrophobic surface. Glow-discharge is used to negatively charge the carbon grid so polar substances (purified His-Ste14) will adhere to the surface. Based on Ohi *et al.*, the His-Ste14 (0.05mg/mL) was quickly enveloped with uranyl formate (0.75%), an electron-dense stain.²⁸ The addition of the stain created a high contrast between the background (stained) and the protein (unstained) due to His-Ste14 being less electron dense than the uranyl formate background.²⁶ After the sample was washed with ddH₂O and dried, the grid was imaged with the FEI Tecnai G2 T20, transmission electron microscope (TEM) (Figure 4.2). Distinguishable particles were seen in all images similar in size to MSP 1E3D1 nanodiscs. Although the particles could be empty nanodiscs, we believe His-Ste14 is still embedded inside the scaffolding protein based on the Coomassie stained SDS-PAGE, immunoblot, and methyltransferase activity assay described in section 4.3.1. His-Ste14 is a small membrane protein that we believe is mostly within membrane and therefore, would not alter the dimensions of an empty nanodisc.^{29,30} Preliminary analysis of the data sets needs to be performed to determine dominant 2D class averages for each set of particles.

4.2.11 Preliminary Cryo-EM Data Analysis

One attempt at cryo-EM was attempted with His-Ste14 incorporated into nanodisc (Figure 4.3). During the screening processes, micrographs contained particles in the ice at approximately the expected size (has a height of 4.5-5.6nm and a diameter of 9.7-12.9nm). Unfortunately, when higher resolution images were obtained, the ice appeared to have melted and 2D reconstructions were not possible. We must repeat this experiment under the exact same conditions to potential obtain critical information about the structure of His-Ste14. These promising results provide evidence that particles are observable and that it could be possible to do a full data collection and possible reconstruction.

4.3 Future Directions

Cryo-EM was attempted due to a promising initial Talos screening. Unfortunately, samples assessed on the Titan transmission electron microscope melted prior to data collection (Figure 4.3). This could occur during grid freezing, transfer from forceps to storage box, grid dripping, or upon transfer of the clipped grid to the autoloader. Thus, it would be advantageous to repeat the same experiment as described above. This is an extremely promising field of study and should be pursued. Additionally, the His-Ste14 negative stain EM could undergo 2D class averaging for each set of particles (Figure 4.2). The goal of this technique is to determine the density of an object from a set of 2D projections obtained from a set of low-dose cryo-EM images of individual macromolecules.³¹⁻³³ Our collected data would be processed with RELION 3.1 and raw micrographs would be CTF-corrected. Although the resolution of negative stain EM structures are low, it would provide better insight on the basic structure and oligomeric state of His-Ste14.

There are also other purification conditions worth troubleshooting to aid in the production of better-quality images for negative staining EM and cryo-EM. His-Ste14 could undergo purification with a lower salt concentration, solubilization with detergents like C₁₂E₇ or LMNG, or the utilization of affinity grids, antibodies, or graphene grids.³⁴⁻³⁸ All of these conditions have improved the quality of protein structures previously published.

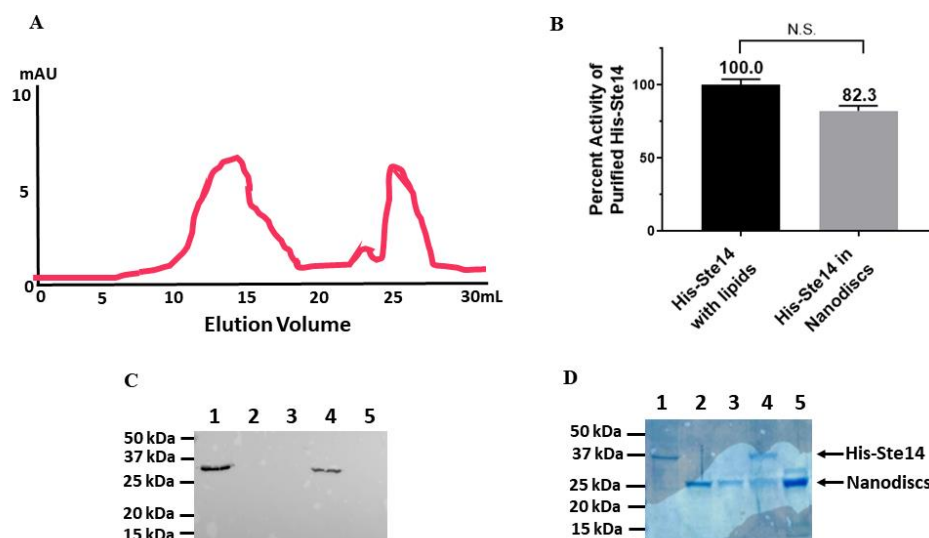


Figure 4.1

Confirmation of His-Ste14 incorporated into MSP 1E3D1 nanodiscs.

(A) Size exclusion chromatography (SEC) analysis of His-Ste14 after solubilization with MSP 1E3D1 nanodiscs. (B) *In vitro* vapor diffusion methyltransferase activity assay of His-Ste14 after SEC purification in nanodiscs. All specific activities were characterized as pmol of methyl groups from [^{14}C]-SAM transferred to the AFC substrate per minute of assay per mg of purified His-Ste14. These activities were calculated after subtracting out background counts from the average of the negative control samples (ddH₂O added to buffer in place of pure protein). Percent activity compared to His-Ste14 purified in DDM at pH 6.0 (not in nanodiscs) was calculated. Each experiment had three samples per condition \pm S.D. No significant difference between the two values based on a ratio paired T-test comparison of His-Ste14 in nanodiscs vs. His-Ste14 control. (C) Expression levels of purified His-Ste14 in nanodiscs indicated by immunoblotting: protein samples (0.1 μg) were run on a 10% SDS-PAGE followed by immunoblot analysis with α -myc (1:10,000) and goat α -mouse IgG-HRP (1:2,500). Immunoblots were visualized with ECL. (D) Determination of purity of protein samples via Coomassie staining on 10% SDS-PAGE. (C & D) Samples in each lane were as followed: Lane 1 – Purified His-Ste14 solubilized in DDM (positive control), Lane 2 and 3 – Buffer from wash steps of Talon® cobalt metal affinity resin after nanodisc solubilization, Lane 4 - Purified His-Ste14 in nanodiscs, peak that eluted at ~12-15mL, and Lane 5 – Empty nanodiscs.

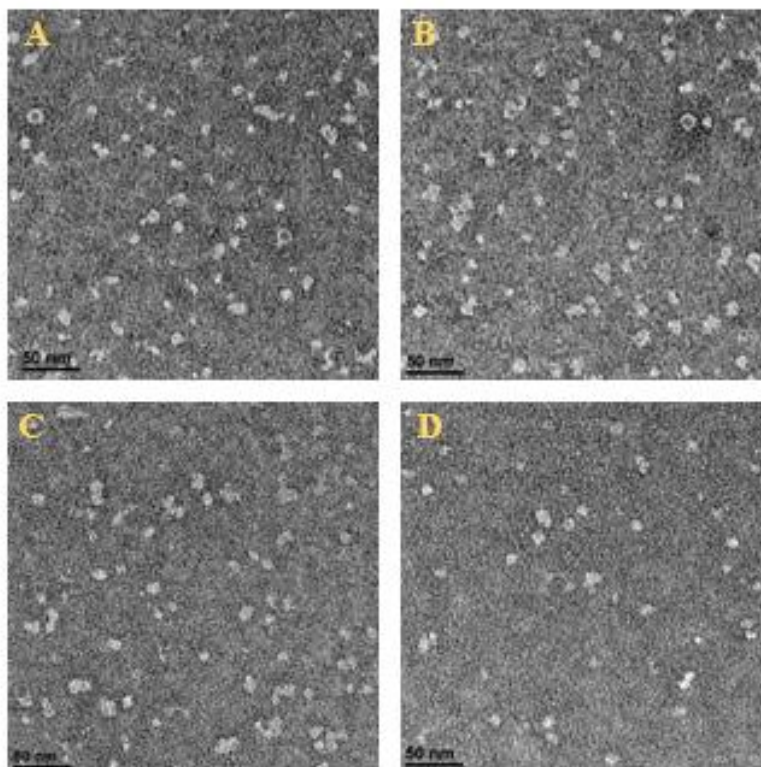


Figure 4.2 Unadjusted micrograph field views of His-Ste14 in nanodiscs using negative staining EM.

(A-D) Different field views of His-Ste14, indicating consistence in size and shape. Scale bars were 50nm and added using Gatan Inc., version 2.31.734.0



Figure 4.3 Unadjusted micrograph field views of His-Ste14 in nanodiscs using negative cryo-EM.

Images were taken with the FEI Talos F200C transmission electron microscope and Falcon 3EC direct electron detector. Homogenous particles with good contrast were observed. Thank you to Prof. Nicolas Noinaj's laboratory for aiding in cryo-EM data collection.

4.4 References

- 1 Yang, J. *et al.* Mechanism of Isoprenylcysteine Carboxyl Methylation from the Crystal Structure of the Integral Membrane Methyltransferase ICMT. *Molecular Cell* **44**, 997-1004, doi:10.1016/j.molcel.2011.10.020 (2011).
- 2 Diver, M. M., Pedi, L., Koide, A., Koide, S. & Long, S. B. Atomic structure of the eukaryotic intramembrane RAS methyltransferase ICMT. *Nature* **553**, 526-+, doi:10.1038/nature25439 (2018).
- 3 Gao, Y., Cao, E., Julius, D. & Cheng, Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature* **534**, 347-+, doi:10.1038/nature17964 (2016).
- 4 Gatsogiannis, C. *et al.* Membrane insertion of a Tc toxin in near-atomic detail. *Nature Structural & Molecular Biology* **23**, 884-890, doi:10.1038/nsmb.3281 (2016).
- 5 Denisov, I. G. & Sligari, S. G. Nanodiscs in Membrane Biochemistry and Biophysics. *Chemical Reviews* **117**, 4669-4713, doi:10.1021/acs.chemrev.6b00690 (2017).
- 6 Noguchi, H. Structure formation in binary mixtures of lipids and detergents: Self-assembly and vesicle division. *Journal of Chemical Physics* **138**, doi:10.1063/1.4774324 (2013).
- 7 Bayburt, T. H. & Sligar, S. G. Membrane protein assembly into Nanodiscs. *Febs Letters* **584**, 1721-1727, doi:10.1016/j.febslet.2009.10.024 (2010).
- 8 Denisov, I. G. & Sligar, S. G. Nanodiscs for structural and functional studies of membrane proteins. *Nature Structural & Molecular Biology* **23**, 481-486, doi:10.1038/nsmb.3195 (2016).
- 9 Nath, A., Atkins, W. M. & Sligar, S. G. Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins. *Biochemistry* **46**, 2059-2069, doi:10.1021/bi602371n (2007).
- 10 Denisov, I. G., Grinkova, Y. V., Lazarides, A. A. & Sligar, S. G. Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *Journal of the American Chemical Society* **126**, 3477-3487, doi:10.1021/ja0393574 (2004).
- 11 De Carlo, S. & Harris, J. R. Negative staining and cryo-negative staining of macromolecules and viruses for TEM. *Micron* **42**, 117-131, doi:10.1016/j.micron.2010.06.003 (2011).
- 12 Ha, J. Y. *et al.* Molecular architecture of the complete COG tethering complex. *Nature Structural & Molecular Biology* **23**, 758-+, doi:10.1038/nsmb.3263 (2016).

- 13 Burgess, S. A., Walker, M. L., Thirumurugan, K., Trinick, J. & Knight, P. J. Use of negative stain and single-particle image processing to explore dynamic properties of flexible macromolecules. *Journal of Structural Biology* **147**, 247-258, doi:10.1016/j.jsb.2004.04.004 (2004).
- 14 Fabre, L., Bao, H., Innes, J., Duong, F. & Rouiller, I. Negative Stain Single-particle EM of the Maltose Transporter in Nanodiscs Reveals Asymmetric Closure of MalK(2) and Catalytic Roles of ATP, MalE, and Maltose. *Journal of Biological Chemistry* **292**, 5457-5464, doi:10.1074/jbc.M116.757898 (2017).
- 15 Dubochet, J. *et al.* Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics* **21**, 129-228, doi:10.1017/s0033583500004297 (1988).
- 16 Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic-resolution microscopy of unstained biological molecules. *Quarterly Reviews of Biophysics* **28**, 171-193, doi:10.1017/s003358350000305x (1995).
- 17 Henderson, R. Realizing the potential of electron cryo-microscopy. *Quarterly Reviews of Biophysics* **37**, 3-13, doi:10.1017/s0033583504003920 (2004).
- 18 Zhang, K. M. *et al.* Structure of the 30 kDa HIV-1 RNA Dimerization Signal by a Hybrid Cryo-EM, NMR, and Molecular Dynamics Approach. *Structure* **26**, 490-+, doi:10.1016/j.str.2018.01.001 (2018).
- 19 Merk, A. *et al.* Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell* **165**, 1698-1707, doi:10.1016/j.cell.2016.05.040 (2016).
- 20 Khoshouei, M., Danev, R., Plitzko, J. M. & Baumeister, W. Revisiting the Structure of Hemoglobin and Myoglobin with Cryo-Electron Microscopy. *Journal of Molecular Biology* **429**, 2611-2618, doi:10.1016/j.jmb.2017.07.004 (2017).
- 21 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254 (1976).
- 22 Anderson, J. L., Frase, H., Michaelis, S. & Hrycyna, C. A. Purification, functional reconstitution, and characterization of the *Saccharomyces cerevisiae* isoprenylcysteine carboxylmethyltransferase Ste14p. *Journal of Biological Chemistry* **280**, 7336-7345, doi:10.1074/jbc.M410292200 (2005).
- 23 Hrycyna, C. A. & Clarke, S. Farnesyl cysteine c-terminal methyltransferase activity is dependent upon the ste14 gene-product in *saccharomyces-cerevisiae*. *Molecular and Cellular Biology* **10**, 5071-5076, doi:10.1128/mcb.10.10.5071 (1990).
- 24 Anderson, J. L., Henriksen, B. S., Gibbs, R. A. & Hrycyna, C. A. The isoprenoid substrate specificity of isoprenylcysteine carboxylmethyltransferase - Development of novel inhibitors. *Journal of Biological Chemistry* **280**, 29454-29461, doi:10.1074/jbc.M504982200 (2005).

- 25 Schaffne.W & Weissman.C. Rapid, sensitive, and specific method for determination of protein in dilute-solution. *Analytical Biochemistry* **56**, 502-514, doi:10.1016/0003-2697(73)90217-0 (1973).
- 26 Ohi, M., Li, Y., Cheng, Y. & Walz, T. Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. *Biological procedures online* **6**, 23-34, doi:10.1251/bpo70 (2004).
- 27 Denisov, I. G., Schuler, M. A. & Sligar, S. G. Nanodiscs as a New Tool to Examine Lipid-Protein Interactions. *Methods in molecular biology (Clifton, N.J.)* **2003**, 645-671, doi:10.1007/978-1-4939-9512-7_25 (2019).
- 28 Scarff, C. A., Fuller, M. J. G., Thompson, R. F. & Iadaza, M. G. Variations on Negative Stain Electron Microscopy Methods: Tools for Tackling Challenging Systems. *Jove-Journal of Visualized Experiments*, doi:10.3791/57199 (2018).
- 29 Romano, J. D. & Michaelis, S. Topological and mutational analysis of *Saccharomyces cerevisiae* Ste14p, founding member of the isoprenylcysteine carboxyl methyltransferase family. *Molecular Biology of the Cell* **12**, 1957-1971, doi:10.1091/mbc.12.7.1957 (2001).
- 30 Wright, L. P. *et al.* Topology of Mammalian Isoprenylcysteine Carboxyl Methyltransferase Determined in Live Cells with a Fluorescent Probe. *Molecular and Cellular Biology* **29**, 1826-1833, doi:10.1128/mcb.01719-08 (2009).
- 31 Penczek, P. A. Fundamentals of three-dimensional reconstruction from projections. *Methods in Enzymology, Vol 482: Cryo-Em, Part B: 3-D Reconstruction* **482**, 1-33, doi:10.1016/s0076-6879(10)82001-4 (2010).
- 32 Grigorieff, N. Three-dimensional structure of bovine NADH : Ubiquinone oxidoreductase (Complex I) at 22 angstrom in ice. *Journal of Molecular Biology* **277**, 1033-1046, doi:10.1006/jmbi.1998.1668 (1998).
- 33 Penczek, P. A., Grassucci, R. A. & Frank, J. The ribosome at improved resolution - new techniques for merging and orientation refinement in 3D cryoelectron microscopy of biological particles. *Ultramicroscopy* **53**, 251-270, doi:10.1016/0304-3991(94)90038-8 (1994).
- 34 Lyumkis, D. *et al.* Cryo-EM Structure of a Fully Glycosylated Soluble Cleaved HIV-1 Envelope Trimer. *Science* **342**, 1484-1490, doi:10.1126/science.1245627 (2013).
- 35 Fernandez-Leiro, R., Conrad, J., Scheres, S. H. W. & Lamers, M. H. cryo-EM structures of the *E. coli* replicative DNA polymerase reveal its dynamic interactions with the DNA sliding clamp, exonuclease and tau. *Elife* **4**, doi:10.7554/eLife.11134 (2015).
- 36 Kelly, D. F., Dukovski, D. & Walz, T. Monolayer purification: A rapid method for isolating protein complexes for single-particle electron microscopy. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4703-4708, doi:10.1073/pnas.0800867105 (2008).

- 37 Earl, L. A., Falconieri, V., Milne, J. L. S. & Subramaniam, S. Cryo-EM: beyond the microscope. *Current Opinion in Structural Biology* **46**, 71-78, doi:10.1016/j.sbi.2017.06.002 (2017).
- 38 Llaguno, M. C. *et al.* Chemically functionalized carbon films for single molecule imaging. *Journal of Structural Biology* **185**, 405-417, doi:10.1016/j.jsb.2014.01.006 (2014).

CHAPTER 5. CHARACTERIZATION OF THE OLIGOMERIC STATE OF STE14 AND THE GXXXG DIMERIZATION MOTIF

5.1 Introduction

Griggs *et al* previously reported that Ste14 functions as a dimer or higher order oligomer using biochemical co-immunoprecipitation and dominant-negative activity experiments.¹ Ste14 possesses a tandem S₂₇G₃₁G₃₅G₃₉ motif located in transmembrane domain one (TM1) (Figure 5.1). This repeating GXXXG motif was determined to be important for the dimerization of several membrane proteins, the first being glycophorin A (GpA).^{2,3} Since the discovery of this repeating GxxxG sequence, other integral membrane proteins with the dimerization motif have been identified including, ATP binding cassette subfamily G member 2 (ABCG2), yeast ATP synthase, human carbonic anhydrase (CA), *Helicobacter pylori* vacuolating toxin, and yeast α -factor receptor.⁴⁻⁹ Senes *et al.* used genetic and evolutionary studies to determine that the GxxxG motif was commonly found in transmembrane domain sequences, suggesting that the sequence had evolved as a necessary motif.¹⁰

The chemical characteristics of the glycine residues in the GxxxG motif are essential for dimerization formation. The glycine amino acid side chain is composed of one hydrogen atom, making the residue small and nonpolar. Since the glycine residues in the repeating GxxxG motif are every four amino acids, they are positioned on the same face of their transmembrane alpha helix tertiary structure. Due to the small amino acid side chain, the glycine's create a flat surface on the alpha helix, which permits the close packing of two GxxxG motifs to form a dimer in the membrane. Additionally, it has been shown that the residues surrounding each glycine, such as serine, threonine, and tyrosine, are also important for dimerization due to the formation of potential intramolecular hydrogen bonds.¹¹ Moreover, leucine, valine, and isoleucine residues can form van der Waals interactions, which can further stabilize the protein dimer.¹¹

We predict Ste14 dimerizes through TM1 at the N-terminal end of two monomers of the enzymes. Overall, there are no crystal structures or three-dimensional models of Ste14 or hIcmt due to the difficulty of purifying and crystallizing integral membrane proteins. Currently, two Icmt crystal structure models exist: an archaea Icmt ortholog, *Methanosarcina acetivorans* (Ma-Icmt), and a eukaryotic Icmt, *Tribolium castaneum* Icmt (beetle-Icmt).^{12,13} Sequence alignment analysis showed that the dimerization motif is present in eukaryotes domain but not in the archaea, Ma-

Icmt (Figure 5.2). Although a useful original model, the *Ma*-Icmt crystal structure shares low sequence identity and similarity to hIcmt, ~32% and ~50%, respectively. Specifically, only residues in the C-terminal half of the enzyme were conserved. This model was not ideal because the enzyme has two fewer helices than Ste14 and does not include the proposed dimerization motif.¹² The beetle Icmt sequence shares 58% identity and 72% similarity with hIcmt and both proteins contain eight transmembrane domains.¹³ Crystal structure of beetle Icmt revealed a monomer.¹³ In this enzyme, a GxxxG motif can be found in TM1 and TM3.¹³ Long *et al.* predicted the two GxxxG stabilize each other rather than form a dimer.¹³ TM3 of beetle Icmt corresponds to TM1 of Ste14, while TM1 of beetle Icmt does not exist in Ste14.¹³

Previously, our lab assessed the effects of changing the amino acid residues of His-Ste14 comprising the GXXXG motif on protein dimerization, expression, activity, and folding (see publications). First, wild-type His-Ste14 and cysteine-less, triple alanine mutant, His-Ste14-TA, were expressed in *Saccharomyces cerevisiae* (*S. cerevisiae*) and we demonstrated their ability to form homodimers following the addition of glutaraldehyde. Additionally, we performed cysteine scanning mutagenesis on residues in TM1, which were used to determine which amino acids are in close enough proximity to promote in dimerization. Briefly, sulfhydryl-specific crosslinkers of varying-length, *Bis*-maleimidoethane (BMOE) and *Bis*-maleimidohexane (BMH), were used to investigate the ability of each amino acid in TM1 to form a dimer. Cross-linking with the sulfhydryl reactive 13 Å cross-linker, BMH, revealed that cysteine mutants on that same side of the alpha helix in TM1 formed strong dimers as observed in immunoblot analysis. Similar results were obtained with the shorter 8 Å cross-linker, BMOE. Together, these results indicate that the residues on the same side of the alpha helix formed by TM1, contribute to the dimer formation of His-Ste14 as shown in helical wheel projections. Finally, these results were confirmed with tryptic digestion experiments followed by immunoblot analysis. During tryptic digestion, trypsin cleaves protein at lysine and arginine amino acids. We found that His-Ste14 and His-Ste14-TA have the same cleavage pattern when visualized by immunoblot. If the His-Ste14-TA cysteine mutant is misfolded due to the introduction of the cysteine residue in TM1, the cleavage sites may be altered. Several mutants had altered cleavage patterns including all the glycine residues in the GxxxG repeating motif (His-G31C-TA, His-G35C-TA, His-G39C-TA), as well as His-S27C-TA, His-Y28C-TA, and His-L30C-TA. Therefore, we believe these residues are important for protein folding and potentially for dimerization.

Here, a combination of size-exclusion chromatography, multi-angle light scattering and small-angle X-ray scattering (SEC-MALS-SAXS) experiments were used to determine the oligomeric state of His-Ste14. These biophysical experiments corroborate the biochemical finding that His-Ste14 forms a dimer.

5.2 Materials and Methods

5.2.1 Materials

Sf9 cells and ES 921 media were purchased from Expression Systems (Davis, CA). The monoclonal α -myc IgG and the polyclonal goat- α -rabbit IgG HRP conjugate were purchased from Invitrogen Life Sciences (Carlsbad, CA). Enhanced chemiluminescence (ECL), The pFastBac-Dual vector was purchased from ThermoFisher (Waltham, MA). BigDye® Terminator v3.1 was obtained from Applied Biosystems (Foster City, CA) through the Purdue University Genomics Facility. 4-(2-aminoethyl)-benzenesulfonylfluoride hydrochloride (AEBSF), aprotinin, and dithiothreitol (DTT) were purchased from Gold Biotechnology (St. Louis, MO). The substrate *N*-acetyl-*S*-farnesyl-*L*-cysteine (AFC) was purchased from Enzo Life Sciences (Farmingdale, NY). 14 C-labelled *S*-adenosyl-*L*-methionine (14 C]-SAM) was purchased from Perkin Elmer (Waltham, MA). *N*-Dodecyl- β -D-maltopyranoside (DDM) was purchased from Anatrace (Maumee, OH). Nitrocellulose membrane (0.22 μ m) and Superdex S200 column were purchased from GE Healthcare Life Sciences (Pittsburg, PA). Talon® cobalt metal affinity resin was purchased from Clontech (Mountain View, CA). Amicon Ultra 30,000 MWCO concentrators and centrifugal filter units (0.22 μ m) GV Durapore® were purchased from Millipore Co. (Billerica, MA). Poly-Prep® chromatography columns were purchased from Bio-Rad Laboratories (Hercules, CA). Biosafe II scintillation fluid and Bio-Beads™ SM-2 resin were purchased from Research Products International (Mount Prospect, IL). All other materials and reagents were purchased from Fisher Scientific (Hampton, NH).

5.2.2 Expression and Membrane Preparation of His-Ste14 from Insect Cell Lines

cDNA encoding *S. cerevisiae* His-Ste14 was subcloned into the pFastBac-Dual vector with the addition of ten histidine residues (His₁₀-), and three *myc* tags with the sequence 5' EQKLISEEDL 3' (*myc3*-). The gene and vector sequence were confirmed with bidirectional dye-terminator sequencing (BigDye® Terminator v3.1). Protein was expressed in baculovirus-infected

Sf9 and harvested from ES 921 media with 1x penicillin-streptomycin at 48 h post-infection. Pelleted cells were flash frozen for crude membrane preparation. Briefly, the insect cell pellets were thawed and resuspended in ~20mL of lysis buffer per 20g of cell pellet harvested (50 mM mannitol, 50 mM Tris-HCl, pH 7.5, 1% aprotinin, 1 mM AEBSF and 2mM DTT) and were rocked for 15 min at 4°C. Typically, 20g of cell pellet is equivalent to 750mL of cells harvested. The resuspended pellets were lysed via dounce homogenization and isolated via ultracentrifugation at 100,000 x g for 1 hour at 4°C. Following centrifugation, the membrane pellet was resuspended in ~2mL buffer (300 mM mannitol, 50 mM Tris-HCl, pH 7.5, 1% aprotinin, 1 mM AEBSF and 1mM DTT, 10% glycerol), aliquoted, flash frozen in liquid nitrogen and stored at -80°C. Total protein concentration was determined using the Bradford protein assay.¹⁴

5.2.3 *In vitro* Methyltransferase Vapor Diffusion Assay of Yeast and Insect Cell Crude Membranes

An *in vitro* methyltransferase vapor diffusion assay was performed as previously described.¹⁵ Briefly, reactions were prepared in a 100mM Tris-HCl, pH 7.5 buffer, containing crude membranes overexpressing His-Ste14 (CH2704) or His-Ste14 isolated from insect cell lines (5µg), AFC (200µM), and [¹⁴C]-SAM (20µM). The negative control reaction contained His-Ste14 (CH2714, 5µg), a yeast deletion strain of *ste14Δ*, and all other components described above. Samples were incubated for 30 mins at 30°C and then were stopped with 50 µL of 1 M NaOH/1% SDS (v/v). Each reaction (100µL) was spotted onto a filter and placed in the neck of a scintillation vial filled with Biosafe II scintillation fluid (10mL) and capped at room temperature. After three hours, the filters are thrown out and the vials were placed in a Packard Tri-Carb 1600CA liquid scintillation counter in order to detect the [¹⁴C]-methyl that diffused off the filters. All specific activities were characterized as pmol of methyl groups from [¹⁴C]-SAM transferred to the AFC substrate per minute of assay per mg of His-Ste14 (CH2704) or His-Ste14 isolated from insect cell lines. These activities were calculated after subtracting out background counts from the average of the negative control samples. Each experiment was performed in triplicate.

5.2.4 SDS-PAGE Staining and Immunoblot Analysis

Crude membrane extracts overexpressing His-Ste14 from Sf9 cells (1µg) or purified His-Ste14 from Sf9 cells (0.1µg for immunoblot and 1.0µg for Coomassie stain) were separated on a

10% SDS-PAGE. Gels incubated with Coomassie stain (0.25% (w/v) Coomassie Brilliant Blue R-250, 80% methanol and 20% acetic acid) overnight, then incubated with Destaining Solution (30% methanol, 10% acetic acid) until all excess dye was removed. Once bands were visualized, the gel was rinsed in ddH₂O, and then incubated in Casting Solution (5% glycerol, 20% ethanol) and preserved in cellophane. For immunoblot analysis, gels were transferred to a nitrocellulose membrane (0.22 μ m) and the nitrocellulose membrane was blocked in 20% (w/v) non-fat dry milk in phosphate-buffered saline (2.70 mM KCl, 137 mM NaCl, 4.00 mM Na₂HPO₄, 1.80 mM KH₂PO₄, pH 7.4) with 0.05% (v/v) Tween-20 (PBST) at room temperature for 3 hours. After three washes with PBST, the membranes were incubated with α -myc (1:10,000) antibody in 5% (w/v) non-fat dry milk in PBST at 4°C overnight. After an additional three washes with PBST, the membranes were incubated with goat α -mouse IgG-HRP (1:2,500) antibody in 5% (w/v) non-fat dry milk in PBST at room temperature for 1 hour. The membrane was washed three more times with PBST and visualized with enhanced chemiluminescence (ECL) on a GeneGnome XRQ.

5.2.5 Purification of His-Ste14

Purification of His-Ste14 was performed as previously described.¹⁵ Briefly, the crude membranes overexpressing the protein of interest were solubilized in a solution of 0.02M imidazole, 1% DDM, 1xEQ buffer (50mM Na₂HPO₄, 300 mM NaCl, pH 6.0) with the addition of 10% glycerol and 2mM AEBSF. It is essential for the pH to be 6.0 or else protein will aggregate (see chapter 2). For the solubilization reaction, crude membranes were at a final concentration of 5 mg/mL. The reaction was rocked at 4°C for 1 hr. The resulting mixture was centrifuged at 100,000 xg for 45 mins. The soluble fraction was incubated for 1 hour at 4°C with Talon® cobalt metal affinity beads equilibrated in 1xEQ buffer (pH 6.0) with the addition of 10% glycerol and 2mM AEBSF. After the protein was bound to the resin, the solution was centrifugated at 350 xg for 2 mins at 4°C and the remaining buffer was removed via aspiration. Next, the beads underwent four wash steps: 1xEQ buffer (pH 6.0) buffer was added to the resin and was allowed the rock at 4°C for 10 mins followed by centrifugation at 350 x g for 2 min before aspiration of excess buffer. The resin was washed twice with buffer A (1xEQ buffer (pH 6.0) with the addition of 1% DDM and 40mM imidazole), once with buffer B (1xEQ buffer (pH 6.0) with the addition of 1% DDM, 500mM KCl, and 40mM imidazole), and once with buffer C (1xEQ buffer (pH 6.0) with the addition of 0.1% DDM, 500mM KCl, and 40mM imidazole). After the final wash, buffer C was

not aspirated, but instead and loaded onto a Poly-Prep® chromatography column. His-Ste14 was eluted from the cobalt resin with elution buffer (1xEQ buffer (pH 6.0), 0.1% DDM, 1M imidazole, and 2mM AEBSF).

Collected protein was concentrated at 5000 xg in a 30,000 MWCO concentrator at 4°C until the final volume was ~1mL. Concentrated protein was filtered in centrifugal filter units (0.22 µm) GV Durapore® by centrifuging at 12,000 xg for 2 mins at 4°C. Once filtered, the His-Ste14 sample was applied to a Superdex S200 column pre-equilibrated with filtered 1xEQ buffer (50mM Na₂HPO₄, 0.3 M NaCl buffer, pH 6.0) with the addition of 0.05% DDM and 2mM DTT at 4°C and spectra was collected at 280nm. Fractions containing purified protein were pooled, concentrated to a volume under 300µL in a new 30,000 MWCO concentrator, immediately flash frozen, and stored at -80°C. Pure protein concentration was determined using an Amido Black protein assay.

5.2.6 SEC-MALS-SAXS data collection and analysis

SEC-MALS-SAXS experiments were performed at beamline 18-ID (BioCAT) of the advanced photon source (APS) at Argonne National Laboratory using a Pilatus3 1M detector (160 µm x 75 µm) and a 12 KeV X-ray with 1.033 wavelength and ~3.5 m sample-to-detector distance. All protein was purified in a buffer comprised of 50mM Na₂HPO₄, 0.3M NaCl, 2mM DTT, and 0.05% DDM, pH 6 and concentrated to ~5.0 mg/mL at 300µL. To avoid aggregation, samples should never exceed 7.0 mg/mL at 300µL. The purified protein samples were loaded on the SEC column with an exclusion limit of 2.5 MDa (WTC-030S5, Wyatt Technologies), which was followed by a UV detector, a multi-angle light scattering (MALS) detector, a dynamic light scattering (DLS) detector (DAWN Helios II, Wyatt Technologies), a differential refractive index detector (Optilab T-rEX, Wyatt Technologies) and finally the SAXS flow cell. 0.5 s exposures were acquired every 3 s at a flow rate of 0.8mL/min. The exposures either flanking or on either side of the elution peak were averaged to generate a buffer file that was subtracted from the exposures corresponding to selected regions of the elution peak to get the I(q) vs q SAXS curves. Data analysis was performed using the ATSAS package.¹⁷ PRIMUS from the ATSAS program suite was used to perform buffer subtraction, Guinier approximation based forward scattering (I₀) analysis, and radius of gyration (R_g) determination.¹⁸ Scattering curves were further analyzed using GNOM and RAW for the calculation of I₀, R_g, distance distribution P(r), maximum dimension

(D_{\max}).^{18,19} Molecular weights were calculated from the MALS data using the ASTRA software (Wyatt Technologies). Plots were generated from buffer subtracted averaged data and plotted using GraphPad Prism v.7.0.

5.3 Results and Discussion

Size exclusion chromatography (SEC) was used to identify the oligomeric state and homogeneity of His-Ste14. The fractions corresponding to the largest peak (and percentage of His-Ste14) was collected and analyzed with MALS and SAXS (Figure 5.3). The largest peak automatically proceeds to MALS and SAXS without manual manipulation. Additionally, it is important to understand that this is the second time the sample undergoes SEC. SEC must be performed in the initial purification process before SEC-MALS-SAXS experiments are performed at Argonne National Laboratory to remove potential contaminants. In the UV trace, the largest percent of His-Ste14 elutes ~16 mins. Other peaks of lesser heights can be seen at other elution times, yet several of them are excess DDM micelles. DDM micelles can absorb UV light at 280nm.

Next, multi-angle light scattering occurred when laser light was passed through the purified His-Ste14 sample. The oscillating electric field from the light induced an oscillating dipole within the His-Ste14 sample. This oscillating dipole re-radiated the light and the intensity of that light is dependent on the magnitude of the dipole induced by His-Ste14. The MALS detector can measure the intensity of the scattered light of the His-Ste14 sample at specific SEC elution times. The largest peak for light scattering was also detected at ~16mins, aligning with the UV spectra (Figure 5.3A). Light scattering and UV spectra will absorb at the same time if protein is in the sample. With the ATSUS program, the MALS data was used to calculate the molecular weight of His-Ste14 (at ~16mins) to be $67.26\text{kDa} \pm 2.811$.¹⁷ His-Ste14 is predicted to have a molecular weight of ~35kDa and ~70kDa for monomer and dimer, respectively (Figure 5.3A). Therefore, our His-Ste14 samples are most likely in a native dimeric conformation.

Immediately following MALS, SAXS was performed by measuring the intensity and scattering angle of an X-ray after passing through the His-Ste14 sample from the ~16min elution. The raw scattering data of His-Ste14 was used to better understand its structural characteristics (Figure 5.3B). Guinier analysis, $\ln(I)$ (beam intensity) vs. q^2 (scattering angle), at low values of q , was performed. The Guinier plot was consistent with the absence of aggregates and was also used to determine the radius of gyration (R_g) to be 37.6 ± 0.303 (Figure 5.3C, Table 5.1). The linearity

of the relationship between $\ln(I)$ and q^2 indicates that purified His-Ste14 is not aggregated. Pair-distance distributions, $P(r)$, with maximum intramolecular distances (D_{\max}) shown for His-Ste14, indicates an oligomeric state due to the two distinct peaks.¹⁶ The D_{\max} was determined to be 104Å (Figure 5.3D, Table 5.1). The D_{\max} calculated with MALS was consistent with the dimensional maximum vector length in our three-dimensional His-Ste14 dimeric model (~100kDa) utilizing Pymol and GalaxyWEB.²⁰

5.4 Future Directions

To confirm that His-Ste14 functions as a dimer, it would be interesting to create mutations in the enzyme in order to isolate the His-Ste14 monomer. One or more mutations could be made in the GXXXG repeating motif to break the interactions between two His-Ste14 enzymes, followed by analysis with SEC-MALS-SAXS (Appendix A). Alternatively, truncations of TM1 may be interesting to analyze as well. Although it would not be function, we would be able to compare the dimeric SEC-MALS-SAXS data to monomer SEC-MALS-SAXS data. As discussed in Chapter four, His-Ste14 can be incorporated into nanodiscs. Analyzing His-Ste14 with SEC-MALS-SAXS in a more native and stable environment may be an interesting continuation of this research.

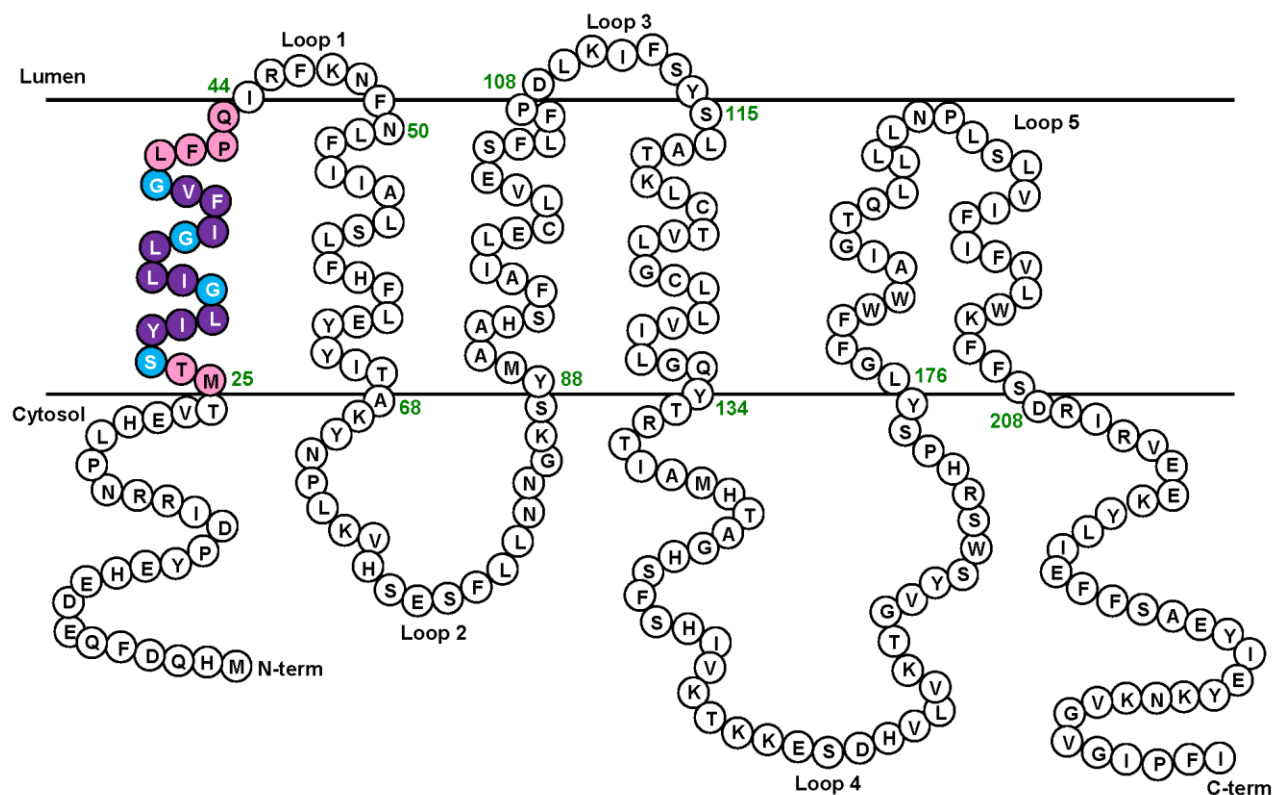


Figure 5.1 Topology of *Saccharomyces cerevisiae* Icmt, Ste14.

Ste14 contains a motif of interest for dimerization, SXXXGXXXGXXXG, consisting of residues 27 to 39. The glycine and serine residues allow for helix-helix interaction between monomer Ste14 protein to form hydrogen bonds without causing steric hindrance. Light blue residues represent the glycine or serine amino acids in the repeating GXXXG motif, while purple residues represent the “X” amino acids. Pink residues indicate the remainder of the TM1 residues in Ste14, residues 25 to 43, utilized in experiments described under Publications.

| | | | | | | | | | | | |
|--------------------|-----|----------|-----------|---------|---------|-------|---------|----------|-----------|-----------|-----|
| O60725 ICMT_HUMAN | 1 | MAGCAARA | PPGSEARLS | LATFLL | GASVLAL | LLT | ---- | RAGLQGR | -TGLALYVA | LNALLLL | 58 |
| P32584 STE14_YEAST | 1 | ----- | ----- | ----- | ----- | ----- | ----- | MHQDFQED | ----- | EHEYD | 16 |
| D6WJ77 ICMT_BEETLE | 1 | ----- | MLS | PAGK | I | S | QSFTGSS | LVFFV | ICMFNHYYG | I | 56 |
| Q8TMG0 Ag_ICMT | | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | |
| Q7PXA7 Ma_ICMT | 1 | ---- | MAIMLCYE | GR | LS | YCFL | GGMAS | LLVFY | GAEAF | LHEES | 59 |
| O60725 ICMT_HUMAN | 59 | YRPPRY | QIAIRACFL | GFVF | CGGTLLS | - | FS | SSASHFG | WMC | SLSLFHYSE | 119 |
| P32584 STE14_YEAST | 17 | RN | -PLHEV | TMT | SYILGI | -- | LLGI | FVGLF | PCIR | FKNFNLFI | 75 |
| D6WJ77 ICMT_BEETLE | 57 | YNEFAFA | IAIRAAFL | GLVLVL | GLYIKL | V | APPNI | QIFGGYMS | VMALFHYSE | EFLAIAI | 118 |
| Q8TMG0 Ag_ICMT | 1 | MNENLWK | ICFI | VMF | - | ----- | ----- | I | WVFRK | VYGT | 37 |
| Q7PXA7 Ma_ICMT | 60 | YNPRDY | QIAVRAAFL | GTVLS | AGI | VVVL | YAEQY | KSGI | YATLMAL | FHYTEY | 121 |
| O60725 ICMT_HUMAN | 20 | SLDSFLLN | HSLEY | TVAAALS | SWLEFT | LEN | I | FWPEL | KQIT | WLS | 171 |
| P32584 STE14_YEAST | 76 | HSESFLLN | NGKSY | MAAHSA | FAILECL | VES | SFLPDL | KIF | SYSLAT | ----- | 133 |
| D6WJ77 ICMT_BEETLE | 19 | STDSFV | INHSPQYT | IAAVS | SWEFFI | ET | YFFPGL | KEI | HMLS | ----- | 170 |
| Q8TMG0 Ag_ICMT | 38 | RPN | ----- | FEK | ----- | SLVFL | -NF | IGMVFL | PLTAV | FS | 89 |
| Q7PXA7 Ma_ICMT | 22 | SPDSF | ILNHS | IHYGL | AAAS | SWIEY | FVEVHY | FEMKT | YKSW | ----- | 173 |
| O60725 ICMT_HUMAN | 172 | CLRKAAMF | TAGS | NFNHVV | ONEKSD | HTLV | T | SGVYA | WFRHPS | SYGAF | 235 |
| P32584 STE14_YEAST | 134 | YTRTI | AMHTAGHS | FSHIV | KTKES | DHVL | VKTGV | YSW | SRHPS | YLGFF | 197 |
| D6WJ77 ICMT_BEETLE | 171 | VLRKTA | ILTAGS | NFNHLV | CEKSSD | HVLV | T | HGVYA | WFRHPS | SYGAF | 234 |
| Q8TMG0 Ag_ICMT | 90 | GLFTK | IHKDLGN | MSAIL | -- | EIKD | GKLVKE | GIYKN | IRHP | MYAHL | 150 |
| Q7PXA7 Ma_ICMT | 174 | SLRKVAMI | TASKN | FSHIV | GERHNE | HELV | T | HGVYG | MRHPS | SYGAF | 237 |
| O60725 ICMT_HUMAN | 236 | YALTWR | -FFRDR | T | EEEE | ISL | I | HFFG | EEYLEY | KKR | 284 |
| P32584 STE14_YEAST | 198 | FIFVLWK | -FFSDR | IRVEEK | YLIEFF | SAEY | I | EYK | NKVGV | GIPFI | 239 |
| D6WJ77 ICMT_BEETLE | 235 | YTLASWM | -FFKER | IYIEES | MLLS | FFG | QYCDY | QQQ | VG | TGIPF | 281 |
| Q8TMG0 Ag_ICMT | 151 | FGIVAW | AILYF | IRVPKE | EEELL | IEEF | GDEY | I | EYMK | TGR | 194 |
| Q7PXA7 Ma_ICMT | 238 | YAIASWK | -FFHDR | ILMEE | ITLLN | FFG | EEYI | Q | QERV | PSGLPY | 286 |

Figure 5.2 Sequence alignment of Icmt orthologs.

Sequence alignment of human Icmt (hIcmt; 284 residues), *Saccharomyces cerevisiae* Icmt (Ste14; 239 residues), *Tribolium castaneum* Icmt (beetle, 281 residues), the Icmt found in *Anopheles gambiae* (Ag-Icmt; 283 residues), and the prokaryotic enzyme from *Methanosarcina acetivorans* (Ma-Icmt; 194 residues). UniProt accession numbers for the Icmt ortholog sequences are as follows: O60725, P32584, D6WJ77, Q8TMG0, and Q7PXA7, respectively. Alignment was made using Clustal Omega and Jalview 2.10.5.^{21,22} Colored boxes denote residues with similar properties: glycine is orange, acidic amino acids are purple, proline is yellow, nonpolar amino acids are light blue, histine and tyrosine are turquoise, basic amino acids are red, and remaining polar, neutrally charged amino acids are green. The red dashed box surrounds the GXXXG motif in Ste14.

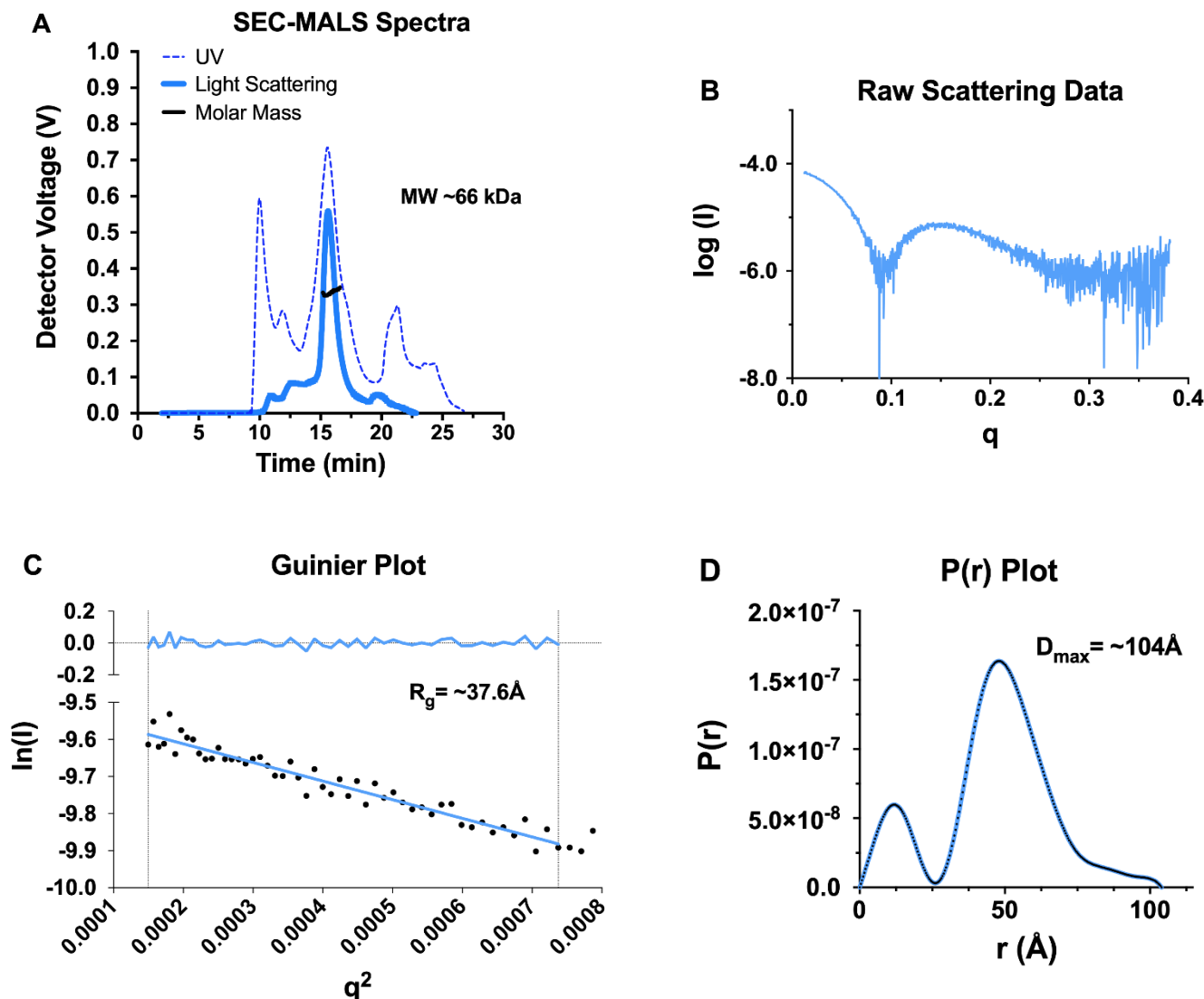


Figure 5.3 Characterization of Ste14 via SEC-MALS-SAXS confirms dimerization state. (A) SEC-MALS indicated one molecular weight for purified Ste14. Extrapolated averaged molecular weights estimate a molecular mass of $67.26 \text{ kDa} \pm 2.811$ and confirms the homodimeric nature of Ste14 ($\sim 70 \text{ kDa}$). (B) Log beam intensity (I) vs scattering angle magnitude (q), after buffer subtraction. (C) Guinier plot of low q values, $\ln(I)$ vs. q^2 , with radius of gyration (R_g) and y-intercept (I_0). Fitting of the blue line to the data is represented by residuals at the top of the plot. R_g was determined to be 37.6 ± 0.303 . (D) $P(r)$ plot with maximum interparticle distance (D_{\max}) of $\sim 104 \text{ \AA}$.

Table 5.1 SAXS structural parameters of His-Ste14

| Pure Ste14 | |
|---------------------------------|---|
| Guinier analysis | |
| $I(0)$ (cm ⁻¹) | $7.32\text{e-}10^{-5} \pm 3.97\text{e-}10^{-7}$ |
| R_g (Å) | 37.6 ± 0.303 |
| q min (Å ⁻¹) | 0.0122 |
| q range (Å ⁻¹) | 0.0122 - 0.0278 |
| r^2 (linear fit) | 0.9284 |
| P(r) analysis | |
| $I(0)$ (cm ⁻¹) | $7.94\text{e-}10^{-5} \pm 6.96\text{e-}10^{-7}$ |
| R_g (Å) | 52.3 |
| D_{max} (Å) | 104 |
| Porod Volume (Å ⁻³) | 52,100 |
| q range (Å ⁻¹) | 0.0122 - 0.0313 |

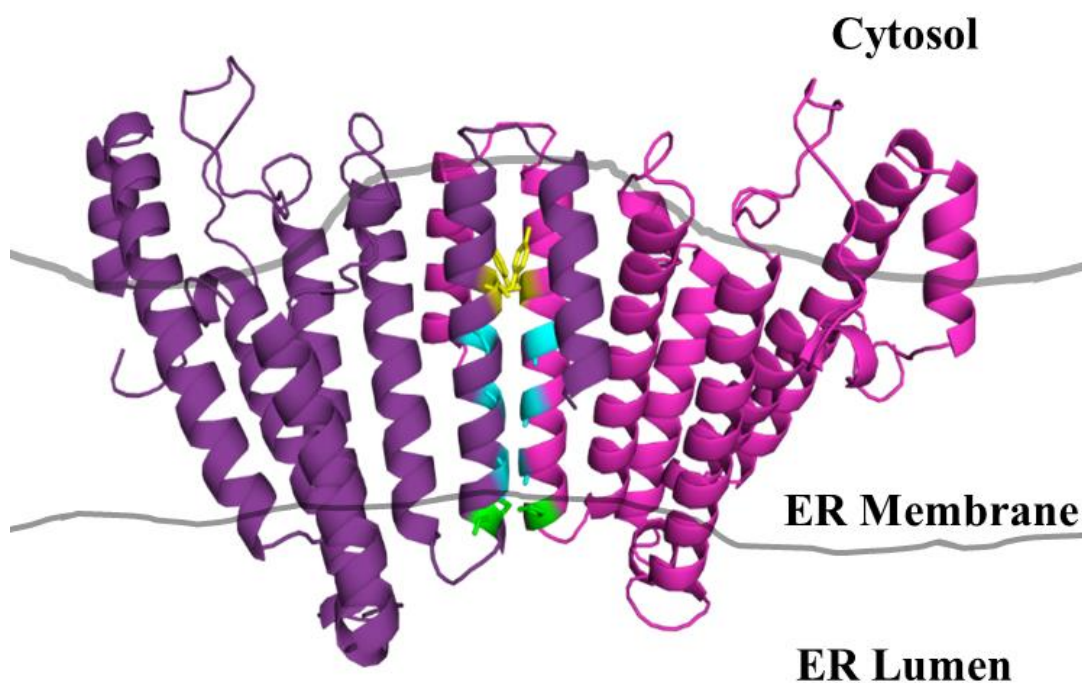


Figure 5.4 Predicted structural model of Ste14 dimerization.

TM1 residues of two Ste14 molecules are shown in pink and purple. The glycine residues (cyan) comprising the repeating GXXXG motif in TM1 create a groove for the close packing of another transmembrane helix. Tyrosine 28 (yellow) and proline 42 (green) may stabilize the dimerization interaction by contacting and holding the other helix in place. A 3D homology alignment was generated with I-TASSER using the beetle Icmt (5V7P) as a structural template for Ste14.²³⁻²⁵ PyMol was used to generate the model. Point of contact were determined with GalaxyWEB.²⁰

5.5 References

- 1 Griggs, A. M., Hahne, K. & Hrycyna, C. A. Functional Oligomerization of the *Saccharomyces cerevisiae* Isoprenylcysteine Carboxyl Methyltransferase, Ste14p. *Journal of Biological Chemistry* **285**, 13380-13387, doi:10.1074/jbc.M109.061366 (2010).
- 2 Gerber, D. & Shai, Y. In vivo detection of hetero-association of glycophorin-A and its mutants within the membrane. *Journal of Biological Chemistry* **276**, 31229-31232, doi:10.1074/jbc.M101889200 (2001).
- 3 MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. A transmembrane helix dimer: Structure and implications. *Science* **276**, 131-133, doi:10.1126/science.276.5309.131 (1997).
- 4 Polgar, O. *et al.* Mutational analysis of ABCG2: Role of the GXXXG motif. *Biochemistry* **43**, 9448-9456, doi:10.1021/bi0497953 (2004).
- 5 Arselin, G. *et al.* The GxxxG motif of the transmembrane domain of subunit e is involved in the dimerization/oligomerization of the yeast ATP synthase complex in the mitochondrial membrane. *European Journal of Biochemistry* **270**, 1875-1884, doi:10.1046/j.1432-1033.2003.03557.x (2003).
- 6 Whittington, D. A. *et al.* Crystal structure of the dimeric extracellular domain of human carbonic anhydrase XII, a bitopic membrane protein overexpressed in certain cancer tumor cells. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 9545-9550, doi:10.1073/pnas.161301298 (2001).
- 7 McClain, M. S. *et al.* Essential role of a GXXXG motif for membrane channel formation by *Helicobacter pylori* vacuolating toxin. *Journal of Biological Chemistry* **278**, 12101-12108, doi:10.1074/jbc.M212595200 (2003).
- 8 Overton, M. C., Chinault, S. L. & Blumer, K. J. Oligomerization, biogenesis, and signaling is promoted by a glycophorin A-like dimerization motif in transmembrane domain 1 of a yeast G protein-coupled receptor. *Journal of Biological Chemistry* **278**, 49369-49377, doi:10.1074/jbc.M308654200 (2003).
- 9 Bhatia, A., Schafer, H. J. & Hrycyna, C. A. Oligomerization of the human ABC transporter ABCG2: Evaluation of the native protein and chimeric dimers. *Biochemistry* **44**, 10893-10904, doi:10.1021/bi0503807 (2005).
- 10 Senes, A., Gerstein, M. & Engelman, D. M. Statistical analysis of amino acid patterns in transmembrane helices: The GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *Journal of Molecular Biology* **296**, 921-936, doi:10.1006/jmbi.1999.3488 (2000).
- 11 MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. A transmembrane helix dimer: structure and implications. *Science* **276**, 131-133 (1997).

- 12 Yang, J. *et al.* Mechanism of Isoprenylcysteine Carboxyl Methylation from the Crystal Structure of the Integral Membrane Methyltransferase ICMT. *Molecular Cell* **44**, 997-1004, doi:10.1016/j.molcel.2011.10.020 (2011).
- 13 Diver, M. M., Pedi, L., Koide, A., Koide, S. & Long, S. B. Atomic structure of the eukaryotic intramembrane RAS methyltransferase ICMT. *Nature* **553**, 526-+, doi:10.1038/nature25439 (2018).
- 14 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254 (1976).
- 15 Anderson, J. L., Frase, H., Michaelis, S. & Hrycyna, C. A. Purification, functional reconstitution, and characterization of the *Saccharomyces cerevisiae* isoprenylcysteine carboxylmethyltransferase Ste14p. *Journal of Biological Chemistry* **280**, 7336-7345, doi:10.1074/jbc.M410292200 (2005).
- 16 Svergun, D. I. & Koch, M. H. Small-angle scattering studies of biological macromolecules in solution *Rep. Prog. Phys.* **66**, 1735-82 (2003).
- 17 Franke, D., Petoukhov, M., Konarev, P., Panjkovich, A., Tuukkanen, A., Mertens, H., Kikhney, A., Hajizadeh, N., Franklin, J. & Jeffries, C. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212-1225 (2017).
- 18 Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., & Svergun, D. I. PRIMUS: a Windows PC -based system for small -angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277 -1282 (2003).
- 19 Svergun D., Determination of the regularization parameter in indirect-transform methods using perceptual criteria *J. Appl. Crystallogr.* **25**, 495-503 (2003).
- 20 Shin, W.H., Lee, G. R., Heo, L., Lee, H., & Seok, C., Prediction of Protein Structure and Interaction by GALAXY Protein Modeling Programs. *Bio Design* **2**, 1-11 (2004).
- 21 Sievers F., Wilm A., Dineen D., Gibson T. J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J. D. & Higgins D. G., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
- 22 Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J., Jalview Version 2 - a multiple sequence alignment editor and analysis workbench Bioinformatics doi: 10.1093/bioinformatics/btp033 (2009).
- 23 Ger Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y., The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods.* **12**, 7-8 (2015).

- 24 Roy, A., Kucukural, A., & Zhang, Y., I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725-738 (2010).
- 25 Zhang, Y., I-TASSER server for protein 3D structure prediction. *BMC Bioinform.* **9**, 40 (2008).

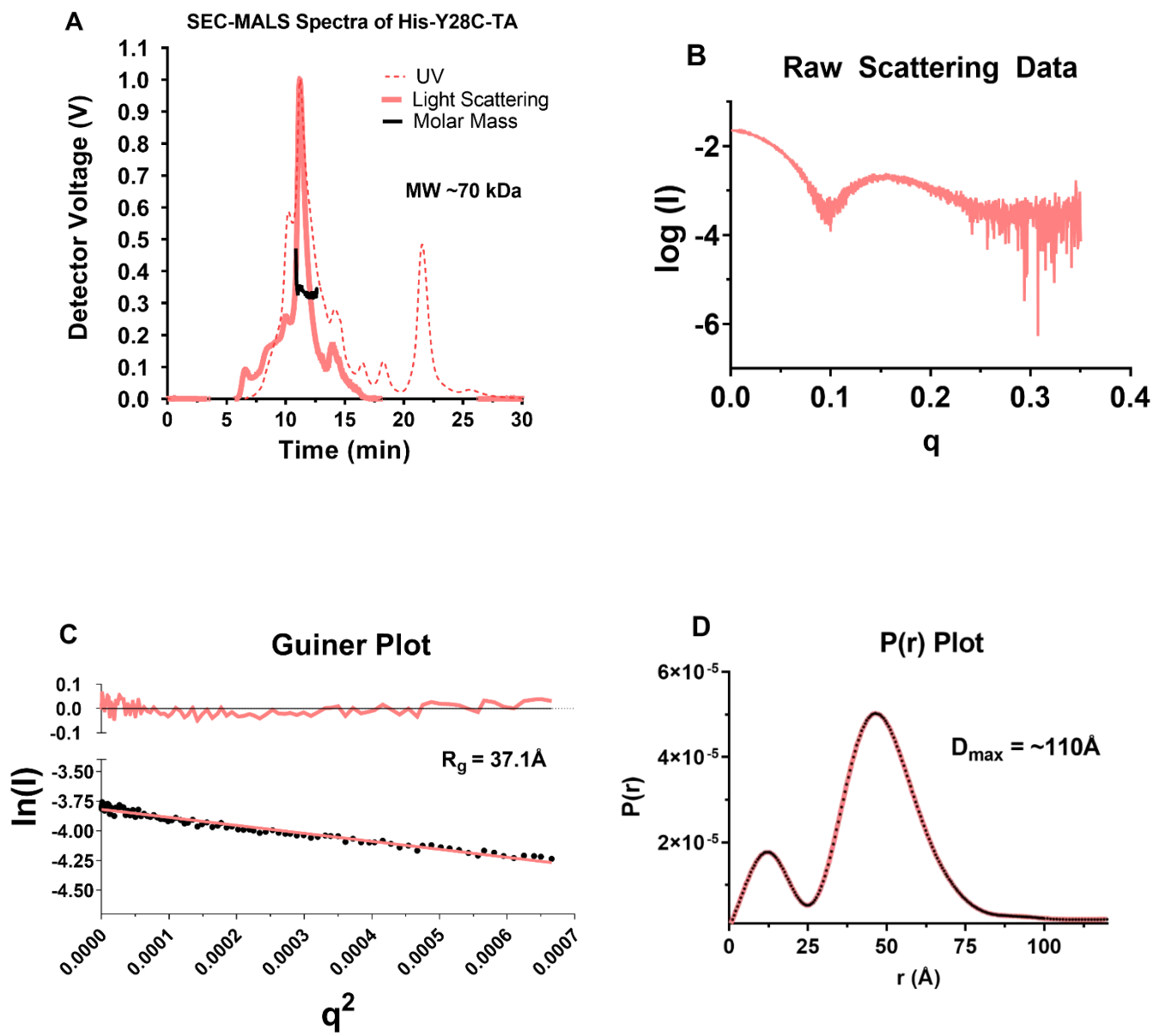
APPENDIX A

Characterization of His-Ste14 mutant Y28C-TA via SEC-MALS-SAXS confirms dimerization state. SEC-MALS-SAXS experiments were performed at beamline 18-ID (BioCAT) of the advanced photon source (APS) at Argonne National Laboratory using a Pilatus3 1M detector (160 μm x 75 μm) and a 12 KeV X-ray with 1.033 wavelength and ~3.5 m sample-to-detector distance. All protein was purified in a buffer comprised of 50mM Na_2HPO_4 , 0.3M NaCl, 2mM dithiothreitol (DTT), and 0.05% DDM, pH 6. The purified protein samples were loaded on the SEC column with an exclusion limit of 2.5 MDa (WTC-030S5, Wyatt Technologies), which was followed by a UV detector, a multi-angle light scattering (MALS) detector, a dynamic light scattering (DLS) detector (DAWN Helios II, Wyatt Technologies), a differential refractive index detector (Optilab T-rEX, Wyatt Technologies) and finally the SAXS flow cell. 0.5 s exposures were acquired every 3 s at a flow rate of 0.8mL/min. The exposures either flanking or on either side of the elution peak were averaged to generate a buffer file that was subtracted from the exposures corresponding to selected regions of the elution peak to get the $I(q)$ vs q SAXS curves. Data analysis was performed using the ATSAS package (Franke, D., Petoukhov, M., Konarev, P., Panjkovich, A., Tuukkanen, A., Mertens, H., Kikhney, A., Hajizadeh, N., Franklin, J., Jeffries, C. (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212-1225).

PRIMUS from the ATSAS program suite was used to perform buffer subtraction, Guinier approximation based forward scattering (I_0) analysis, and radius of gyration (R_g) determination (Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC -based system for small -angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277 -1282). Scattering curves were further analyzed using GNOM (51) for the calculation of I_0 , R_g , distance distribution $P(r)$, maximum dimension (D_{max}). Molecular weights were calculated from the MALS data using the ASTRA software (Wyatt Technologies). Plots were generated from buffer subtracted averaged data and plotted using GraphPad Prism v.7.0.

SEC-MALS indicated one molecular weight for purified His-Y28C-TA. Extrapolated averaged molecular weights estimate a molecular mass of 70.0 kDa and confirms the homodimeric nature of this mutant. Wild-type His-Ste14 is ~70kDa as a homodimer, thus the mutation of tyrosine to cysteine does not break the higher order oligomer into a monomer. Log beam initially

(I) vs scattering angle magnitude (q) was plotted after buffer subtraction, which was termed “Raw Scattering Data”. Then, the Guinier plot was created at low q values, $\ln(I)$ vs. q^2 , with radius of gyration (R_g) and y-intercept (I_0). Fitting of the pink line to the data is represented by residuals at the top of the plot. The R_g was determined to be 37.1 ± 0.262 . The linearity of the relationship between $\ln(I)$ and q^2 indicates that purified Ste14 is not aggregated. Pair-distance distributions, $P(r)$, with maximum intramolecular distances (D_{max}) shown for Ste14, indicates a dimeric state due to the two distinct peaks. $P(r)$ plot with maximum interparticle distance (D_{max}) of $\sim 104 \text{ \AA}$.



| Pure His-Y28C-TA | |
|------------------------------|-----------------------------------|
| Guinier analysis | |
| $I(0)$ (cm ⁻¹) | $0.0229 \pm 8.429\text{e}10^{-5}$ |
| R_g (Å) | 37.0649 ± 0.2621 |
| q min (Å ⁻¹) | 0.00653 |
| q range (Å ⁻¹) | 0.00653 - 0.03235 |
| r^2 (linear fit) | 0.9780 |
| P(r) analysis | |
| $I(0)$ (cm ⁻¹) | 0.0233 |
| R_g (Å) | 40.7337 ± 0.6930 |
| D_{max} (Å) | ~110 |
| q range (Å ⁻¹) | 0.0043 - 0.03522 |

APPENDIX B

A preliminary X-ray crystallography experiment was performed with the following protocol. Briefly, the crude membranes previously isolated were solubilized in a solution of 0.02M imidazole, 1% (w/v) *n*-dodecyl- β -D-maltopyranoside (DDM), and 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (20mM HEPES, 50mM NaCl, pH 7.2) with the addition of 2mM 4-(2-Aminoethyl)benzensulfonyl fluoride hydrochloride (AEBSF). For the solubilization reaction, crude membranes were at a final concentration of 5 mg/mL. The reaction was rocked at 4°C for 1 hour. The resulting mixture was centrifuged at 100,000 xg for 45 mins. The soluble fraction was incubated for 1 hour at 4°C with Talon® cobalt metal affinity beads equilibrated in HEPES buffer with the addition of 2mM AEBSF. After the protein was incorporated with the resin, the solution was centrifuged at 350 xg for 2 mins at 4°C and the remaining buffer was removed via aspiration. Next, the beads underwent four wash steps: buffer was added to the resin and was allowed to rock at 4°C for 10 mins followed by centrifugation at 350 x g for 2 min before aspiration of excess buffer. The resin was washed twice with buffer A (HEPES buffer with the addition of 1% DDM (w/v) and 40mM imidazole), once with buffer B (HEPES buffer with the addition of 1% DDM (w/v), 500mM KCl, and 40mM imidazole), and once with buffer C (HEPES buffer with the addition of 0.1% DDM (w/v), 500mM KCl, and 40mM imidazole). The solution underwent centrifugation at 350 x g for 2 min and loaded onto a Poly-Prep® chromatography column. Then, the His-Ste14 was eluted from the cobalt resin with elution buffer (0.1% DDM, 1M imidazole, and 2mM AEBSF). Collected protein was concentrated at 5000 xg in a 30,000 molecular weight cutoff (MWCO) concentrator at 4°C until the volume was under 1mL. Concentrated protein was filtered in centrifugal filter units (0.22 μ m) GV Durapore® by centrifuging at 12,000 xg for 2 mins at 4°C. Once filtered, the His-Ste14 was applied to a Superdex S200 column pre-equilibrated with HEPES buffer (with the addition of 0.05% DDM and 2mM Dithiothreitol (DTT)) at 4°C and spectra was collected at 280nm. Fractions containing purified protein were pooled, concentrated in a new 30,000 MWCO concentrator until the volume was under 200 μ L. S-adenosyl-L-homocysteine (SAH) was added to the concentrated protein at a 1:2 ratio.

Broad crystallization screening was performed using hanging drop method on a Mosquito LCP crystallization robot (TTP Labtech) with commercially available crystallization screens.

100ng were spotted for each well of the tray under high humidity and stored at 4°C for crystallization to occur. The trays selected for this screening were: Index, MCSG1, MCSG2, MCSG3, MCSG4, MemFac, MemGold, MemSt/Sys, and TOP96. Although there was crystal development in all of the trays, no protein was observed within crystals. This could be due to a higher salt concentration and would be interesting to repeat at a lower concentration. Additionally, these experiments were not performed with His-Ste14 incorporated into nanodiscs. The next logical experiment would be following the above protocol with His-Ste14 within the nanodisc complex.

VITA

Anna Christine Ratliff was born on April 22nd, 1992 in Elmhurst, Illinois to her parents Chris and Elaine Ratliff. She grew up with her older sister, Kate Ratliff, where they loved being outdoors, going on family road trips, and learning together. From 2006 to 2010, Anna attended York Community High School where she became an engaged member of the community through academics, athletics and various other extracurriculars. She played soccer and French horn in the band, as well as partaking in the National Honors Society and Key Club. Anna also took a wide range of classes from AP French to AP music theory; however, it was in AP Biology where she found a passion and desire to study in the sciences.

In 2010, Anna matriculated at Lawrence University in Appleton, Wisconsin, where she obtained her Bachelor of Arts in Chemistry and Biochemistry, *cum laude*. While at Lawrence, she ran Cross-country and Track & Field all four years and became a captain of both teams her senior year. Additionally, she was active in the community by facilitating a mentorship program with student athletes and elementary school students and organizing Special Olympic events at the University. She also met her fiancé, Graham Jones, at Lawrence University in 2011. Anna's passion of science continued to blossom at Lawrence through her biochemistry course work and during her research as a summer intern in biophysical research with Dr. Doug Martin of the Physics Department.

In 2014, Anna enrolled at Purdue University and joined Dr. Christine Hrycyna's research lab to pursue her Doctoral degree in Chemistry. She studied various methyltransferases, membrane proteins, and gained valuable experience with a variety of biophysical techniques. Anna graduated from Purdue University in December of 2019 with a Ph.D. in Chemistry. In 2020, she will be beginning a postdoctoral research position at the National Institute of Health/ National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) in Bethesda, Maryland under Dr. Susan Buchanan. Her research will entail studying membrane proteins in gram negative bacteria using various structural techniques. After her postdoctoral position, Anna hopes to teach biochemistry at a small, primarily undergraduate university.

PUBLICATIONS

Dimerization motif of yeast ICMT

Formatted for submission to Journal of Biological Chemistry

Characterization of the GXXXG dimerization motif of the yeast isoprenylcysteine carboxyl methyltransferase

Anna C. Ratliff¹, Sahej Bains^{2#}, Srinivas Chakravarthy³, Amy L. Funk^{1#}, Amy Griggs¹⁺, Christine A Hrycyna^{1*}

From the ¹Department of Chemistry, Purdue University, West Lafayette IN 47907, ²Department of Biological Sciences, Purdue University, West Lafayette IN 47907, ³Biophysics Collaborative Access Team, Illinois Institute of Technology, Sector 18ID, Advanced Photon Source, Argonne National Laboratory, Lemont IL 60439

Running title: *Dimerization motif of yeast ICMT*

^{1#}Present address: National Cancer Institute ATRF, 8560 Progress Drive, Frederick MD 21701

^{1#}Present address: Cook Biotech, Inc., 1425 Innovation Place, West Lafayette, IN 47906

^{2#}Present address: Mayo Clinic, 200 1st St SW, Rochester MN 55905

*To whom correspondence should be addressed: Christine A. Hrycyna: Department of Chemistry and the Purdue Cancer Center, Purdue University, West Lafayette, Indiana 47907; hrycyna@purdue.edu; Tel. (765) 494-7322

Keywords: Ste14, Isoprenylcysteine carboxyl methyltransferase (ICMT), dimerization, methyltransferase, K-Ras

ABSTRACT

Isoprenylcysteine carboxyl methyltransferase, IcmT, is an integral membrane protein localized to the endoplasmic reticulum and is responsible for the post-translational α -carboxyl methylesterification of the C-terminus in CaaX proteins. This post-translational modification is important for localization of the CaaX protein the plasma membrane, where upon activation, can signal cellular growth and proliferation. To date, IcmT is the only integral membrane protein methyltransferase, thus, little is known about the structure and function. We previously showed that the homolog of IcmT from *Saccharomyces cerevisiae*, Ste14, is functional as a dimer or oligomer. In this study, we provided further evidence to support that Ste14 is a dimer by utilizing

multi-angle light scattering (MALS) and small angle X-ray scattering (SAXS). Additionally, we discovered a set of possible structural residues for dimerization within the first transmembrane (TM) domain of Ste14, which contains a putative dimerization motif, G₃₁XXXG₃₅XXXG₃₉. Using cysteine-scanning mutagenesis, we characterized the TM1 cysteine mutants for their effects on protein expression, activity, and stability. Residues S27, Y28, L30, G31, G35, and G39 are critical for activity of the enzyme. Trypsin digestion revealed that these same residues are important for maintaining the overall structure of the wild-type protein. With the addition of sulfhydryl-specific cross-linkers and immunoblot analysis, we have determined that residues M25, T26, Y28, F41, P42, and Q43 form strong

dimers. Together, these biochemical and biophysical data suggest that the amino acids comprising and surrounding the GXXXG motif play a pivotal role in dimer formation.

INTRODUCTION

Isoprenylcysteine carboxyl methyltransferase (Icmt) is an enzyme localized in the endoplasmic reticulum membrane that is responsible for the α -carboxyl methyl esterification of CaaX proteins (1-5). CaaX proteins are involved in the regulation of several cellular processes including cell growth and carcinogenesis. These proteins contain a C-terminal "CaaX motif" where "C" is a cysteine, "a" is an aliphatic amino acid, and "X" is any amino acid (6-15). This CaaX motif signals these proteins to undergo three distinct post-translational modifications. First, the cysteine residue is isoprenylated by either farnesyltransferase (FTase) or geranylgeranyltransferase I (GGTase I), then the -aaX sequence is proteolyzed by Ras converting enzyme 1 (Rce1), and lastly, the isoprenylated cysteine undergoes methyl esterification by Icmt (4,16). Over 300 CaaX proteins have been identified including Ras proteins, Rheb proteins, nuclear lamins, and the α -factor yeast mating peptide (17). The Ras proteins are of particular interest to us because mutations in K-Ras are associated with 90% of pancreatic cancers and 30% of all cancers (18). The carboxyl methylation of Ras by Icmt was found to be important for cell signaling and the proper localization of Ras proteins within cells (4,19-23). Interestingly, Icmt is the only identified enzyme to date that is responsible for this C-terminal methyl esterification of the cysteine residue of CaaX proteins. Thus, elucidation of the molecular details and mechanism of Icmt will allow for the development of therapeutics that effectively inhibit Icmt causing the mislocalization of Ras, and in turn, aid in the treatment of K-Ras driven

cancers (23-28). To better understand the structural nature of Icmt, we used the *S. cerevisiae* homolog of Icmt, Ste14, for our biochemical and structural studies because the human enzyme has yet to be purified while remaining functionally active (29).

Ste14 is a 26-kDa integral membrane protein that contains 239 amino acids and has six transmembrane segments (3). It is a functional homolog of human Icmt and is 41% identical and 63% similar to the human protein (29). Ste14 possesses a tandem S₂₇XXXG₃₁XXXG₃₅XXXG₃₉ motif located in transmembrane domain one (TM1). The GXXXG motif has been shown to be important in the dimerization of several membrane proteins including glycophorin A (GpA) (30, 31), ATP binding cassette subfamily G member 2 (ABCG2) (32, 33), yeast ATP synthase (34), human carbonic anhydrase (CA) (35), *Helicobacter pylori* vacuolating toxin (36), and yeast α -factor receptor (37). The spacing of the glycine residues allows them to be positioned on the same face of the helix and provide a flat surface to permit the close packing of TM1 of two Ste14 proteins. Additionally, it has been shown that the residues surrounding each glycine, such as serine, threonine, and tyrosine, are also important for dimerization due to the formation of potential intramolecular hydrogen bonds. Moreover, leucine, valine, and isoleucine residues can form van der Waals interactions which can further stabilize the protein dimer (31).

The three-dimensional characterization of Ste14 has been challenging, due to the difficulty associated with crystallizing an integral membrane protein. Currently, two Icmt crystal structure models exist: an Icmt ortholog from the prokaryote *Methanosarcina acetivorans* (Ma-Icmt) and *Tribolium castaneum* Icmt (beetle-Icmt) (38, 39). The Ma-Icmt model contains a similar C-terminus as Ste14 and other Icmt homologs, but does not have any

sequence identify in the N-terminus. We could not accurately use this model because it has less helices than Ste14 and does not include the proposed dimerization site. The beetle Icmt model has 58% identity with the active site region of human Icmt and both proteins contain eight transmembrane domains (39). Long *et al.* demonstrated that beetle Icmt is monomeric due to a GXXXG-like helical packing motif that is predicted to stabilize the two additional N-terminal transmembrane domains that do not exist in Ste14 (39). Sequence alignment analysis showed that the dimerization motif can be seen across the eukaryotic domain and not in the archaea, *Ma-Icmt* (Figure S1). Here, we provide further evidence to support our previous study that Ste14 exists as a dimer or higher order oligomer (40).

We provided evidence of the oligomeric state of Ste14 with a combination of size-exclusion chromatography, multi-angle light scattering and small-angle X-ray scattering (SEC-MALS-SAXS). These biophysical experiments confirmed the molecular weight of dimer formation and overall dimensions of Ste14. We also assessed the effects of the amino acid residues comprising the GXXXG motif on protein dimerization, expression, activity, and folding. In this study, we utilized varying-length cross-linkers to investigate the ability of each amino acid in TM1 to form a dimer. We generated wild-type (WT) and cysteine-less, triple alanine mutant, (TA) Ste14 and demonstrated their ability to form homodimers following the addition of glutaraldehyde. We further utilized cysteine scanning mutagenesis and the addition of sulfhydryl-specific crosslinkers, *Bis*-maleimidoethane (BMOE) and *Bis*-maleimidoethane (BMH) to determine which amino acids are in close enough proximity to aid in dimerization. Cross-linking with the sulfhydryl reactive 13 Å cross-linker, BMH, revealed that cysteine

mutants on that same side of the alpha helix in TM1 formed strong dimers as observed in immunoblot analysis. Similar results were obtained with the shorter 8 Å cross-linker, BMOE. Together, these results indicate that the residues on the same side of the alpha helix contribute to the dimer formation of Ste14 as shown in the helical wheel projections. These results were confirmed with tryptic digestion experiments followed by immunoblot analysis and methyltransferase activity assays.

RESULTS

Purification of wild-type Ste14 via SEC-MALS-SAXS confirms homodimeric state

Size exclusion chromatography (SEC) was used to identify the oligomeric states of Ste14 and homogeneity of Ste14. The fractions corresponding to the largest peak (and percentage of Ste14) was collected and analyzed with MALS and SAXS (Figure 1). MALS was used to calculate the molecular weight of Ste14 to be $67.26\text{kDa} \pm 2.811$; Ste14 is predicted to have a molecular weight of $\sim 35\text{kDa}$ and $\sim 70\text{kDa}$ for monomer and dimer, respectively (Figure 1A).

Additionally, the raw scattering data of Ste14 was used to better understand its structural characteristics. Guinier analysis, $\ln(I)$ (beam intensity) vs. q^2 (scattering angle), at low values of q , was performed. The Guinier plot was consistent with the absence of aggregates and was also used to determine the radius of gyration (R_g) to be 37.6 ± 0.303 (Figure 1C, Table S1). The linearity of the relationship between $\ln(I)$ and q^2 indicates that purified Ste14 is not aggregated.

Pair-distance distributions, $P(r)$, with maximum intramolecular distances (D_{max}) shown for Ste14, indicates a dimeric state due to the two distinct peaks. The D_{max} was determined to be 104\AA (Figure 1D, Table S1).

Crude membrane and purified preparations of wild-type and cys-less Ste14 form a homodimer upon the addition of glutaraldehyde

We have successfully constructed a biologically active cysteine-less (TA) Ste14 mutant that is expressed at similar levels to WT protein (data not shown). In our cysteine-less protein, the three native cysteine residues of Ste14 at positions 99, 121, and 126 were replaced with alanine residues using site-directed mutagenesis. All of our constructs contain ten histidine residues and 3 myc epitopes at the N-terminus which facilitate the purification and identification of Ste14 by immunoblot analysis (His₁₀myc₃N-Ste14). To assess whether TA Ste14 forms a homodimer, we analyzed its migration during gel electrophoresis after the addition of 0.2% (wt/v) glutaraldehyde, a non-specific lysine crosslinker. The expected band size of monomeric Ste14 is ~35kDa and the dimer is anticipated to migrate at ~70kDa. Immunoblot analysis with α -myc showed that crude membrane preparations of TA Ste14 formed a dimer (Figure 2A). To further determine whether TA forms a homodimer, we cross-linked both pure WT and pure TA Ste14 with 0.2% (wt/v) glutaraldehyde. Immunoblot analysis confirmed that both were able to form a dimer or higher order oligomers (Figure 2B). Additionally, both purified WT Ste14 and purified TA Ste14 formed dimers without the addition of glutaraldehyde, suggesting that a stronger dimer is formed with pure protein. Because the TA mutant was able to form a homodimer, we next generated cysteine mutants to determine which specific residues are important for dimerization. We utilized crude membranes for our assays to mimic the state of Ste14 in its native environment.

Cys-less and TM1 Ste14 mutants have differential expression and varied in vitro activities

Again, TA Ste14 permitted the use of cysteine scanning mutagenesis to characterize each residue in TM1 of Ste14 for its ability to form a homodimer and its effect on expression, activity, and structure (Figure S2). All TM1 TA Ste14 mutants, except for G35C, expressed at a similar level as WT (Figure 3A).

To determine if the mutated TM1 residues allowed Ste14 to retain its methyltransferase activity, each mutant was analyzed by the *in vitro* vapor diffusion methyltransferase assay. Crude membrane preparations lacking Ste14 were used as negative controls (Ste14). G31, G35, and G39 comprise the proposed tandem dimerization motif. Although G31 and G35 are highly conserved throughout several species of methyltransferases, G39 is not (Figure S1). However, all three residues are important for Icmt methyltransferase activity. Mutation of any of the three glycine residues to cysteine resulted in a complete and significant loss of methyltransferase activity (Figure 3B). To determine whether the loss in activity could be attributed to a bulkier cysteine side chain, we mutated G31, G35, and G39 to alanine. Interestingly, G31A and G35A were unable to regain activity, but G39A restored 35% methyltransferase activity as compared to WT (data not shown).

Glycine residues comprising the GXXXG dimerization motif are not the only residues proposed to be important for dimerization. The amino acids adjacent to the dimerization motif are thought to provide stabilizing hydrogen bonds as well as van der Waals interactions (31, 41-45). The specific activities of these TM1 mutants were examined, and while other mutants retained more than 40% activity, protein with mutations in residues 25-31 had considerably lower activities, except I29C (Figure 3B).

M25C lost 68% of its activity while T26C lost 73%. Both S27C and L30C lost 97% of their activities. Moreover, the non-conserved Y28C lost 100% of its methyltransferase activity. The Y28C TM1 mutant is the only other residue besides G31C, G35C, and G39C that lost all measurable methyltransferase activity (Figure 3B).

Partial trypsin digestion of Ste14 TM1 mutants revealed differentially folded proteins

The TM1 mutants were digested with trypsin to determine if they properly folded after the introduction of their respective mutations compared to WT. TA exhibited the same cleavage pattern as WT. Ste14 possesses several trypsin cleavage sites within the loop regions of the protein including R16, R17, K153, K154, K155, K163 and R171. R16 and R17 are located at the N-terminus of Ste14 and the start of TM1, thus can be exposed to trypsin. The other cleavage sites are located within the fourth cytosolic loop where SAM binding is proposed to occur (38,39). WT, TA and all of the TM1 mutants were subjected to proteolytic cleavage by trypsin and analyzed via immunoblot (Figure 4). The TM1 mutants with cleavage patterns different from WT and TA Ste14 were S27C, Y28C, L30C, G31A, G31C, G35A, G35C, and G39C (Figure 4). S27C, G31A, G31C, G35A, and G35C were not cleaved by trypsin to yield the lower molecular weight ~25kDa band as seen utilizing the α -myc antibody. Y28C, L30C, and G39C yielded a small amount of the ~25kDa band. However, the G39A mutant showed a cleavage pattern similar to WT Ste14, suggesting that it was properly folded. Notably, we see that trypsin cleaved WT Ste14 at the C-terminus, presumably at one of the five sites in the fourth cytosolic loop because the lower ~25kDa band is not recognized by the α -Ste14 antibody which recognizes the last 42 amino acids of Ste14

(data not shown). These data suggest that G31A, G31C, and G35A mutants are misfolded and the trypsin cleavage site in the fourth cytosolic loop is masked. As expected, the G39A mutant reveals the same cleavage pattern as WT Ste14, as it has 35% WT activity. G39 may play a lesser role in protein folding, because mutation this residue to alanine does recover some methyltransferase activity and permits protein folding comparable to WT. However, it is clear that all glycine residues in the tandem binding motif, as well as S27, Y28, and L30, are critical for both protein folding and methyltransferase activity.

Cysteine crosslinking with bis-maleimidoethane (BMH) or bis-maleimidoethane (BMOE) identifies residues with the ability to aid in Ste14 dimerization

To detect the residues that interact during dimerization, we used cysteine scanning mutagenesis and the bi-functional sulfhydryl reactive cross-linkers, BMH and BMOE. TM1 mutants were crosslinked with either BMH or BMOE and then subjected to SDS-PAGE and immunoblot analysis. Active TA Ste14, which does not contain cysteine residues for crosslinking, and P42C, which was discovered earlier to form a very strong homodimer, were used as negative and positive controls for cross-linking, respectively.

First, all TM1 Ste14 mutants were subjected to cross-linking with BMH, which has a spacer arm of 13 Å. M25C, T26C, Y28C, F41C, and P42C all formed homodimers (Figure 5). Q43C was the only mutant that formed a homodimer without the addition of BMH. S27C, L30C, G31C, G35C, G39C, and L40C also formed homodimer, but at lower expression level than the previously mentioned mutants. I29C, I32C, L33C, L34C, I36C, F37C, and V38C

were not cross-linked under the conditions tested.

To further investigate the dimer formation of TM1 residues, a shorter 8 Å cross-linker, BMOE, was used. M25C, T26C, S27C, Y28C, F41C, P42C and Q43C were able to form dimers. (Figure 6). L30C, G31C, G35C, V38C, G39C, and L40C displayed weak cross-linking with BMOE. No dimer formation was observed with I29C, I32C, L33C, L34C, I36C, and F37C. Based on these results, we have determined more specifically that the residues that interact to form the dimerization interface in TM1 are Y28, G39, L40, P42 and Q43. The other two glycine mutants, G31C and G35C, are misfolded which makes it difficult to determine if they would form strong dimers if folded properly, but they are critical for the structural stability of Ste14 (Figure 3B & 4).

DISCUSSION

The GXXXG motif is a highly conserved pattern of amino acids found in 80% of proteins that form dimers. We further characterized the oligomerization of Ste14 by investigating the contributions of the GXXXG motif in TM1 and determining the role of the peripheral GXXXG residues in the processes of homodimerization, protein expression, activity, stability, and folding. Here, we provide evidence that the GXXXG motif and TM1 of Ste14 comprise the dimerization interface and several of these residues are vital to enzymatic activity. The construction of a cysteine-less, TA, variant of Ste14 that displayed all the same properties of WT Ste14, allowed us to explore part of this dimerization motif by utilizing cysteine scanning mutagenesis to determine which residues in TM1 were important for homodimer formation.

A helical wheel has been constructed for each cross-linker and the strength of cross-linking provides insights into how the two TM1 domains of two different Ste14

monomers interact (Figure 7A). A sulfhydryl reactive cross-linker, BMH, which has a spacer arm of 13 Å, revealed that mutants M25C, T26C, Y28C, F41C, and P42C and Q43C formed strong homodimers. In addition, S27C, L30C, G31C, G35C, G39C, and L40C formed weak homodimers. Of the mutants that were cross-linked in low yields, S27C, L30C and L40C, lie on the opposite side of the helix and potentially form weak homodimers due to the extended length of the BMH cross-linker. High yield of dimer formation was observed with F41C, which is also potentially on the other face of the helix. Overall, the BMH cross-linking experiments revealed that most of the residues that formed strong homodimers were on the same face of the helix as the GXXXG motif. Because BMH is 13 Å long, it is more flexible and can potentially reach to the other side of the helix and cross-link residues that would not normally interact. Therefore, we performed the original cross-linking experiments again with a shorter, more specific cross-linker, BMOE, which is 8 Å long.

We determined that M25C, T26C, S27C, Y28C, F41C, P42C and Q43C all form strong homodimers when cross-linked with BMOE. With the exception of T26, all of these residues lie on the same face of the helix as the three glycine residues, G31, G35, and G39. Y28 and P42 are of particular interest because they surround the G₃₁XXXG₃₅XXXG₃₉ dimerization interface and may play a crucial role in dimerization. We predict Y28 stabilizes Ste14 through intramolecular hydrogen bonds with the backbone carbonyl groups of TM1 and P42 is necessary for proper alpha helix formation (Figure 7B & C). Moreover, we were able to demonstrate that G31C and G35C mutants cross-linked to form a weak dimer after the addition of BMOE. This is expected because G31 and G35 are misfolded and therefore may not be able to interact in the correct conformation to form a strong dimer. These

residues are critical to every other property of the protein including expression, methyltransferase activity, stability, and folding, so it is likely that they are fundamental to dimer formation even though both cross-linkers were able to induce only weak dimers.

G31, G35, and G39 form a groove in TM1 (Figure 7B & C). If any of these glycine residues are mutated, then it is possible that the interacting helix can no longer come into close enough contact to form a stable homodimer. Previous studies of other GXXXG proteins, such as the amyloid precursor protein (APP), support this finding. Specifically, when glycine residues in the GXXXG motif were mutated, APP could no longer form a dimer and had a varied cleavage pattern by α -secretase (46). Additionally, the glycine residues of GpA formed a groove in its GXXXG helices, but mutation of these residues abolished the depth of the groove and strength of dimerization (31). The glycine residues are could also be fundamentally important for the overall three-dimensional structure of the monomer. Without a properly folded monomer structure, Ste14 would be unable to form a strong homodimer. Our trypsin digestion experiments revealed that G31C, G35C, and G39C are critically misfolded and that only G39A could regain a folded structure similar to WT Ste14. Therefore, the inability of G31C, G35C, and G39C to form strong homodimers is partially explained by the misfolded monomeric structure.

To further our knowledge about the Ste14 TM1 residues surrounding the dimerization motif, we investigated the effects of these mutations on the methyltransferase activity of Ste14. It is not surprising that G31C, G35C, or G39C were inactive because of their decreased expression and misfolded state. However, it was hypothesized that G31A, G35A, and G39A could restore methyltransferase

activity because TM1 of human Icm1 contains an AXXXA motif and the structural similarity between a glycine and an alanine amino acid would minimize the overall change in the tertiary protein structure. This was found only to be true for G39A, though the activity was restored only to less than half of WT activity. It is worth noting that T26C, S27C, Y28C, and L30C displayed a loss in methyltransferase activity as significant as the glycine residues. Y28C and L30C displayed a similar cleavage pattern compared to G39C upon trypsin digestion. Meanwhile S27C revealed a similar cleavage to pattern as G31C and G35C. Interestingly, T26C, S27C, Y28C, and L30C formed a homodimer with either BMH or BMOE. Therefore, these mutants must be able to make necessary dimer contacts but are also misfolded enough to abrogate methyltransferase activity.

Overall, this study provides evidence of the importance of the GXXXG motif in dimer formation of Ste14. The three glycine residues, G31, G35, and G39, along with the peripheral Y28 and P42 residues potentially help stabilize the Ste14 homodimer. Additionally, the residues in TM1 comprising and surrounding the GXXXG domain play a vital role in protein expression, stability, and folding. Utilizing the results of this investigation in combination with future studies to provide a more in-depth structural analysis of Icm1, we can design more potent therapeutics to inhibit the function of Icm1 in oncogenic cells with the end goal of minimizing overactive K-Ras signaling in cancer.

EXPERIMENTAL PROCEDURES

Materials

Primers were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). Restriction enzymes were purchased from New England Biolabs (Beverly, MA). The Dr. Richard Gibbs

laboratory (Purdue University, West Lafayette, IN) kindly synthesized N-acetyl-S-farnesyl-L-cysteine (AFC). *S*-adenosyl-L-(¹⁴C-*methyl*) methionine was purchased from Perkin Elmer (Waltham, MA). *Bis*-maleimidoethane, BMH, *Bis*-maleimidoethane, BMOE, and enhanced chemiluminescence (ECL) were purchased from Pierce (Rockford, IL). Trypsin was purchased from Sigma-Aldrich (St. Louis, MO). Micrococcal nuclease was purchased from Worthington Biochemical Corp. (Lakewood, NJ). Aprotinin was purchased from MP Biomedical (Irvine, CA). The SM1188 yeast strain and the α -Ste14 polyclonal antibody were gifts from Dr. S. Michaelis (The Johns Hopkins University School of Medicine, Baltimore, MD). 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF), aprotinin, and dithiothreitol (DTT) were purchased from Gold Biotechnology (Olivette, MO). The *E. coli* DH5 α subcloning efficiency cells, the monoclonal α -myc IgG antibody, the goat anti-mouse IgG, and the goat α -rabbit IgG were all purchased from Invitrogen (Carlsbad, CA). Nitrocellulose membranes (0.22 μ m) were purchased from GE Healthcare Life Sciences (Pittsburg, PA). N-Dodecyl- β -D-maltopyranoside (DDM) was purchased from Anatrace, Inc. (Maumee, OH). TALON metal affinity resin was purchased from Clontech (Mountain View, CA). Amicon Ultra 30,000 MWCO concentrators were purchased from Millipore Co. (Billerica, MA). pFastBac vectors were purchased from ThermoFisher (Waltham, MA). Sf9 cells and ES 921 media were purchased from Expression Systems (Davis, CA). All other materials and reagents were purchased from Fisher Scientific (Pittsburgh, PA).

Site-Directed Mutagenesis and Cloning

Construction of the pCHH₁₀m₃N plasmid encoding WT Ste14 tagged at the N-

terminus with ten histidine repeats and three myc epitopes under the control of the PGK promoter has been described previously (29). Two-step PCR was utilized to mutate all three native cysteine residues (C99, C121, and C126) to alanine. Subsequently, residues 25-43 were mutated to cysteine and/or alanine residues using site-directed mutagenesis. The resulting PCR products were digested sequentially with restriction enzymes *Eag* I and *Sac* II, ligated into the digested pCHH₁₀m₃N and then transformed into DH5 α subcloning efficiency cells. The Ste14 mutations were confirmed by bidirectional sequencing.

Yeast Strains

Each TM1 Ste14 mutant plasmid strain (2μ *URA3* *P_{PGK}* *His₁₀-myc₃-STE14-X* mutation) was transformed into a Δ *ste14* deletion strain, SM1188 (*STE14-3::TRP1*, *MATa* *trp1* *leu2* *ura3* *his4* *can1*) by the Elble method (47). Efficiency of transformation was increased by adding dithiothreitol to a final concentration of 50 mM. The yeast cells were grown in SCURA to mid-log phase (OD₆₀₀ 3-5), pelleted by centrifugation at 5316 x g and stored at 80°C.

Crude Membrane Preparations from Yeast Cells

For experiments using crude membrane protein, the membranes and proteins were prepared as previously described with minor modifications (29). Briefly, the yeast cell pellets were resuspended in lysis buffer (300 mM sorbitol, 100 mM NaCl, 6 mM MgCl₂·6H₂O, 10 mM Tris-HCl, pH 7.5, 1% aprotinin, 1 mM AEBSF and 2mM DTT) and incubated on ice for 15 min to swell. The cell pellets underwent two freeze-thaw cycles with liquid nitrogen. The cells were then lysed with a French press at 12,000 p.s.i. The cellular mixture was spun down twice at 500 x g for 10 min at 4°C and the supernatant was

collected. The membrane fraction was isolated through ultracentrifugation at 100,000 x g for 1 hour at 4°C. Following centrifugation, the membrane pellet was resuspended in 10 mM Tris-HCl, pH 7.5, aliquoted, flash frozen in liquid nitrogen and stored at -80°C. Total protein concentration was determined using the Bradford protein assay.

Expression and Membrane Preparation of Ste14 from Insect Cells

cDNA encoding *Saccharomyces cerevisiae* Ste14 were subcloned into the pFastBac-Dual vector and confirmed by sequencing. Protein was expressed in baculovirus-infected Sf9 cells and harvested from ES 921 media with 1x Pencillin-Septamycin at 48 h post-infection. Pelleted cells were flash frozen for membrane preparation. Briefly, the insect cell pellets were thawed and resuspended in lysis buffer (50 mM mannitol, 50 mM Tris-HCl, pH 7.5, 1% aprotinin, 1 mM AEBSF and 2mM DTT) and were rocked for 15 min at 4°. The resuspended pellets were lysed via dounce homogenization and isolated through ultracentrifugation at 100,000 x g for 1 hour at 4°C. Following centrifugation, the membrane pellet was resuspended in buffer (300 mM mannitol, 50 mM Tris-HCl, pH 7.5, 1% aprotinin, 1 mM AEBSF and 1mM DTT, 10% glycerol), aliquoted, flash frozen in liquid nitrogen and stored at -80°C. Total protein concentration was determined using the Bradford protein assay.

Purification of Histidine-tagged Ste14

Histidine-tagged Ste14 was purified as previously described (29). Briefly, crude membrane protein was solubilized in lysis buffer (0.45 M Na₂HPO₄, 0.23 M NaCl, 1% aprotinin, 2 mM AEBSF, and 10% glycerol) with 20 mM imidazole and 1% DDM (w/v) for 1 hour at 4°C. The insoluble fraction was removed by ultracentrifugation at 100,000 x

g for 45 min at 4°C. The soluble fraction was incubated for 1 hour at 4°C with Talon metal affinity beads in lysis buffer. The resin underwent four wash steps with gentle rocking at 4°C and centrifugation at 350 x g for 2 min before aspiration of excess buffer. The steps included two washes with buffer 1 (lysis buffer plus 1% DDM (w/v) and 40 mM imidazole), one wash with buffer 2 (lysis buffer plus 1% DDM (w/v), 500 mM KCl and 40 mM imidazole) and one wash with buffer 3 (lysis buffer plus 0.1% DDM (w/v), 515 mM KCl, 20 mM imidazole). Protein was eluted with buffer 4 (lysis buffer plus 0.1% DDM and 1 M imidazole) and concentrated using a 30,000 MWCO concentrator at 5,316 x g for 20 min at 4°C. 1mL of purified protein was loaded onto one Superdex 200 Increase column (GE Healthcare) in a 0.45 M Na₂HPO₄, 0.23 M NaCl, 2mM DTT, and 0.05% DDM (w/v) buffer, pH 6. The protein was eluted, pooled, and concentrated as described above. Protein concentration was determined using an Amido Black protein assay (48).

SEC-MALS-SAXS data collection and analysis

SEC-MALS-SAXS experiments were performed at beamline 18-ID (BioCAT) of the advanced photon source (APS) at Argonne National Laboratory using a Pilatus3 1M detector (160 µm x 75 µm) and a 12 KeV X-ray with 1.033 wavelength and ~3.5 m sample-to-detector distance. All protein was purified in a buffer comprised of 50mM Na₂HPO₄, 0.3M NaCl, 2mM DTT, and 0.05% DDM, pH 6. The purified protein samples were loaded on the SEC column with an exclusion limit of 2.5 MDa (WTC-030S5, Wyatt Technologies), which was followed by a UV detector, a multi-angle light scattering (MALS) detector, a dynamic light scattering (DLS) detector (DAWN Helios II, Wyatt Technologies), a differential refractive index detector (Optilab T-rEX, Wyatt

Technologies) and finally the SAXS flow cell. 0.5 s exposures were acquired every 3 s at a flow rate of 0.8 mL/min. The exposures either flanking or on either side of the elution peak were averaged to generate a buffer file that was subtracted from the exposures corresponding to selected regions of the elution peak to get the $I(q)$ vs q SAXS curves. Data analysis was performed using the ATSAS package (49). PRIMUS from the ATSAS program suite was used to perform buffer subtraction, Guinier approximation based forward scattering (I_0) analysis, and radius of gyration (R_g) determination (50). Scattering curves were further analyzed using GNOM (51) for the calculation of I_0 , R_g , distance distribution $P(r)$, maximum dimension (D_{max}). Molecular weights were calculated from the MALS data using the ASTRA software (Wyatt Technologies). Plots were generated from buffer subtracted averaged data and plotted using GraphPad Prism v.7.0.

In Vitro Methyltransferase Vapor Diffusion Assay

Reactions were carried out with 5 μ g of crude membranes, 200 μ M *N*-acetyl-*S*-farnesyl-L-cysteine (AFC), and 20 μ M (14 C-methyl) *S*-adenosyl-L-methionine (SAM) in 100 mM Tris-HCl, pH 7.5 in 60 μ L. The reactions were incubated at 30 °C for 30 min and then terminated by the addition of 1M NaOH/1% SDS. Each reaction (100 μ L) was then spotted on pleated filter papers and allowed to diffuse at room temperature in scintillation vials containing 10 mL scintillation fluid. After 3 h the filters were removed and the base labile (14 C) methyl groups transferred were counted by a Packard Tri-carb 1600CA liquid scintillation counter. All assays were performed in triplicate. Graph of methyltransferase data was generated using GraphPadPrism version 7.0. One way ANOVA was performed with

Prism7.0 and all error bars represent standard deviation.

SDS-PAGE and Immunoblot Analysis

Protein samples were placed into a 65°C heat block for 30 min after the addition of SDS sample buffer to a final concentration of 2X SDS. The proteins were separated by a 7.5%, 10%, or 12% acrylamide SDS-PAGE gel and transferred to nitrocellulose membranes (0.22 μ m) at 100 V for 90 min. The membranes were blocked at rt for 2 h in 20% (w/v) nonfat dry milk in phosphate-buffered saline with Tween-20 (137 mM NaCl, 2.7 mM KCl, 4 mM Na₂HPO₄, 1.8 mM KH₂PO₄, and 0.05% (v/v) Tween-20, pH 7.4) (PBST). Then, the membrane was incubated with primary antibody (1:500 α -Ste14 or 1:10000 α -myc) and secondary antibody (1:10,000 goat α -rabbit HRP or 1:4,000 goat α -mouse HRP, respectively) diluted in 5% (w/v) dry milk dissolved in PBST. The membranes were washed three times with PBST and visualized on a GeneGnome imaging system after the addition of SuperSignal West Pico enhanced chemiluminescence.

Limited Trypsin Digestion of TM1 Ste14 Mutants

Limited trypsin digestion was carried out as previously described (52) with the following modifications. Assays contained 20 μ g of crude membrane protein, 0.25 mg/ml trypsin, and 10 mM Tris-HCl, pH 7.5. Reactions were incubated at 37 °C for 15 min and stopped by the addition 1 mM AEBSF. Total protein was precipitated by addition of 10% TCA and incubated at 4 °C for 10 min. The reaction was centrifuged for 5 min at 4 °C and the protein pellet was resuspended in reducing SDS-PAGE sample buffer. The samples were heated at 65 °C for 30 min before being separated on a 12% SDS-PAGE gel and followed the protocol previously stated. Immunoblot analysis was performed

using α -myc and α -mouse antibodies. All gels were performed in triplicate.

Glutaraldehyde Crosslinking Analysis

Assays contained 100 μ g of crude membrane protein, 0.2% (wt) glutaraldehyde, 40 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 25 mM KCl, 0.5 mM DTT, and 10% glycerol. Reactions were incubated at 37 °C for 5 min and then stopped by the addition of 1M Tris-HCl, pH 8.0 and reducing SDS-PAGE sample buffer. The samples were then heated for 30 min at 65 °C and separated on a 7.5% SDS-PAGE gel. Cross-linking reactions were detected by immunoblot analysis using α -myc and α -mouse antibodies. 5 μ g pure Ste14 and TA were

analyzed with the same protocol as above. All assays were performed in triplicate.

Cysteine Crosslinking Analysis

All crosslinking assays contained 80 μ g of crude membrane protein, 0.4 mM BMOE or BMH (bi-functional sulfhydryl reactive cross-linkers with a spacer arm of 8 and 13 Å, respectively), 10 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS) pH 7.0, and 5 mM MgCl₂. Reactions were incubated at room temperature for 30 min and stopped by the addition non-reducing SDS-PAGE sample buffer containing 40 mM cysteine. The samples were separated on a 7.5% SDS-PAGE gel and cross-linking was detected by immunoblot analysis using α -myc and α -mouse antibodies. All assays were performed in triplicate.

Acknowledgements: We would like to thank Dr. S. Michaelis (The Johns Hopkins University School of Medicine) for α -Ste14 polyclonal antibodies and Dr. Richard Gibbs laboratory (Purdue University) for synthesizing N-acetyl-S-farnesyl-L-cysteine (AFC). We thank Dr. A. Lyon for the helpful discussion of our data. This research used resources of the Advanced Photon Source, a United States Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract DE-AC02-06CH11357.

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Hrycyna, C. A., Sapperstein, S. K., Clarke, S., and Michaelis, S. (1991) The *Saccharomyces cerevisiae* STE14 gene encodes a methyltransferase that mediates C-terminal methylation of afactor and RAS proteins. *EMBO J.* **10**, 1699–1709
2. Romano, J. D., and Michaelis, S. (2001) Topological and Mutational Analysis of *Saccharomyces cerevisiae* Ste14p, Founding Member of the Isoprenylcysteine Carboxyl Methyltransferase Family. *Mol. Biol. Cell.* **12**, 1957–1971
3. Romano, J. D., Schmidt, W. K., and Michaelis, S. (1998) The *Saccharomyces cerevisiae* Prenylcysteine Carboxyl Methyltransferase Ste14p Is in the Endoplasmic Reticulum Membrane. *Mol. Biol. Cell.* **9**, 2231–2247
4. Philips, M., and Wright, L. (2006) CAAX modification and membrane targeting of Ras. *J. Lipid Res.* **47**, 883–891
5. Wright, L. P., Court, H., Mor, A., Ahearn, I. M., Casey, P. J., and Philips, M. R. (2009) Topology of Mammalian Isoprenylcysteine Carboxyl Methyltransferase Determined in Live Cells with a Fluorescent Probe. *Mol. Cell. Biol.* **29**, 1826–1833
6. Xu, N., Shen, N., Wang, X., Jiang, S., Xue, B., and Li, C. (2015) Protein prenylation and human diseases: a balance of protein farnesylation and geranylgeranylation. *Sci. China Life Sci.* **58**, 328–335
7. Perez-Sala, D. (2007) Protein isoprenylation in biology and disease: general overview and perspectives from studies with genetically engineered animals. *Front Biosci.* **12**, 4456–4472
8. Berndt, N., Hamilton, A. D., and Sebt, S. M. (2011) Targeting protein prenylation for cancer therapy. *Nat. Rev. Cancer.* **11**, 775–791
9. Barbacid, M. (1995) Farnesyltransferase inhibitors are inhibitors of ras but not r-ras2/tc21, transformation. *Oncogene.* **10**, 1905–1913
10. Zhang, F. L., and Casey, P. J. (1996) Protein Prenylation: Molecular Mechanisms and Functional Consequences. *Annu. Rev. Biochem.* **65**, 241–269
11. Moores, S. L., Schaber, M. D., Mosser, S. D., Rands, E., O'Hara, M. B., Garsky, V. M., Marshall, M. S., Pompliano, D. L., and Gibbs, J. B. (1991) Sequence dependence of protein isoprenylation. *J Biol Chem.* **22**, 14603-14610
12. Zhang, F. L., and Casey, P. J. (1996) Protein prenylation: molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* **65**, 241–269
13. Hrycyna, C. A., and Clarke, S. (1993) Modification of eukaryotic signaling proteins by C-terminal methylation reactions. *Pharmacol. Ther.* **59**, 281–300
14. Casey, P. J., and Seabra, M. C. (1996) Protein prenyltransferases. *J. Biol. Chem.* **271**, 5289–5292
15. Court, H., Hahne, K., Philips, M. R., and Hrycyna, C. A. (2011) Biochemical and Biological Functions of Isoprenylcysteine Carboxyl Methyltransferase. *Enzymes.* **30**, 71–90

16. Winter-Vann, A. M., and Casey, P. J. (2005) Post-prenylation-processing enzymes as new targets in oncogenesis. *Nat Rev Cancer*. **5**, 405–412
17. Gao, J., Liao, J., and Yang, G.-Y. (2009) CAAAX-box protein, prenylation process and carcinogenesis. *Am. J. Transl. Res.* **1**, 312–325
18. O'Hagan, R. C., and Heyer, J. (2011) KRAS Mouse Models: Modeling Cancer Harboring KRAS Mutations. *Genes Cancer*. **2**, 335–343.
19. Bergo, M. O., Leung, G. K., Ambroziak, P., Otto, J. C., Casey, P. J., and Young, S. G. (2000) Targeted Inactivation of the Isoprenylcysteine Carboxyl Methyltransferase Gene Causes Mislocalization of K-Ras in Mammalian Cells. *J. Biol. Chem.* **275**, 17605–17610
20. Bergo, M. O., Gavino, B. J., Hong, C., Beigneux, A. P., McMahon, M., Casey, P. J., and Young, S. G. (2004) Inactivation of Icmt inhibits transformation by oncogenic K-Ras and B-Raf. *J. Clin. Invest.* **113**, 539–550
21. Scott Reid, T., Terry, K. L., Casey, P. J., and Beese, L. S. (2004) Crystallographic Analysis of CaaX Prenyltransferases Complexed with Substrates Defines Rules of Protein Substrate Selectivity. *J. Mol. Biol.* **343**, 417–433
22. Michaelson, D., Ali, W., Chiu, V. K., Bergo, M., Silletti, J., Wright, L., Young, S. G., and Philips, M. (2005) Postprenylation CAAAX processing is required for proper localization of Ras but not Rho GTPases. *Mol. Biol. Cell.* **16**, 1606–1616
23. Bergo, M. O., Leung, G. K., Ambroziak, P., Otto, J. C., Casey, P. J., Gomes, A. Q., Seabra, M. C., and Young, S. G. (2001) Isoprenylcysteine Carboxyl Methyltransferase Deficiency in Mice. *J. Biol. Chem.* **276**, 5841–5845
24. Lau, H. Y., Ramanujulu, P. M., Guo, D., Yang, T., Wirawan, M., Casey, P. J., Go, M.-L., and Wang, M. (2014) An improved isoprenylcysteine carboxylmethyltransferase inhibitor induces cancer cell death and attenuates tumor growth in vivo. *Cancer Biol. Ther.* **15**, 1280–1291
25. Karnoub, A. E., and Weinberg, R. A. (2008) Ras oncogenes: split personalities. *Nat Rev Mol Cell Biol.* **9**, 517–531
26. Prior, I. A., Lewis, P. D., and Mattos, C. (2012) A Comprehensive Survey of Ras Mutations in Cancer. *Cancer Res.* **72**, 2457–2467
27. Bos, J. L. (1989) ras Oncogenes in Human Cancer: A Review. *Cancer Res.* **49**, 4682–4689
28. Bos, J. L. (1988) The ras gene family and human carcinogenesis. *Mutat. Res. Genet. Toxicol.* **195**, 255–271
29. Anderson, J. L., Frase, H., Michaelis, S., and Hrycyna, C. A. (2005) Purification, Functional Reconstitution, and Characterization of the *Saccharomyces cerevisiae* Isoprenylcysteine Carboxylmethyltransferase Ste14p. *J. Biol. Chem.* **280**, 7336–7345
30. Gerber, D. and Shai, Y. (2001) In vivo detection of hetero-association of glycophorin-A and its mutants within the membrane. *J Biol Chem.* **276**, 31229–31232
31. MacKenzie, K. R., Prestegard, J. H., and Engelman, D. M. (1997) A transmembrane helix dimer: structure and implications. *Science*. **276**, 131–133
32. Bhatia, A., Schafer, H. J., and Hrycyna, C. A. (2005) Oligomerization of the human ABC transporter ABCG2: evaluation of the native protein and chimeric dimers. *Biochemistry*. **44**, 10893–10904

33. Polgar, O., Robey, R. W., Morisaki, K., Dean, M., Michejda, C., Sauna, Z. E., Ambudkar, S. V., Tarasova, N., and Bates, S. E. (2004) Mutational analysis of ABCG2: role of the GXXXG motif. *Biochemistry*. **43**, 9448-9456.
34. Arselin, G., Giraud, M. F., Dautant, A., Vaillier, J., Brethes, D., Coulary-Salin, B., Schaeffer, J., and Velours J. (2003) The GxxxG motif of the transmembrane domain of subunit e is involved in the dimerization/oligomerization of the yeast ATP synthase complex in the mitochondrial membrane. *Eur. J. Biochem.* **270**, 1875-1884
35. Whittington, D. A., Waheed, A., Ulmasov, B., Shah G. N., Grubb, J. H., Sly, W. S., and Christianson, D. W. (2001) Crystal structure of the dimeric extracellular domain of human carbonic anhydrase XII, a bitopic membrane protein overexpressed in certain cancer tumor cells. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 9545-9550
36. McClain, M. S., Iwamoto, H., Cao, P., Vinion-Dubiel, A. D., Li, Y., Szabo, G., Shao, Z., and Cover, T. L. (2003) Essential role of a GXXXG motif for membrane channel formation by *Helicobacter pylori* vacuolating toxin. *J. Biol. Chem.* **278**, 12101-12108
37. Overton, M. C., Chinault, S. L., and Blumer, K. J., (2003) Oligomerization, biogenesis, and signaling is promoted by a glycophorin A-like dimerization motif in transmembrane domain 1 of a yeast G protein-coupled receptor. *J. Biol. Chem.* **278**, 49369-49377
38. Yang, J., and Barford, D. (2011) Mechanism of Isoprenylcysteine Carboxyl Methylation from the Crystal Structure of the Integral Membrane Methyltransferase ICMT. *Mol. Cell.* **44**, 997-1004
39. Diver, M. M., Pedi, L., Koide, A., Koide, S., and Long, S. B. (2018) Atomic structure of the eukaryotic intramembrane RAS methyltransferase ICMT. *Nature* **553**, 526-529
40. Griggs, A. M., Hahne, K., and Hrycyna, C. A. (2010) Functional oligomerization of the *Saccharomyces cerevisiae* isoprenylcysteine carboxyl methyltransferase, Ste14p. *J. Biol. Chem.* **285**, 13380-13388
41. Lemmon, M. A., Flanagan, J. M., Treutlein, H. R., Zhang, J., and Engelman D. M. (1992) Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry*. **31**, 12719-12725
42. Lemmon, M. A., Treutlein, H. R., Adams, P. D., Brunger, A. T., and Engelman D. M. (1994) A dimerization motif for transmembrane alpha-helices. *Nat. Struct. Biol.* **1**, 157-163
43. Russ, W. P. and Engelman, D. M. (2000) The GxxxG motif: a framework for transmembrane helix-helix association. *J. Mol. Biol.* **296**, 911-919
44. Cuthbertson, J. M., Bond, P. J., and Sansom, M. S. (2006) Transmembrane helix-helix interactions: comparative simulations of the glycophorin a dimer. *Biochemistry* **45**, 14298-14310.
45. Melnyk, R. A., Kim, S., Curran, A. R., Engelman, D. M., Bowie, J. U., and Deber, D. M. (2004) The affinity of GXXXG motifs in transmembrane helix-helix interactions is modulated by long-range communication. *J. Biol. Chem.* **279**, 16591-16597
46. Munter, L. M., Voigt, P., Harmeier, A., Kaden, D., Gottschalk, K. E., Weise, C., Pipkorn, R., Schaefer, M., Langosch, D., and Multhaup, G. (2007) GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of Abeta42. *Embo. J.* **26**, 1702-1712
47. Elble, R. (1992) A simple and efficient procedure for transformation of yeasts. *Biotechniques*. **13**, 18-20

48. Schaffner, W., and Weissmann, C. (1973) A rapid, sensitive, and specific method for the determination of protein in dilute solution. *Anal. Biochem.* **56**, 502–514
49. Franke, D., Petoukhov, M., Konarev, P., Panjkovich, A., Tuukkanen, A., Mertens, H., Kikhney, A., Hajizadeh, N., Franklin, J., Jeffries, C. (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212-1225
50. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC -based system for small -angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277 -1282
51. D. Svergun (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria *J. Appl. Crystallogr.* **25**, 495-503
52. Fung, B. K., and Nash, C. R. (1983) Characterization of transducin from bovine retinal rod outer segments. II. Evidence for distinct binding sites and conformational changes revealed by limited proteolysis with trypsin. *J. Biol. Chem.* **258**, 10503-10510
53. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods.* **12**, 7-8
54. Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725-738
55. Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinform.* **9**, 40
56. Sievers F., Wilm A., Dineen D., Gibson T. J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J. D. and Higgins D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539.
57. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. and Barton, G. J. (2009) Jalview Version 2 - a multiple sequence alignment editor and analysis workbench *Bioinformatics* doi: 10.1093/bioinformatics/btp033

FOOTNOTES

Funding was provided by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health Grant (R01 GM106082-05), the Purdue Research Foundation, and Purdue University Center for Cancer Research. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health and the National Institute of General Medical Sciences.

The abbreviations used are: AEBSF, 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride; AFC, N-acetyl-S-farnesyl-L-cysteine; *Ag-Icmt*, *Anopheles gambiae* isoprenylcysteine carboxyl methyltransferase; amyloid precursor protein, APP; ATP binding cassette subfamily G member 2, ABCG2; BMH, *bis*-maleimidoethane; BMOE, *bis*-maleimidoethane; carbonic anhydrase, CA; DTT, dithiothreitol; DDM, *n*-dodecyl- β -D-maltopyranoside; D_{max} , dimensional maximum; ECL, enhanced chemiluminescence; ER, endoplasmic reticulum; FTase, farnesyltransferase, GGTase1, geranylgeranyltransferase I; GpA, glycophorin A; hlcmt, human isoprenylcysteine carboxyl methyltransferase; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HPR, horseradish peroxidase; HCl, hydrochloric acid; Icmt, isoprenylcysteine carboxyl methyltransferase; IgG, immunoglobulin G; KCl, potassium chloride; kDa, kilodalton; *Ma-Icmt*, *Methanosarcina acetivorans* isoprenylcysteine carboxyl methyltransferase; MALS, multi-angle light scattering; MgCl₂, magnesium chloride; MOPS, 3-(*N*-morpholino)propanesulfonic acid; MWCO, molecular weight cutoff; NaCl, sodium chloride; NaOH, Sodium hydroxide; Na₂HPO₄, sodium phosphate dibasic; OD, optical density; PBS, phosphate buffered saline; PBST, Phosphate buffered saline + 0.05% Tween 20; PCR, polymerase chain reaction; P(r), pair-distance distributions; p.s.i, pounds per square inch; PTM, post-translational modification; Rce1, ras converting enzyme 1; R_g, radius of gyration; SAH, S-adenosyl-homocysteine; SAM, S-adenosyl-L-methionine; [¹⁴C]-SAM, S-adenosyl-L-[methyl-¹⁴C] methionine; SAXS, small angle light scattering; SC-URA, synthetic complete minus uracil; SDS, sodium dodecyl sulfate; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; S.D., standard deviation; SEC, size exclusion chromatography; TCA, trichloroacetic acid; TM, transmembrane; UV, ultraviolet; v/v, volume per volume; WT, wild-type; wt, Weight; wt/v, weight per volume; YPD, yeast peptone dextrose

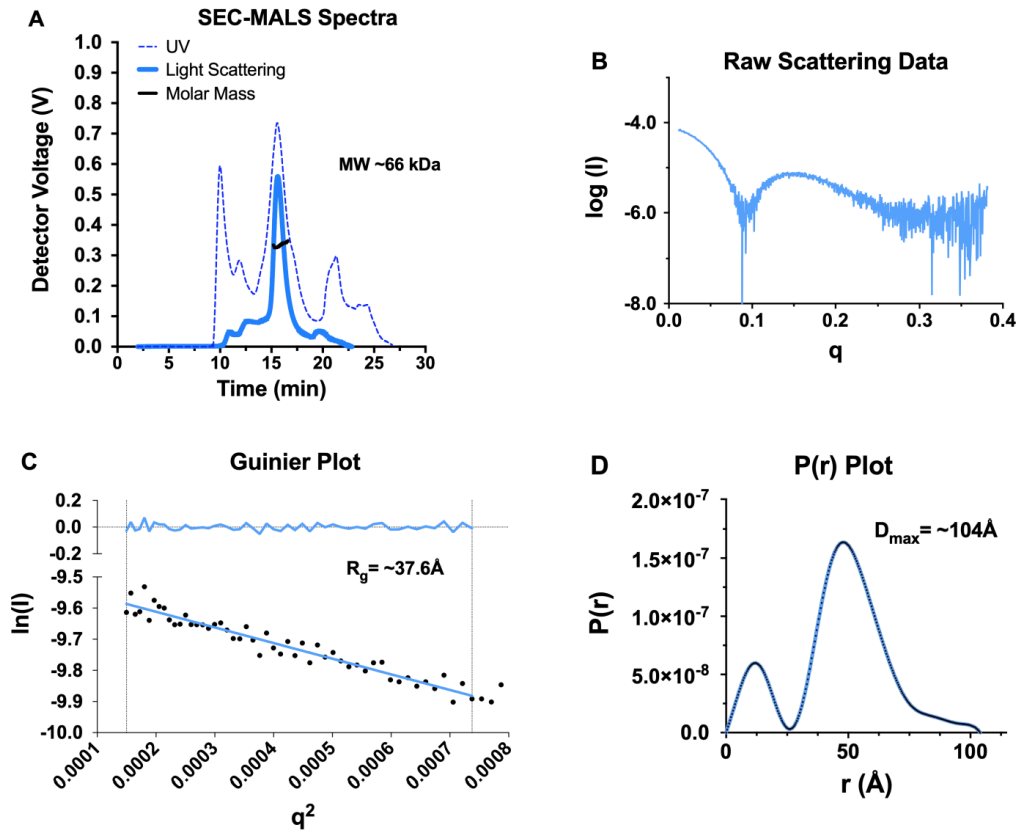


Figure 1: Characterization of Ste14 via SEC-MALS-SAXS confirms dimerization state. (A) SEC-MALS indicated one molecular weight for purified Ste14. Extrapolated averaged molecular weights estimate a molecular mass of $67.26 \text{ kDa} \pm 2.811$ and confirms the homodimeric nature of Ste14 ($\sim 70 \text{ kDa}$). (B) Log beam intensity (I) vs scattering angle magnitude (q), after buffer subtraction. (C) Guinier plot of low q values, $\ln(I)$ vs. q^2 , with radius of gyration (R_g) and y-intercept (I_0). Fitting of the blue line to the data is represented by residuals at the top of the plot. R_g was determined to be 37.6 ± 0.303 . (D) $P(r)$ plot with maximum interparticle distance (D_{\max}) of $\sim 104 \text{ \AA}$.

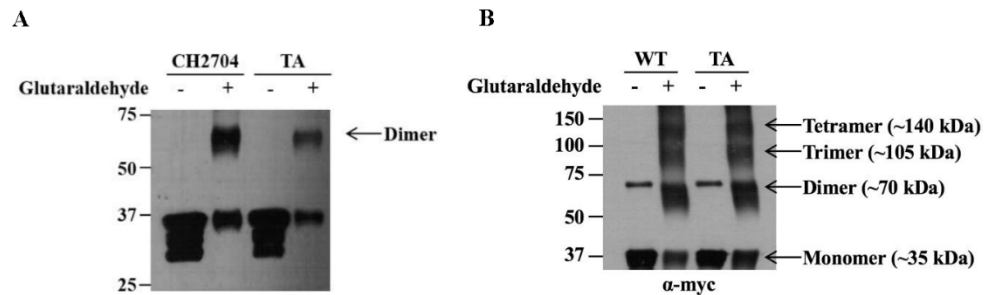


Figure 2: Ste14 forms a homodimer in both crude membranes and pure samples. (A) 100 μ g of crude membrane protein expressing WT Icmt was crosslinked at room temperature using 0.2% (wt/v) glutaraldehyde. Following cross-linking reactions, samples were analyzed via immunoblot on a 7.5% SDS-PAGE gels, using α -myc (1:5,000) and α -mouse (1:4,000). (B) 5 μ g of pure protein was crosslinked at room temperature using 0.2% (wt/v) glutaraldehyde. Following cross-linking reactions, samples were analyzed via immunoblot on a 7.5% SDS-PAGE gels and detected using α -myc (1:10,000) and α -mouse (1:4,000). The molecular weight of Ste14 is ~35kDa, ~70kDa, ~105kDa, and ~140kDa as a monomer, dimer, trimer, and tetramer, respectively.

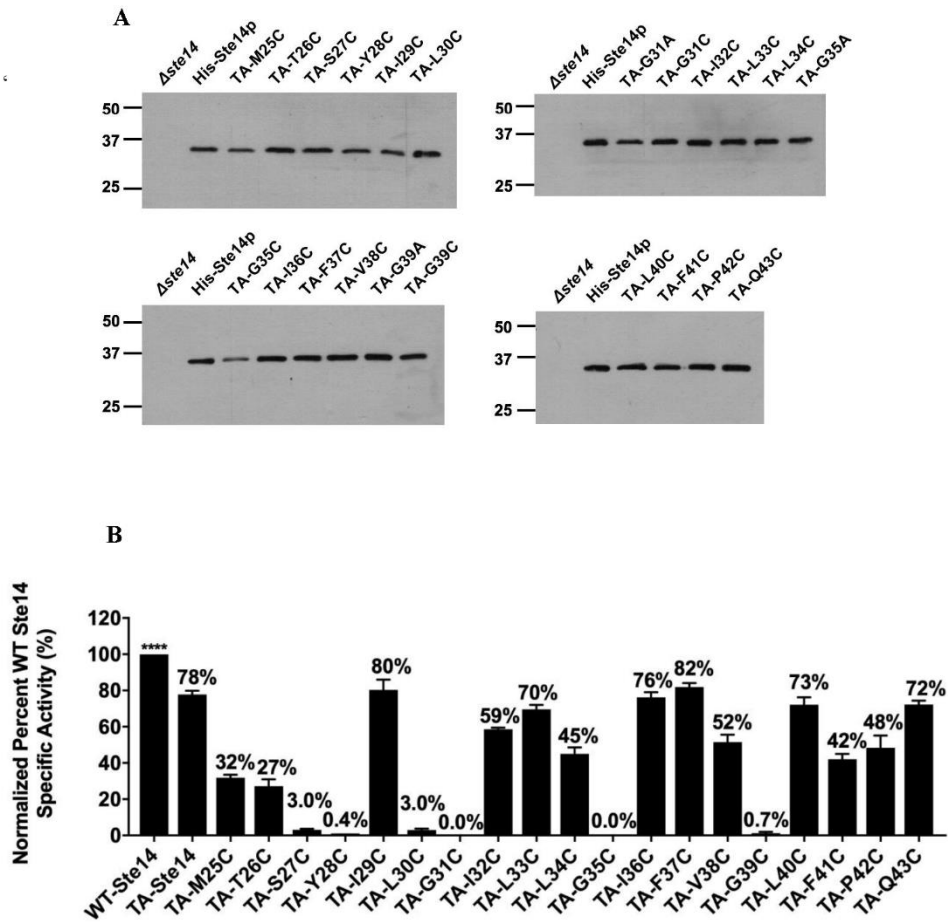


Figure 3: Expression and methyltransferase activity of TM1 cysteine and alanine mutants compared to WT in proposed dimerization site of Ste14 (A) 1 μ g of crude membrane protein was loaded onto a 10% SDS-PAGE gel, followed by immunoblot analysis using α -Ste14 (1:500) and α -rabbit (1:10,000). (B) Methyltransferase assay of all Ste14 TM1 cysteine and alanine mutants. The methyltransferase activity assay was performed with 5 μ g of crude membranes, 200 μ M AFC and 20 μ M 14 C-SAM. Methyltransferase activity assays were performed in triplicate. Graph was generated using GraphPadPrism version 7.0. One way ANOVA was performed with Prism7.0 and all error bars represent standard deviation.

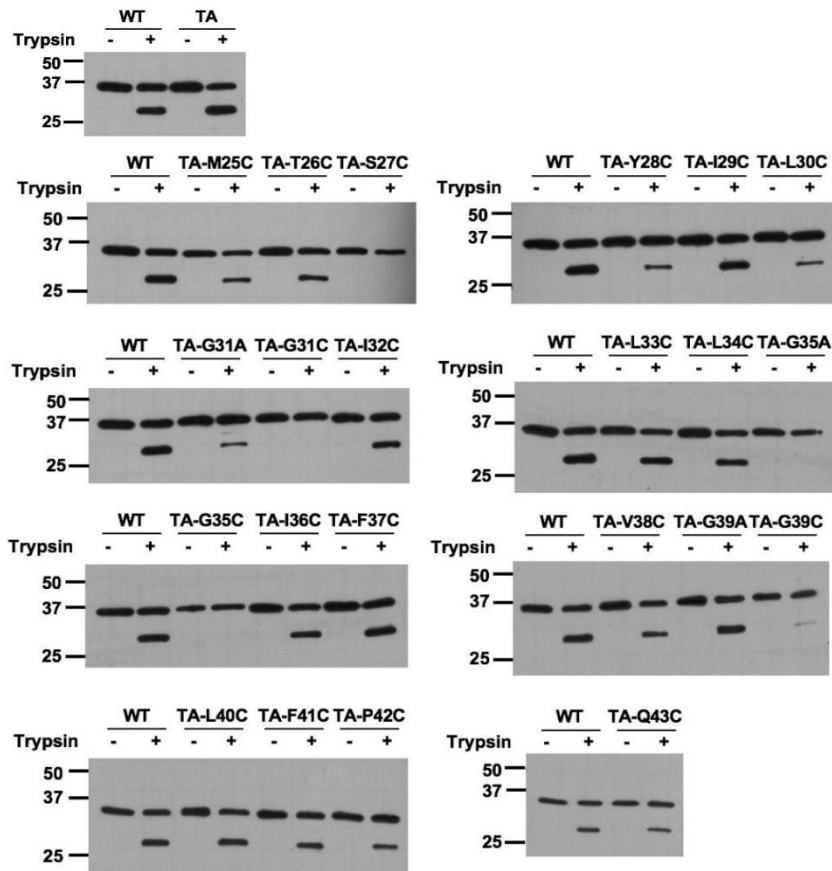


Figure 4: Trypsin digestion of all TM1 cysteine and alanine mutants compared to WT. 2.5 μ g of crude membrane protein was loaded onto a 12% SDS-PAGE gel, followed by immunoblot analysis with α -myc (1:10,000) and α -mouse (1:4000). Trypsin digestion revealed that the TM1 Ste14 mutants retain largely the same folded structure as WT except S27C, Y28C, L30C, G31A, G31C, G35A, G35C, and G39C.

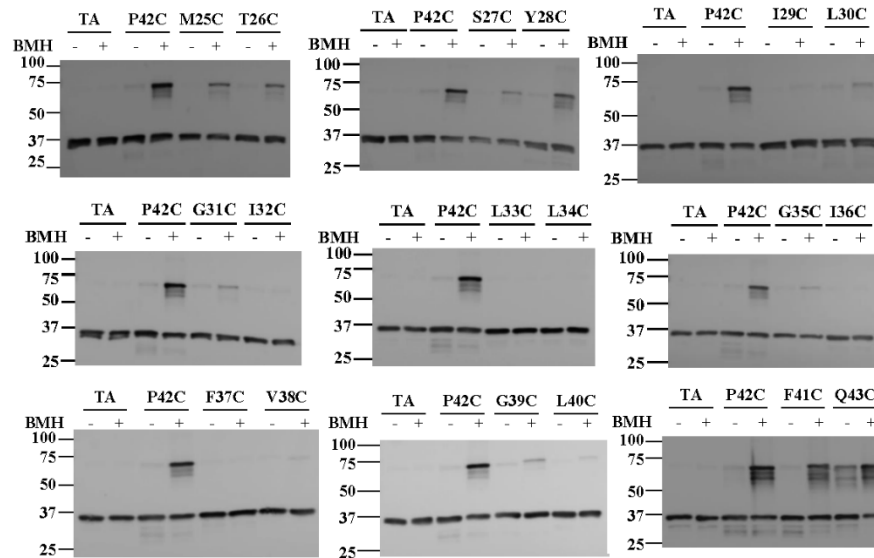


Figure 5: Crosslinking analysis of TM1 Ste14 mutant upon the addition of BMH. 80 μ g protein from crude membranes was crosslinked using 0.4 mM BMH. Reactions were incubated at room temperature for 30 min and stopped by the addition 2X non-reducing SDS containing 40 mM cysteine. The samples were separated on a 7.5% SDS-PAGE gel, followed by an immunoblot analysis using α -myc (1:5,000) and α -mouse (1:4,000).

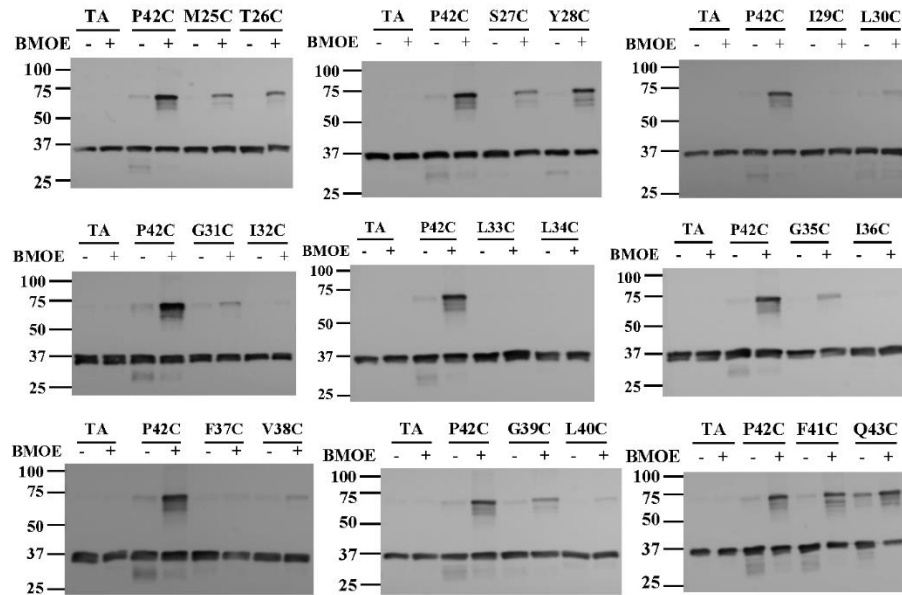


Figure 6: Crosslinking analysis of TM1 Ste14 mutant upon the addition of BMOE. 80 μ g of protein from crude membranes was crosslinked using 0.4 mM BMOE. Reactions were incubated at room temperature for 30 min and stopped by the addition 2X non-reducing SDS containing 40mM cysteine. The samples were separated on a 7.5% SDS-PAGE gel, followed by an immunoblot analysis using α -myc (1:5,000) and α -mouse (1:4,000).

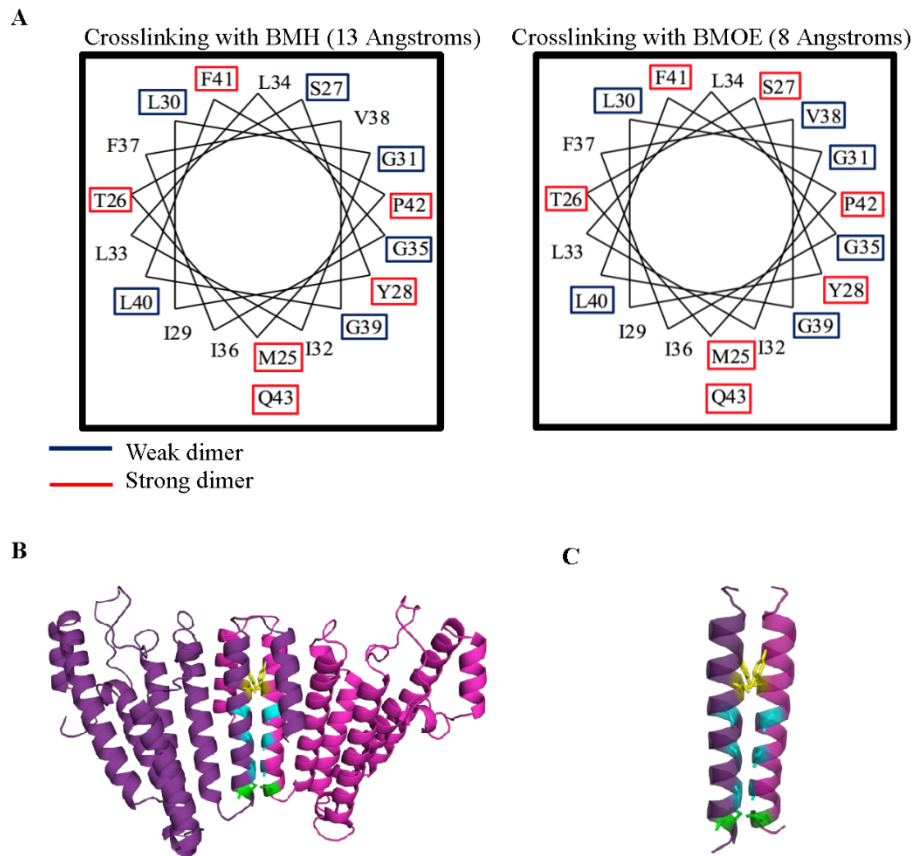
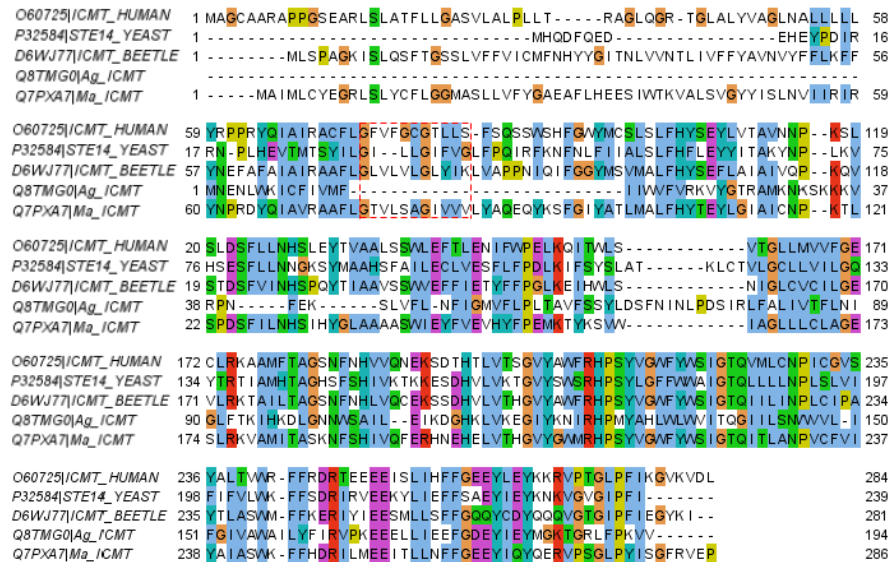
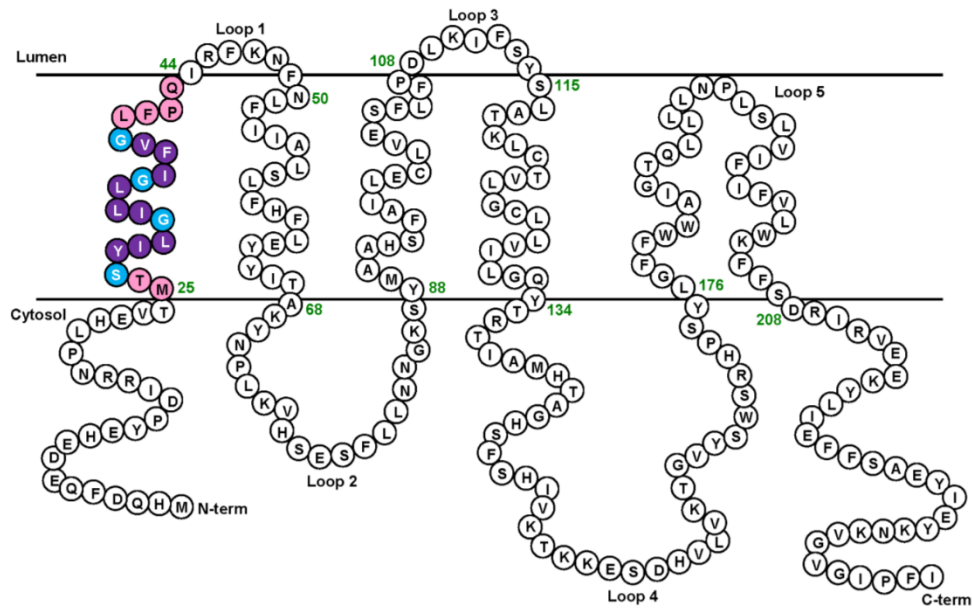


Figure 7: Modeling of Ste14 dimerization in TM1 (A) Helical wheel diagrams of TM1 crosslinking of Ste14 with BMH or BMOE. The helical wheel diagram of TM1 residues of Ste14 was created using the WinPep Software. Strong and weak dimer formations upon the addition of sulfhydryl specific cross-linkers BMH and BMOE are shown by red and blue boxes, respectively. (B) Our predicted model of Ste14 homodimerization, where TM1 residues of two Ste14 molecules are shown (pink and purple). (C) The glycine residues (cyan) comprising the repeating GXXXG motif in TM1 create a groove for the close packing of another transmembrane helix. Tyrosine 28 (yellow) and proline 42 (green) may stabilize the dimerization interaction by contacting and holding the other helix in place. A 3D homology alignment was generated with I-TASSER using the beetle Icm1 (5V7P) as a structural template for Ste14 (52-55). PyMol was used to generate the model.

Dimerization motif of yeast ICMT



Supporting Figure 1: Sequence alignment of IcmT orthologs. Sequence alignment of human IcmT (hIcmT; 284 residues), *Saccharomyces cerevisiae* IcmT (Ste14p; 239 residues), *Tribolium castaneum* IcmT (beetle, 281 residues), the IcmT found in *Anopheles gambiae* (Ag-IcmT; 283 residues), and the prokaryotic enzyme from *Methanosarcina acetivorans* (Ma-IcmT; 194 residues). UniProt accession numbers for the IcmT ortholog sequences are as follows: O60725, P32584, D6WJ77, Q8TMG0, and Q7PXA7, respectively. Alignment was made using Clustal Omega (56) and Jalview 2.10.5 (57). Colored boxes denote residues with similar properties: glycine is orange, acidic amino acids are purple, proline is yellow, nonpolar amino acids are light blue, histidine and tyrosine are turquoise, basic amino acids are red, and remaining polar, neutrally charged amino acids are green. The red dashed box surrounds the GXXXG motif in Ste14.



Supporting Figure 2: Topology of *Saccharomyces cerevisiae* Icm1, Ste14 with highlighted location of dimer motif in TM1 (15). Ste14 contains a motif of interest for dimerization, SXXXGXXXGXXXG, consisting of residues 27 to 39. The glycine and serine residues allow for helix-helix interaction between monomer Ste14 protein to form hydrogen bonds without causing steric hindrance. Light blue residues represent the glycine or serine amino acids in the repeating GXXXG motif, while purple residues represent the “X” amino acids. Pink residues indicate the remainder of the TM1 residues in Ste14, residues 25 to 43, utilized in experiments later described.

Supporting Table 1: SAXS structural parameters of Ste14

| Pure Ste14 | |
|---------------------------------|---|
| Guinier analysis | |
| $I(0)$ (cm ⁻¹) | $7.32\text{e-}10^{-5} \pm 3.97\text{e-}10^{-7}$ |
| R_g (Å) | 37.6 ± 0.303 |
| q min (Å ⁻¹) | 0.0122 |
| q range (Å ⁻¹) | 0.0122 - 0.0278 |
| r^2 (linear fit) | 0.9284 |
| P(r) analysis | |
| $I(0)$ (cm ⁻¹) | $7.94\text{e-}10^{-5} \pm 6.96\text{e-}10^{-7}$ |
| R_g (Å) | 52.3 |
| D_{max} (Å) | 104 |
| Porod Volume (Å ⁻³) | 52,100 |
| q range (Å ⁻¹) | 0.0122 - 0.0313 |

Methyltransferase-like 21e inhibits 26S proteasome activity to facilitate hypertrophy of type IIb myofibers

Chao Wang,^{*,1} Bin Zhang,[†] Anna C. Ratliff,[‡] Justine Arrington,[‡] Jingjuan Chen,^{*} Yan Xiong,[†] Feng Yue,^{*} Yaohui Nie,^{*} Keping Hu,[‡] Wen Jin,[†] W. Andy Tao,^{§,¶} Christine A. Hrycyna,^{‡,¶} Xiaobo Sun,[‡] and Shihuan Kuang^{*,¶,2}

^{*}Department of Animal Sciences, [†]Department of Chemistry, [‡]Department of Biochemistry, and [¶]Center for Cancer Research, Purdue University, West Lafayette, Indiana, USA; and [§]Institute of Medicinal Plant Development, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China

ABSTRACT: Skeletal muscles contain heterogeneous myofibers that are different in size and contractile speed, with type IIb myofiber being the largest and fastest. Here, we identify methyltransferase-like 21e (Mettl21e), a member of newly classified nonhistone methyltransferases, as a gene enriched in type IIb myofibers. The expression of *Mettl21e* was strikingly up-regulated in hypertrophic muscles and during myogenic differentiation *in vitro* and *in vivo*. Knockdown (KD) of *Mettl21e* led to atrophy of cultured myotubes, and targeted mutation of *Mettl21e* in mice reduced the size of IIb myofibers without affecting the composition of myofiber types. Mass spectrometry and methyltransferase assay revealed that Mettl21e methylated valosin-containing protein (Vcp/p97), a key component of the ubiquitin-proteasome system. KD or knockout of Mettl21e resulted in elevated 26S proteasome activity, and inhibition of proteasome activity prevented atrophy of *Mettl21e* KD myotubes. These results demonstrate that Mettl21e functions to maintain myofiber size through inhibiting proteasome-mediated protein degradation.—Wang, C., Zhang, B., Ratliff, A. C., Arrington, J., Chen, J., Xiong, Y., Yue, F., Nie, Y., Hu, K., Jin, W., Tao, W. A., Hrycyna, C. A., Sun, X., Kuang, S. Methyltransferase-like 21e inhibits 26S proteasome activity to facilitate hypertrophy of type IIb myofibers. *FASEB J.* 33, 9672–9684 (2019). www.fasebj.org

KEY WORDS: Mettl21e · Myostatin · non-histone methylation · muscle atrophy · VCP/p97

Mammalian skeletal muscles are mainly composed of mature muscle cells (myofibers) that are heterogeneous in the expressions of myosin heavy chain (Myh) isoforms (1). Based on the expression of Myh7, Myh2, Myh1, or Myh4, rodent limb muscle myofibers are classified as type I, IIa, IIx/d, or IIb, respectively. These 4 types of myofibers also exhibit distinct physiologic and metabolic characteristics (2). The contractile speed of myofibers ranks in the order of IIb > IIx > IIa > I, whereas the fatigue resistance of myofibers is negatively correlated with the contractile speed (2). Biochemically and metabolically, I and IIa myofibers have high oxidative capacity with a large number of

mitochondria. In contrast, IIx and IIb myofibers mainly utilize glycolysis to generate ATP. Besides the physiologic and metabolic differences, the size of myofibers are generally in the order of IIb > IIx > IIa > I (3). However, it is not clear why IIb myofibers have the largest size compared with the other types of myofibers.

The identity of a myofiber is controlled by several factors, including motor neuron innervation, hormonal regulation, progenitor cell origin, and myofiber type-specific gene expression (2). Each mature myofiber is monoinnervated by a single motor neuron, which transmits electrical impulses to a subset of myofibers (4). Firing of slow motor neurons (motor neurons innervating slow myofibers) generates sustained and low-amplitude elevations in intracellular calcium concentrations, leading to the activation of calcineurin (5). In contrast, fast motor neuron firing leads to brief calcium transients of high amplitude, which are insufficient to activate calcineurin (5). Activated calcineurin activates Nuclear factor of activated T-cells (NFAT) and myocyte enhancer factor-2 (MEF2) transcription factors to promote transcription of slow myofiber-related genes (6–9). Besides motor neurons, Twist Family BHLH Transcription Factor 2 (Twist2)-dependent progenitor cells are recently shown to maintain IIb myofibers (10). In addition, myofiber-type specific

ABBREVIATIONS: EDL, extensor digitorum longus; FLAG, DYKDDDDK epitope; GA, gastrocnemius; KO, knockout; KD, knockdown; Mettl21e, methyltransferase-like 21e; Mstn, myostatin; Myh, myosin heavy chain; SAM, S-adenosyl-L-methionine; sh, short hairpin; SOL, soleus; TA, tibialis anterior; TALEN, transcription activator-like effector nuclease; TCA, trichloroacetic acid; Vcp, valosin-containing protein; WT, wild type

¹ Correspondence: Purdue University, 174B Smith Hall, 901 West State St., West Lafayette, IN 47907, USA. E-mail: wang1438@purdue.edu

² Correspondence: Purdue University, 174B Smith Hall, 901 West State St., West Lafayette, IN 47907, USA. E-mail: skuang@purdue.edu

doi: 10.1096/fj.201900582R

This article includes supplemental data. Please visit <http://www.fasebj.org> to obtain this information.

genes or microRNAs are shown to contribute to myofiber specifications, such as type I-specific ephrin-A3 and microRNA-208/microRNA-499 (11, 12), oxidative myofiber-enriched PGC1- α (13), glycolytic myofiber-specific T-Box 15 (14), and fast myofiber-enriched SRY-Box 6 (15). However, gene programs that regulate subtypes of fast myofibers are poorly understood.

Methyltransferases are enzymes that transfer a methyl group (CH₃) from a donor, generally S-adenosyl-L-methionine (SAM), to lysine and arginine residues of substrate proteins (16). Methylation of histone proteins are often related to changes in chromatin state and gene expression, whereas methylation of non-histone proteins may modify their activity, stability, and molecular interactions (17). In humans, a group of methyltransferase-like 21 (METTL21) proteins (A–D) has recently been identified to mediate lysine methylation of molecular chaperones and eukaryotic translation elongation factor 1A (eEF1A) (18–22). Specifically, METTL21A (HSPA lysine methyltransferase, or HSPA-KMT) methylates several heat shock protein 70 (HSPA/Hsp70) proteins (20), METTL21B methylates eEF1A at Lys¹⁶⁵ (K165) (21, 22), and METTL21C and METTL21D (VCP-KMT) methylate the ATP-dependent chaperone VCP at Lys³¹⁵ (K315) (18, 19, 23). In addition, a yeast Mettl21-like protein is shown to methylate eEF1A (24), suggesting a potential role of Mettl21 family methyltransferases in protein synthesis, a pivotal process for muscle hypertrophy.

In searching for myofiber type-specific regulators, we analyzed microarray and RNA-sequencing data comparing gene expression in the skeletal muscles of control and myostatin (Mstn) knockout (KO) mice, which exhibit robust myofiber hypertrophy and increased abundance of IIb myofibers (25, 26). This analysis led to the identification of *Mettl21e* gene that was among the top up-regulated genes in Mstn^{KO} muscles. Intriguingly, the expression of *Mettl21e* is also increased in the Callipyge sheep model of muscle hypertrophy and in response to β -adrenergic agonist-induced muscle protein accretion in lambs (27, 28), suggesting a conserved link between Mettl21e and muscle hypertrophy in multiple species. In this study, we used loss-of-function assays in cell culture and in a novel mouse model to investigate the function of Mettl21e *in vivo*. Our results demonstrate that Mettl21e is a type IIb myofiber-enriched protein that functions to maintain myofiber size through suppression of 26S proteasome activity.

MATERIALS AND METHODS

Mice

Constitutive Mstn KO mice were generated by Dr. S. E. Jin-Lee (Johns Hopkins University, Baltimore, MD, USA) (29). Heterozygous Mstn^{+/-} mice were bred to generate Mstn KO and wild-type (WT) littermates. Unless otherwise indicated, we used 2-mo-old mice for experiments. All procedures involving the use of animals were performed in accordance with the guidelines

presented by Purdue University's Animal Care and Use Committee.

Generation of Transcription activator-like effector nuclease-mediated *Mettl21e* KO mice

The Transcription activator-like effector nuclease (TALEN)-mediated Mettl21e KO mice were produced by Beijing View-Solid Biotechnology (Beijing, China). The TALEN plasmid pCS5-eTALEN-T was designed to induce *Mettl21e* frameshift mutation. TALEN-left targets the sequence (5'-GGTCGCAGATCATGG-3') of the sense strand, and TALEN-right targets the sequence (5'-AGTCGTTATCAGAGTTG-3') of the antisense strand. Mutated mice were generated by pronuclear injection using standard methods. Founder mice were screened for the presence of mutation by sequencing the PCR products amplified by the primers for *Mettl21e*-sens (5'-ATGGACCTCACAGTA-CTCACAT-3') and *Mettl21e*-anti (5'-GCTTGCACAAATG-GAGACAAG-3'). Mutant mice were mated with WT C57BL/6 mice to obtain heterozygous Mettl21e^{flie} mice. Mettl21e^{flie} mice were then mated at least 3 generations to obtain Mettl21e^{KO} and WT mice.

Primary myoblast isolation, culture, and differentiation

Primary myoblasts were isolated from 5-wk-old WT female mice. The hind limb skeletal muscles were minced and digested in type I collagenase and dispase B mixture (Roche Applied Science, Penzberg, Germany). The digestion was stopped with F-10 Ham's medium containing 17% fetal bovine serum, and the cells were filtered from debris, centrifuged, and cultured in growth medium (F-10 Ham's medium supplemented with 17% fetal bovine serum, 4 ng/ml basic fibroblast growth factor, and 1% penicillin-streptomycin) on uncoated dishes for 3 d when 5 ml growth medium was added each day. Then, the supernatant were collected, centrifuged, and trypsinized with 0.25% trypsin. After washing off the trypsin, primary myoblasts were seeded on collagen-coated dishes, and the growth medium was changed every 2 d. Myoblasts were induced to differentiate on matrigel-coated dishes and cultured in differentiation medium (DMEM supplemented with 2% horse serum and 1% penicillin-streptomycin). Differentiation medium was replaced every day. Primary myoblasts were cultured in normal humidified tissue culture incubators with 5% CO₂. To inhibit Mstn signaling, cells were treated with 200 ng/ml follistatin (SRP3045; MilliporeSigma, Burlington, MA, USA). To inhibit proteasome activity, differentiated myoblasts were treated with 25 nM bortezomib (B-1408; LC Laboratories, Woburn, MA, USA).

Immunostaining and image acquisition

Muscle samples were processed following the protocol described by Wang *et al.* (30). Myoblasts and myotubes were fixed with 4% paraformaldehyde and then blocked with blocking buffer (5% goat serum, 2% bovine serum albumin, 0.2% Triton X-100, and 0.1% sodium azide in PBS) for at least 1 h. Then, the samples were incubated with primary antibodies [1:200 in blocking buffer; sarcomeric myosin heavy chain antibody (clone MF20) was from Developmental Studies Hybridoma Bank (University of Iowa, Iowa City, Iowa, USA), and anti-DYKDDDDK epitope (FLAG) antibody (F1804) was from MilliporeSigma] overnight. After washing with PBS, the samples were incubated with respective secondary antibodies and DAPI for 45 min at room temperature. Fluorescent images were captured using a Leica DM 6000B fluorescent microscope (Leica Microsystems, Wetzlar, Germany).

RNA extraction and real-time quantitative PCR

Total RNA of muscles or myoblasts were extracted using Trizol Reagent (15596-018; Thermo Fisher Scientific, Waltham, MA, USA). RNA was treated with RNase-free DNase I (AM2224; Thermo Fisher Scientific) to remove genomic DNA. The purity and concentration of total RNA were measured by Nanodrop 3000 (Thermo Fisher Scientific). Random primers and Moloney murine leukemia virus reverse transcriptase were used to convert RNA into cDNA. Real-time PCR was performed using Roche Lightcycler 480 PCR System with SYBR Green Master Mix (04707516001; Roche Applied Science). Primers used were listed in Supplemental Table S1 and ref. 31. C_t value of 18S rRNA was used as internal control, and $2^{-\Delta\Delta C_t}$ method was used to analyze the relative mRNA expression of various genes.

Single myofiber isolation

Extensor digitorum longus (EDL) and soleus (SOL) muscles were removed carefully and digested with 2 mg/ml collagenase type 1 (CLS-1; Worthington Biochemical, Lakewood, NJ, USA) in DMEM (MilliporeSigma) for 45 min at 37°C. Digestion was stopped by carefully transferring EDL or SOL muscles to a horse serum-coated Petri dish (60-mm) with DMEM. Myofibers were released by gently flushing muscles with a large bore glass pipette. The released single myofiber was washed in PBS and then transferred to a 0.2-ml PCR tube. The residual PBS was removed from the PCR tube. RNA of a single myofiber was extracted using a PicoPure RNA Isolation Kit (KIT0204; Thermo Fisher Scientific) according to the manufacturer's protocol. Generally, we pipetted 50 μ l extraction buffer into the PCR tube containing a single myofiber and incubated for 30 min at 42°C to extract cellular contents. Then, we gently mixed 50 μ l 70% ethanol into the cell extracts, and the mixture was added into preconditioned purification column. After 2 rounds of wash, RNA was eluted with 11 μ l elution buffer. The eluted RNA was used directly in reverse transcription to generate cDNA for real-time PCR analyses.

Protein extraction and Western blot analysis

Muscle samples and cultured myoblasts were washed with PBS and homogenized with radioimmune precipitation assay buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS). Protein concentrations were determined using Pierce BCA Protein Assay Reagent (Pierce Biotechnology, Rockford, IL, USA). Proteins (100 μ g) were separated by 10% SDS-PAGE, electrotransferred onto PVDF membrane (MilliporeSigma), and incubated with specific primary antibodies. Glyceraldehyde 3-phosphate dehydrogenase (6C5) antibodies (1:1000 in 5% w/v nonfat dry milk) were from Santa Cruz Biotechnology (Dallas, TX, USA), and MF20 (1:100 in 5% w/v nonfat dry milk) was from Developmental Studies Hybridoma Bank. Immunodetection was performed using ECL Western blotting substrate (Pierce Biotechnology) and detected with FluorChem R imaging system (ProteinSimple, San Jose, CA, USA).

Adenovirus generation

The adenoviruses with short hairpin (sh)-RNAs, *Mstn*, or *Mettl21e-FLAG* were generated using the AdEasy system (240009; Agilent, Santa Clara, CA, USA). We subcloned the *U6-shRNA* cassettes from pLKO.1-*U6-shRNA* plasmids (Sh1: TRCN0000176763; Sh2: TRCN0000177496; Sh3: TRCN0000177877; Sh4: TRCN0000176682; MilliporeSigma) with primers (pLKO.1-f/r, Supplemental Table S1). *Mstn* ORF and *Mettl21e-FLAG* ORF were cloned with primers

(*Mstn*-f/r for *Mstn*; *Mettl21e*-f/*Mettl21e-FLAG*-r for *Mettl21e-FLAG*; Supplemental Table S1). These cloned DNA sequences were inserted into pAdTrack-CMV plasmid (16405; Addgene, Watertown, MA, USA) (The cloned *Mettl21e-FLAG* ORF was also inserted into pcDNA3.1 for intracellular location detection) and then were digested by PmeI and transfected with the DH5a competent cell with pAdEasy-1. The following steps were exactly according to the methods generating the *MyoD*-overexpression adenovirus (31).

RNA sequencing

RNA was extracted from tibialis anterior (TA) muscles of three 8-wk-old WT and *Mettl21e*^{KO} mice. Sequencing libraries were generated using NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's recommendations. RNA was sequenced and analyzed by Novogene (Beijing, China) using the Illumina HiSeq4000 platform (Illumina, San Diego, CA, USA). The original data of RNA sequencing are deposited to the Gene Expression Omnibus (GEO) data set (GSE122024, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122024>; Token: qrqjeqcnfmvxct).

Immunoprecipitation to pull down *Mettl21e* binding complex

Vectors containing *Mettl21e-FLAG* or *GFP-FLAG* were transduced into primary myoblasts (70–80% confluent) by adenovirus. Two days posttransduction, cells were induced to differentiation for 3 d and then were scraped with ice-cold PBS (from 10 \times 100-mm plates for each plasmid) and were centrifuged. The cell pellet was completely resuspended with 1 ml lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 1% NP40 with 1 time Protease inhibitor cocktail) on ice and sonicated with a 1-s pulse at 5-s intervals 10 times. After 5 min on ice, cell lysate was centrifuged with 14,000 rpm at 4°C, and then the supernatant (containing 5 mg proteins determined by the Pierce bicinchoninic acid assay) was transferred to 50 μ l pre-cleaned anti-FLAG magnetic beads slurry (M8823; MilliporeSigma). After incubation on rotator for 3 h in a cold room (4°C), the magnetic beads were washed 3 times with lysis buffer. Then, we added 500 μ l ddH₂O, and pipetted up and down several times to remove salt or solvent remainder (twice). Bead-captured proteins were eluted with 100 μ l of 50 mM triethyl amine and 5 mM DIT on a thermal shaker (99°C, 5 min). The protein elution was centrifuged with CentrVap Concentrator (Labconco, Kansas City, MO, USA) to partially remove triethyl amine, adjusted to 100 μ l and pH 8.0 with 1% acetic acid and 15 mM Iodoacetamide, and placed in the dark for 1 h. Before desalting the protein elution with C18 Zip tips (NT3C18.96; Glygen, Columbia, MD, USA), 1 μ g trypsin was reacted with the 100 μ l adjusted protein elution for 16 h at 37°C. The following steps for mass spectrometric data acquisition and analysis were according to the descriptions by Wang *et al.* (31).

Recombinant protein production and purification

pETDuet-1-derived plasmids incorporating *His₆-VCP*, *His₆-Mettl21e*, or *His₆-Hsp90ab1* were transformed into the *Escherichia coli* expression strain BL21 (DE3; Thermo Fisher Scientific). Cells were cultured in LB medium with 0.1 mg/ml ampicillin at 37°C in a shaking incubator at 220 rpm until the absorbance at 600 nm reached 0.6 optical density (OD). The culture was induced with 100 μ M isopropyl β -D-thiogalactoside (Gold Biotechnology, St. Louis, MO, USA), and the temperature was lowered to 18°C for 18 h. Cells were harvested by centrifugation at 7000 g. Cell pellets were resuspended in lysis buffer [50 mM Tris (pH 7.5), 500 mM

NaCl, 10% (w/v) glycerol, 30 mM imidazole, 3 mM 2-ME, 0.5% Triton X-100, 1 tablet cComplete EDTA-free Protease Inhibitor Cocktail (Roche, Basel, Switzerland), and 2 mM 4-benzene-sulfonyl fluoride hydrochloride (Gold Biotechnology)]. Cells were lysed *via* probe sonication and underwent centrifugation at 100,000 *g* for 1 h. The soluble protein in the supernatant was rocked with Ni-NTA resin (Thermo Fisher Scientific) at 4°C. After 1 h, resin was washed with buffer (50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 10% glycerol, and 30 mM imidazole) followed by an additional wash with the addition of 0.5 M KCl. Recombinant proteins were removed from resin with elution buffer (50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 300 mM imidazole). The eluted protein was concentrated in Amicon Ultra MWCO concentrators columns (MilliporeSigma) with appropriate molecular mass cutoffs. Proteins were divided into aliquots and stored at -80°C and were thawed on ice prior to use in assays. Concentrations of each protein were determined *via* Bradford protein assay (Thermo Fisher Scientific). Purity was confirmed by loading purified protein (1 µg) on 10% SDS-PAGE and either stained with Coomassie [0.25% (w/v) Coomassie Brilliant Blue R-250 (20278; Thermo Fisher Scientific), 80% methanol, and 20% acetic acid] or transferred to a nitrocellulose membrane (0.22 µm; GE Healthcare, Waukesha, WI, USA). The nitrocellulose membrane was blocked at 4°C overnight in 20% (w/v) nonfat dry milk in PBS with Tween 20 [137 mM NaCl, 2.7 mM KCl, 4 mM Na₂HPO₄, 1.8 mM KH₂PO₄ and 0.05% (v/v) Tween-20, pH 7.4]. The blocked membrane was incubated for 1 h at room temperature with α-His-hrp (1:10,000; MilliporeSigma) antibody in 5% (w/v) nonfat dry milk in PBS with Tween 20. The membranes were washed 3 times with PBS with Tween 20, and the protein bands were visualized using ECL (Pierce Biotechnology).

In vitro methyltransferase reaction

Methyltransferase reactions were performed in 50-µl volumes for 1 h at 37°C in methyltransferase reaction buffer [50 mM Tris (pH 7.5), 50 mM KCl, 5 mM MgCl₂, 1 mM ATP], 13 µM [¹⁴C] SAM (2 µCi), 1 µM methyltransferase enzyme or 4 µM substrates, and varying concentrations of substrates or methyltransferase enzyme. The reactions were stopped by precipitating proteins with 50-µl 10% (v/v) Trichloroacetic acid (TCA) at 4°C for 1 h. The reactions were spotted onto glass fiber filters (45 µm; Whatman, Maidstone, United Kingdom), and the acid-insoluble material was retained during vacuum filtration. The reaction tubes were rinsed with an additional 50 µl 10% (v/v) TCA and applied to the filters. The filters were then washed with 1 ml of 10% (v/v) TCA followed by 1 ml 100% ethanol and left to dry for 10 min at room temperature. The dried filters were placed in scintillation vials with 10 ml Bio-Safe II biodegradable scintillation cocktail (Research Products International, Mt Prospect, IL, USA), and radioactivity was measured by scintillation counting.

Proteasome activity assay

TA muscles or differentiated myoblasts were homogenized with 0.5% NP-40. Then, the tissue or cell lysates were centrifuged at 4°C at 13,000 rpm for 10 min. The supernatants were collected and measured using a Proteasome Activity Assay Kit (ab107921; Abcam, Cambridge, MA, USA) according to the manufacturer's protocol.

Statistical analysis

The data were presented with mean and SD. *P* values were calculated using unpaired 2-tailed Student's *t* test for 2 group

comparison and 1-way ANOVA for multiple group comparison. Values of *P* < 0.05 were considered to be statistically significant.

RESULTS

Mettl21e expression is elevated in hypertrophic Mstn KO muscles

We analyzed microarray data comparing gene expression of gastrocnemius (GA) muscles from WT and Mstn conditional KO mice (25). Three months after Cre-mediated deletion of Mstn in adult mice, the expression of *Mettl21e* in GA muscles was 12-fold higher in Mstn KO mice than in WT mice (Fig. 1A). Consistently, a higher level of *Mettl21e* expression was detected in the TA muscles of constitutive Mstn KO mice relative to the same muscles of WT mice (Fig. 1B). These results suggest that the expression of *Mettl21e* is negatively correlated with Mstn but positively associated with muscle hypertrophy.

To determine whether *Mettl21e* expression is directly regulated by Mstn downstream signaling, we examined acute responses to perturbations of Mstn signaling. We first treated primary myoblasts with an Mstn inhibitor (follistatin) in differentiation medium for 5 d. The treatment did not alter the expression of *Mettl21e* (Fig. 1C). We also used adenoviral vectors to overexpress Mstn or GFP (control) in primary myoblasts cultured in growth medium. Overexpression of Mstn did not affect the expression of *Mettl21e* (Fig. 1D). We further analyzed microarray data comparing gene expression of GA muscles of mice injected with PBS or an Mstn antibody (clone JA-16 or PF-354) for 4 d, at which time point muscle hypertrophy did not occur (32). The short-term injection of JA16 did not change the expression of *Mettl21e* (Fig. 1E), though it blocked Mstn signaling (32). These results suggest that *Mettl21e* expression is associated with muscle hypertrophy rather than directly regulated by Mstn downstream signaling.

Muscle hypertrophy is often associated with changes in myofiber type distribution. We found that there was an apparent increase in the abundance of IIb myofibers in the Mstn KO EDL muscles relative to WT muscles (Fig. 1F), accompanied by an increased mRNA level of *Myh4* (Fig. 1G), which encodes the IIb isoform Myh. Interestingly, the mRNA levels of *Mettl21e* were also evaluated by 8-fold in Mstn KO compared with WT EDL muscles (Fig. 1H). In contrast, *Mettl21e* levels in the SOL muscle that lacks IIb myofibers were comparable in Mstn KO and WT mice (Fig. 1I–K). These results indicate that the up-regulation of *Mettl21e* in the Mstn KO muscles is correlated to an increase in IIb myofibers.

Mettl21e is enriched in type IIb myofibers

We next directly determined the expression of *Mettl21e* in different types of myofibers. We used *Mettl21c* as a control, which is specifically expressed in type I myofibers (23). We found that both *Mettl21c* and *Mettl21e* were primarily expressed in the skeletal muscle of mice (Fig. 2A). Within various muscle groups, the levels of *Mettl21c* were found

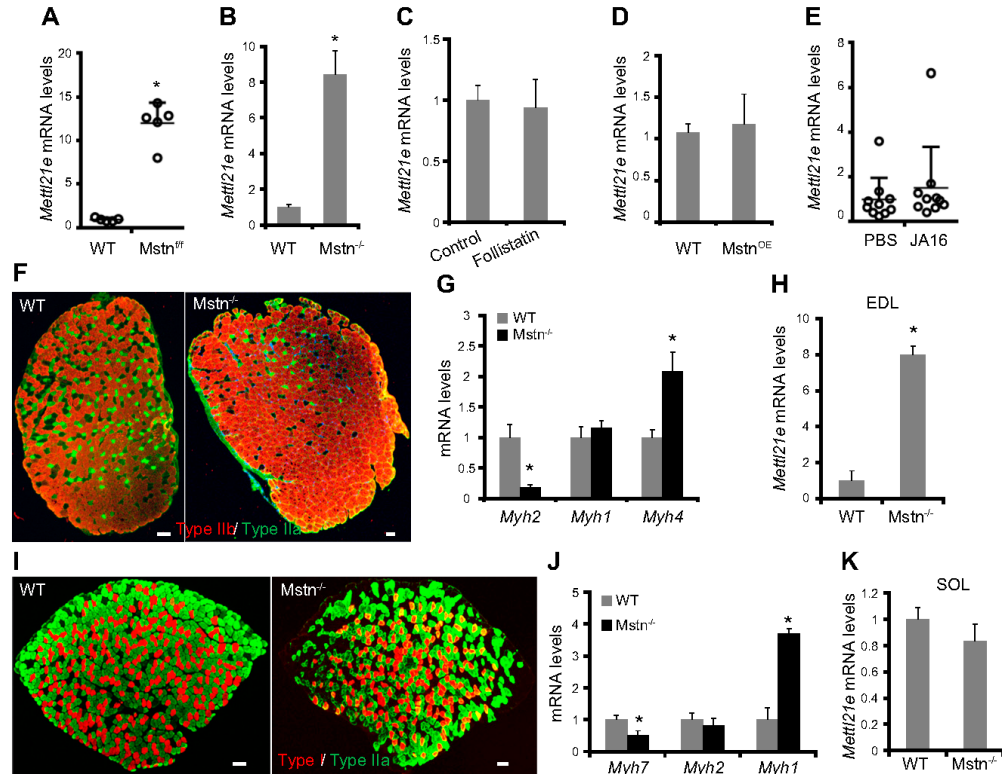


Figure 1. The expression of *Mettl21e* is associated with *Mstn* KO in fast muscles. *A*) *Mettl21e* expression in GA muscles from WT and postdevelopmental *Mstn* KO mice. *B*) *Mettl21e* expression in TA muscles from WT and constitutive *Mstn* KO mice. *C*) *Mettl21e* expression in differentiated myoblasts treated with follistatin. *D*) *Mettl21e* expression in myoblasts with *Mstn* overexpression. *E*) *Mettl21e* expression in GA muscles with anti-*Mstn* antibody JA16 or PBS. *F*) Immunostaining of type IIb (Myh4, Red) and type IIa (Myh2, Green) myofibers in EDL muscles of WT and *Mstn* KO mice. *G*, *H*) qPCR analysis of *Myh* genes (*G*) and *Mettl21e* (*H*) in EDL muscles of WT and *Mstn* KO mice. Scale bars, 100 μ m. *I*) Immunostaining of type I (Myh7, Red) and type IIa (Myh2, Green) myofibers in SOL muscles of WT and *Mstn* KO mice. Scale bars, 100 μ m. *J*, *K*) qPCR analysis of *Myh* genes (*J*) and *Mettl21e* (*K*) in SOL muscles of WT and *Mstn* KO mice. The values in *A* and *E* were originated from the published microarray data. Error bars represent mean \pm SD of 6 mice (*B*, *G*, *H*, *J*, *K*) or 5 independent biologic experiments (*C*, *D*). * P < 0.05 (Student's *t* test).

to be SOL > GA > TA and EDL (Fig. 2B), a pattern consistent with the relative abundance of type I myofibers in these muscles. In contrast, the expression of *Mettl21e* was much higher in TA, EDL, and GA muscles than in SOL muscles, which lack type IIb myofibers (Fig. 2B). To directly examine the expression of *Mettl21e* in various types of myofibers, we isolated individual myofibers from SOL and EDL muscles and performed quantitative (q) PCR analysis on single myofibers. All individual myofibers predominantly express one of *Myh* isoforms, and based on the relative levels of *Myh7*, *Myh2*, *Myh1*, and *Myh4*, individual myofibers can be readily grouped into type I (blue column), IIa (red column), IIx (green column), or IIb (gray column) myofibers (Fig. 2C). Interestingly, all IIb myofibers also expressed a medium level of *Myh1* (Fig. 2C). Strikingly, whereas *Mettl21c* expression was restricted to type I myofibers, *Mettl21e* expression was mainly

enriched in IIb myofibers (Fig. 2D). On average, the mRNA level of *Mettl21c* was 12-fold in type I myofibers as in the other types (IIa, IIx, IIb) of myofibers, and *Mettl21e* expression was about 7-fold in type IIb myofibers as in the other types (I, IIa, IIx) of myofibers (Fig. 2D, inset). The single-cell-analysis results demonstrate that *Mettl21e* is enriched in type IIb myofibers.

Mettl21e regulates myotube size in myoblast culture

Mettl21e expression was up-regulated by 37-fold at 3 d postinduction of differentiation in primary myoblasts (Fig. 3A), suggesting it mainly functions in differentiated myotubes. We thus performed shRNA to knock down (KD) *Mettl21e* in myotubes. Four adenoviral shRNA-GFP vectors (Sh1-Sh4) were constructed; 3 of them target coding DNA

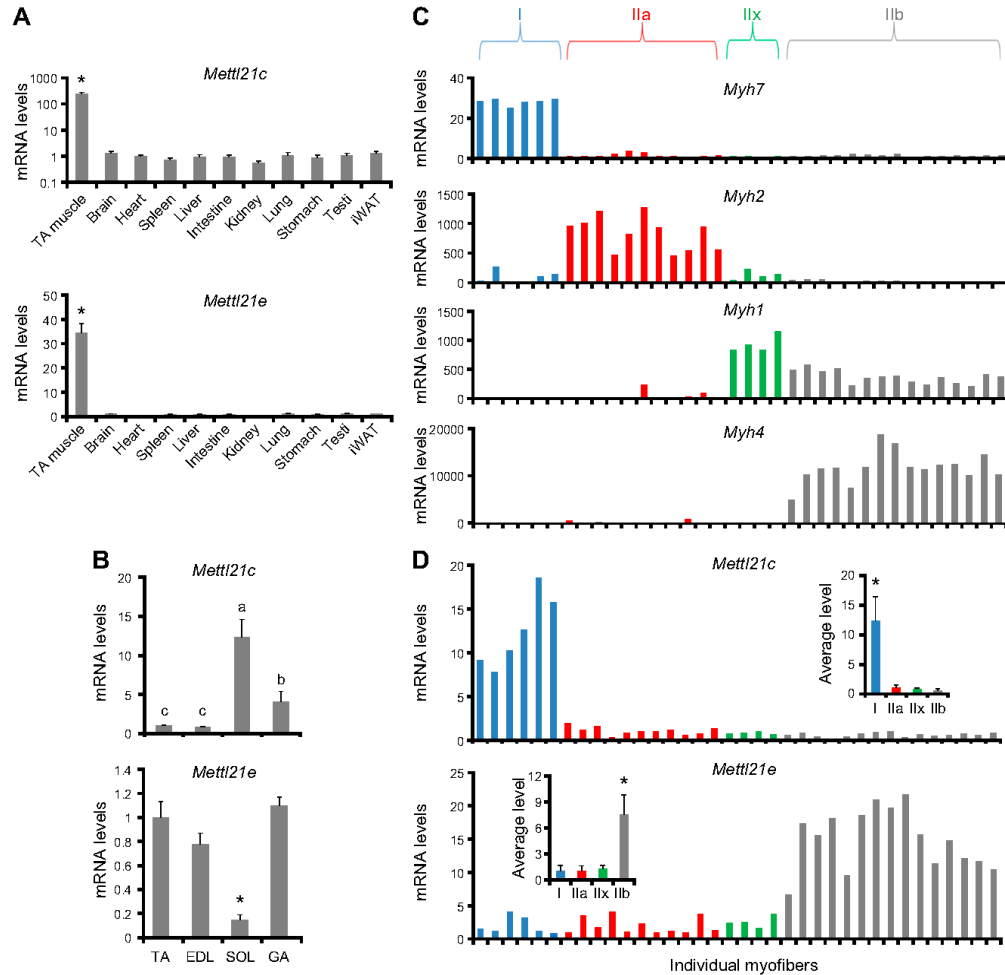


Figure 2. The expression *Mettl21e* in different types of myofibers. **A**) qPCR showing relative mRNA levels of *Mettl21c* and *Mettl21e* in different tissues. **B**) qPCR analysis of *Mettl21c* and *Mettl21e* in fast-twitch (TA, EDL, and GA) and slow-twitch (SOL) muscles; $n = 6$ independent biologic experiments with 3 technical repeats. **C**) qPCR analysis of Myh genes in isolated myofibers from SOL and EDL muscles of 4 mice. Individual myofibers were grouped into type I (blue), IIa (red), IIx (green), and IIb (gray). **D**) qPCR analysis of *Mettl21c* and *Mettl21e* in individual myofibers. Inset: the average expression of *Mettl21c* and *Mettl21e* in different types of myofibers. Error bars represent mean + sd. * $P < 0.05$ (1-way ANOVA).

sequence, and 1 targets 3'UTR (Fig. 3B). Sh1–Sh4 reduced the expression of *Mettl21e* by 99%, 0%, 52%, and 92%, respectively (Fig. 3B). Sh1 and Sh4 were thus chosen to KD *Mettl21e*, and Sh2 was used as a negative control. Myoblasts were transduced with the Sh1, Sh2, Sh4, or control (GFP only) adenoviral vectors for 2 d and then induced to differentiate for 3 d. As expected, myotubes treated with Sh2 were morphologically similar to control myotubes (Fig. 3C). In contrast, myotubes treated with Sh1 or Sh4 were thinner than control myotubes (Fig. 3C). Quantitative analysis showed that *Mettl21e*^{KD} decreased the fusion index and the diameter of myotubes (Fig. 3D, E). *Mettl21e*^{KD} myotubes also expressed lower levels of Myh than did the control

myotubes (Fig. 3F). Our cell culture data suggest a function of *Mettl21e* in regulating myofiber size.

Mutation of *Mettl21e* decreases the size of type IIb myofibers in mice

We then generated *Mettl21e* KO mice using TALEN technique, creating a single nucleotide deletion in the 5' region of the *Mettl21e* gene (Fig. 4A). This frameshift mutation generated 4 mutated amino acids (³⁵FIPT³⁸ to ³⁵LSQL³⁸) and a stop codon (TGA), leading to a truncated *Mettl21e* protein with only 38 amino acids (Fig. 4B). The truncated *Mettl21e* protein does not contain the functional domain of

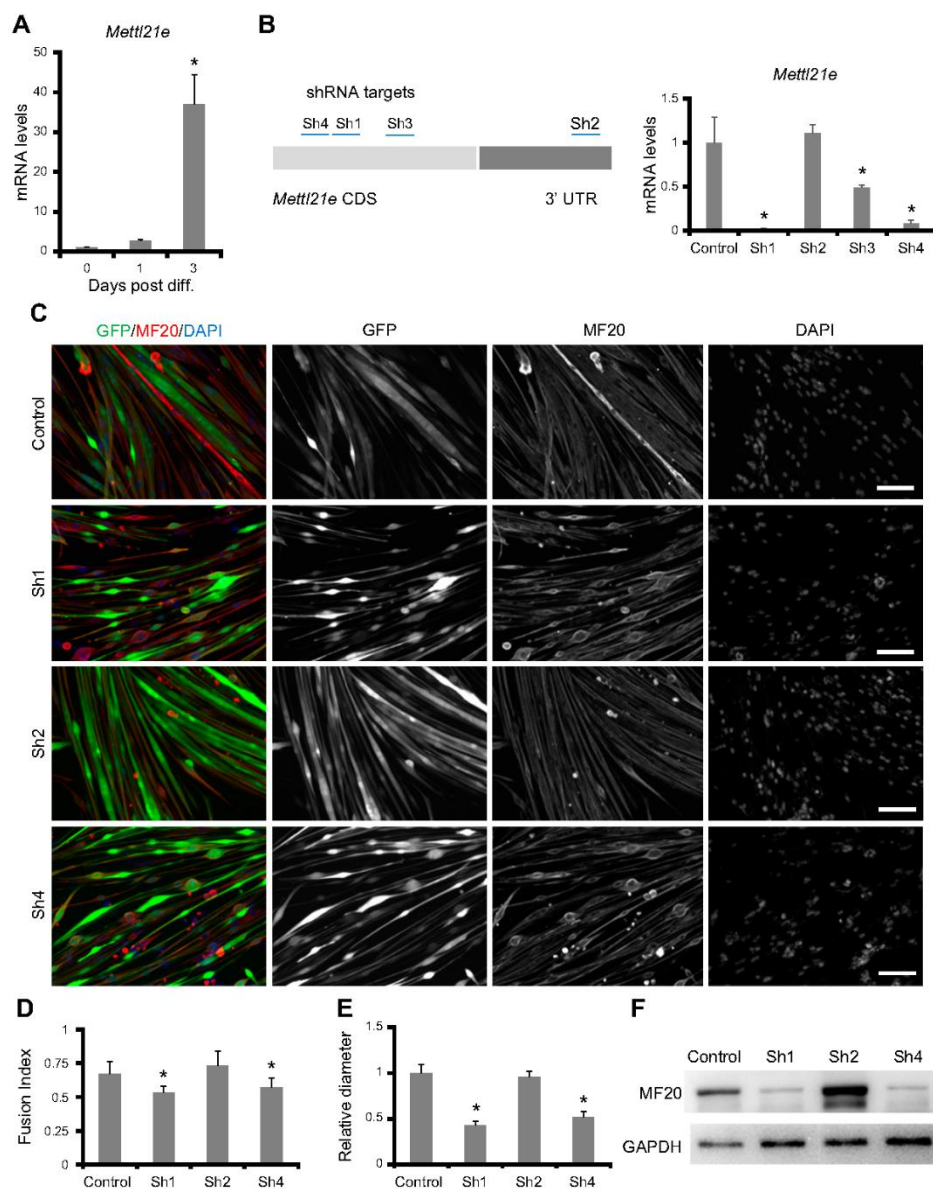


Figure 3. Knockdown (KD) of *Mettl21e* induces myotube atrophy *in vitro*. *A*) Relative levels of *Mettl21e* in myoblasts during differentiation; $n = 5$ independent biologic experiments with 3 technical repeats. *B*) Relative levels of *Mettl21e* in myotubes treated with the adenoviral shRNAs; $n = 5$ independent biologic experiments with 3 technical repeats. *C*) Myotubes treated with shRNAs were stained by MF20 (red), and nuclei were counterstained by DAPI (blue). Scale bar, 100 μ m. Myoblasts were incubated with adenoviruses for 1 d and cultured in virus-free growth medium for 1 more day before differentiating for 2 d. *D*, *E*) Fusion index (*D*) and relative diameter (*E*) quantified based on staining in *C*. Fusion index measures the percentage of myonuclei present in multinucleated myofibers. Only GFP⁺ cells were used for quantification; $n = 5$ independent biologic experiments, with 5 different areas analyzed in each experiment. *F*) Western blot analysis showing the effect of *Mettl21e* KD on Myh protein expression. Error bars represent mean \pm SD. * $P < 0.05$ (Student's *t* test).

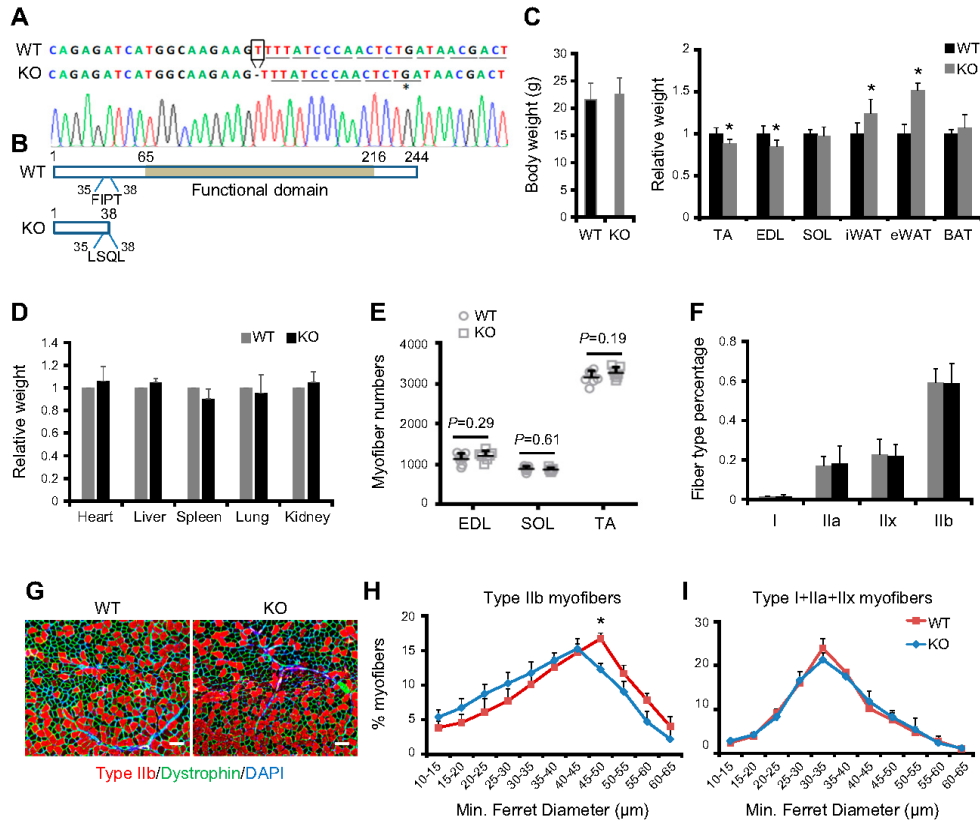


Figure 4. Loss of *Mettl21e* decreases type IIb myofiber size. **A**) Sequencing chromatograms of targeted region of *Mettl21e* gene. Reading frame is underlined. The stop codon is highlighted by line and dot in mutated *Mettl21e* gene. **B**) Predicted *Mettl21e* protein in WT and *Mettl21e*^{KO} mice. Numbers indicate the sites of amino acids in annotated protein. Conserved functional domain of *Mettl21e* is in gray box. **C, D**) Whole body weights (**C**) and relative weights (**D**) of different tissues; *n* = 7 pairs of mice. **E**) Myofiber numbers in skeletal muscles of WT and *Mettl21e*^{KO} mice; *n* = 7 pairs of mice. **F**) Quantification of myofiber types in EDL muscles. **G**) Immunostaining of type IIb (Red) myofibers in WT and *Mettl21e*^{KO} mice. Scale bars, 100 μ m. **H, I**) Size distribution of type IIb myofibers (**H**) or the other types (I + IIa + IIx) of myofibers (**I**) in TA muscles of WT and *Mettl21e*^{KO} mice; *n* = 5 pairs of mice. Error bars represent mean + sd **P* < 0.05 (Student's *t* test).

Mettl21e, indicative of a loss-of-function mutation (Fig. 4B). Importantly, the weights of 2 representative fast muscles (TA and EDL) were significantly lower in the *Mettl21e*^{KO} mice compared with WT littermates (Fig. 4C). Interestingly, the decreases in muscle weight is accompanied by an increase in the weights of subcutaneous and epididymal white adipose tissues (Fig. 4C). As a result, there was no distinguishable difference in the overall body weight and tissue weights of other tissues between WT and *Mettl21e*^{KO} mice (Fig. 4C, D).

We further investigated the histology of fast-twitch muscles. Numbers of myofiber were comparable between WT and *Mettl21e*^{KO} mice in several different muscles (Fig. 4E). Staining for Myh isoforms showed that *Mettl21e*^{KO} did not change myofiber type composition in EDL muscles (Fig. 4F). However, type IIb myofibers had smaller size in *Mettl21e*^{KO} mice than in WT mice (Fig. 4G). Analysis of size distribution of type IIb myofibers showed that

Mettl21e^{KO} myofibers had a peak at 40–45 μ m, whereas the WT myofiber diameter peaked at 45–50 μ m (Fig. 4H). In contrast, diameters of the other myofiber types (type I + IIa + IIx) were comparable between WT and *Mettl21e*^{KO} mice (Fig. 4I). These loss-of-function studies provide compelling evidence that *Mettl21e* is essential for maintaining type IIb myofiber size but dispensable for myofiber patterning.

Mettl21e methylates Vcp and modulates the activity of 26S proteasome

We performed RNA-sequencing analysis to better understand how *Mettl21e* KO alters TA muscle gene expression. Compared with TA muscles of WT mice, there were only 3 genes that were up-regulated and 6 genes that were down-regulated by more than 2-fold (fold change >2 and *P* < 0.01) in *Mettl21e*^{KO} samples in triplicate (Fig. 5A).

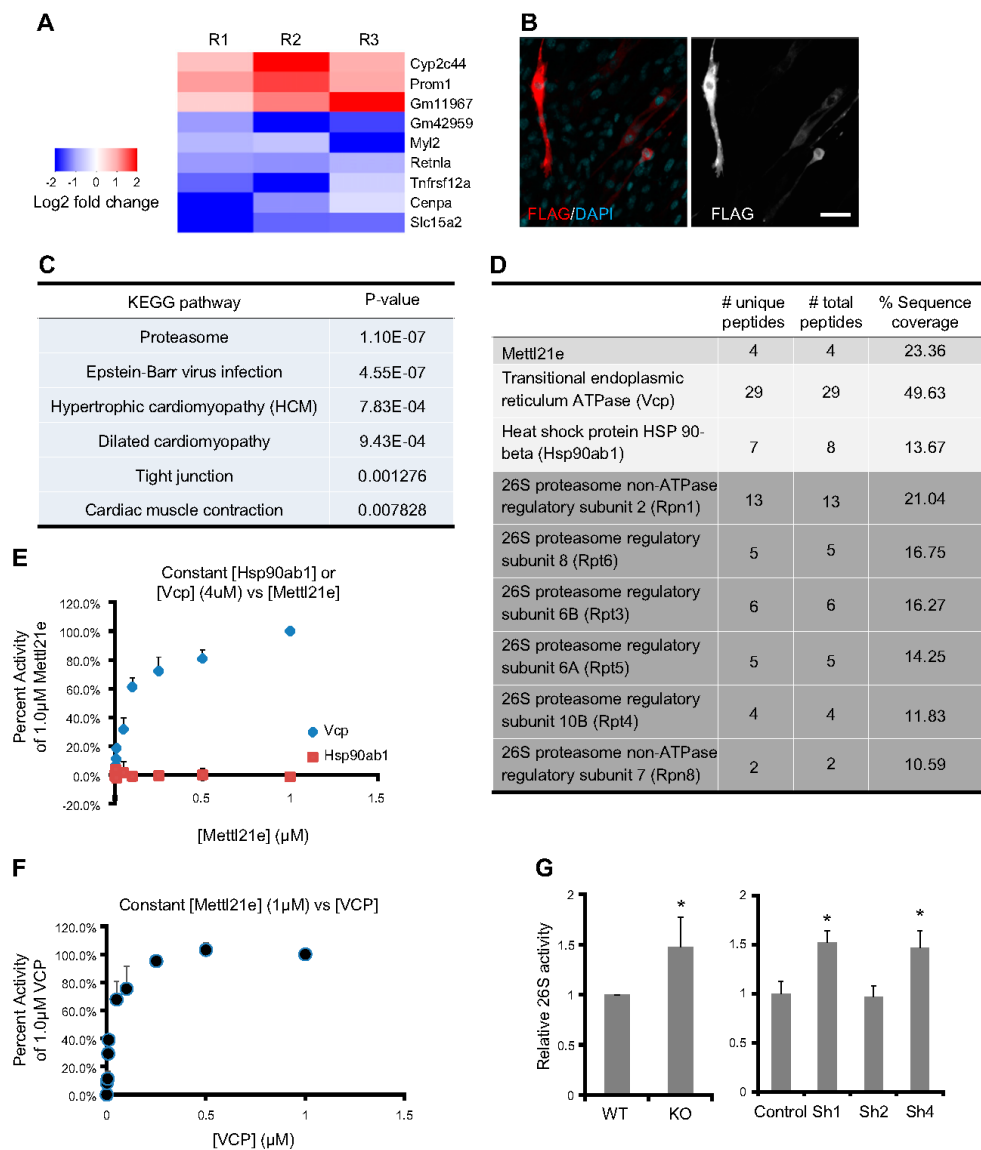


Figure 5. Mettl21e methylates Vcp and regulates 26S proteasomes activity. *A*) Heat map showing genes changed in Mettl21e^{KO} muscles (fold change >2 and $P < 0.01$) in 3 biologic replicates, including 3 up-regulated and 6 down-regulated. Blue indicates the change level of down-regulated genes, and red shows the change level of up-regulated genes. *B*) Immunostaining of FLAG in C2C12 myoblasts transfected with Mettl21e-FLAG plasmid. Scale bar, 50 μ m. Scale bar, 100 μ m. *C*) Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of Mettl21e-associated proteins identified by mass spectrometry. *D*) Mettl21e-associated proteins related to proteasome. *E*) *In vitro* methyltransferase kinetic analysis of Vcp or Hsp90ab1 as a substrate of Mettl21e. Vcp or Hsp90ab1 was held constant at 4 μ M in the presence of 13 μ M [¹⁴C] SAM and increasing concentrations of Mettl21e. The data are represented as a percentage of counts per minute (CPMs, representing methyl groups transferred) at 1 μ M of Mettl21e. Error bars represent mean + sd of 3 experiments performed in duplicate. *F*) *In vitro* methyltransferase kinetic analysis of Vcp as a substrate of Mettl21e. Mettl21e was held constant at 1 μ M in the presence of 13 μ M [¹⁴C] SAM and increasing concentrations of Vcp as the substrate. The data are represented as a percentage of CPMs at 1 μ M of Vcp. Error bars represent mean + sd of 3 experiments performed in duplicate. *G*) The activity of 26S proteasome in TA muscles of WT and Mettl21e^{KO} mice and in control and Mettl21e^{KD} myotubes. Error bars represent mean + sd of 5 independent biologic experiments with 3 technical repeats. * $P < 0.05$ (Student's *t* test).

The small number of genes whose expression is affected by *Mettl21e* KO suggests a post-transcriptional regulatory function of *Mettl21e*. Consistently, we found that *Mettl21e* was localized exclusively in the cytoplasm by using FLAG antibody to stain C2C12 myoblasts transfected with a pcDNA3.1-*Mettl21e*-FLAG plasmid that encodes a *Mettl21e*-FLAG fusion protein (Fig. 5B).

We next sought to identify *Mettl21e*-associated proteins that could be responsible for the regulation of myofiber size. We used FLAG antibody to immunoprecipitate protein complexes from myocytes transduced with *Mettl21e*-FLAG or *GFP*-FLAG adenoviral vectors. Using mass spectrometry, we identified a list of proteins specifically in *Mettl21e*-associated complexes but not in *GFP*-associated complexes (Supplemental Table S2). To narrow down the list of potential substrates, we did Kyoto Encyclopedia of Genes and Genomes (KEGG) signaling pathway analysis of the identified proteins, and the top enriched pathway is proteasome (Fig. 5C). We then prioritized the *Mettl21e*-associated proteins that are related to proteasome complexes, including 6 subunits of 26S proteasome [regulatory particle non-ATPase 1 (Rpn1), Rpn8, regulatory particle triple-A ATPase 3 (Rpt3), Rpt4, Rpt5, and Rpt6] and 2 chaperone proteins [Vcp and heat shock protein HSP 90-beta (Hsp90ab1)] together with *Mettl21e* (Fig. 5D).

Given the established role of Vcp and Hsp90 in 26S proteasome function (33, 34) and the recent report that *Mettl21* proteins preferentially interact with chaperone proteins (18), we examined whether Vcp or Hsp90ab1 is subject to *Mettl21e*-mediated methylation using *in vitro* methylation assay. Recombinant Vcp or Hsp90ab1 protein was held constant at 4 μ M in the presence of [14 C] SAM and increasing concentrations of recombinant *Mettl21e*, and the incorporation of methyl groups into proteins was measured as TCA-insoluble radioactivity. The incorporated radioactivity indicates that Vcp but not Hsp90ab1 is dose-dependently methylated by *Mettl21e* (Fig. 5E). Conversely, we held recombinant *Mettl21e* at 1 μ M with increasing concentrations of Vcp. The radioactivity curve confirms that *Mettl21e* methylates Vcp dose dependently (Fig. 5F).

Vcp is involved in the regulation of the ubiquitin-proteasome system (35). To determine if *Mettl21e* methylation of Vcp leads to alterations in the activity of proteasome, we performed 26S proteasome activity assay. KO of *Mettl21e* increased the activity of 26S proteasome by 47% in TA muscles (Fig. 5G). In addition, *Mettl21e*^{KD} myotubes had about 50% higher proteasome activities than the control myotubes (Fig. 5G). These results indicate that *Mettl21e* methylates Vcp and modulates the activity of 26S proteasome.

Inhibition of proteasome activity prevents atrophy of *Mettl21e* KD myotubes

To verify that *Mettl21e* regulates myotube size through proteasome, we inhibited proteasome activity in control and *Mettl21e*^{KD} myotubes using bortezomib, a Food and Drug Administration-approved 26S proteasome inhibitor. To this end, myoblasts were transduced with the adenoviral-*GFP* vector (control) and Sh1 (*Mettl21e*^{KD}) at d 1 after

serum withdrawal. After differentiation for an extra 2 d, myotubes were treated with DMSO or 25 nM bortezomib for 2 d. In the absence of the inhibitor, *Mettl21e*^{KD} myotubes consistently exhibited thinner morphology than did the control KD myotubes (Fig. 6A). In contrast, there was no distinguishable difference between control and *Mettl21e*^{KD} myotubes in the presence of bortezomib (Fig. 6A). Quantifications of fusion index and relative myotube diameter showed that bortezomib treatment blunted the differences between control and *Mettl21e*^{KD} myotubes (Fig. 6B). Consistently, bortezomib treatment increased the protein levels of Myh in control myotubes and normalized the levels of Myh in *Mettl21e*^{KD} myotubes (Fig. 6C). Together, we conclude that *Mettl21e* maintains myofiber size through a proteasome-dependent pathway.

DISCUSSION

Mettl21e is one of the top up-regulated genes in *Mstn* KO fast-twitch muscles (25), but we show that it is not directly regulated by *Mstn*. Instead, increased expression of *Mettl21e* was due to increases in the abundance of IIb myofibers in *Mstn* KO fast-twitch muscles (36). This finding also explains the up-regulation of *Mettl21e* in hypertrophic muscles of Callipyge sheep (27), in which a substantial increase of type IIb myofibers in hypertrophic muscles was reported (37, 38). Interestingly, the counterpart of *Mettl21e* in humans is denoted as a unitary pseudogene (39). Human *METTL21EP* was possibly inactivated by a splice-junction mutation (AG to TA) located at the acceptor-splice site of its second intron. This splice-junction mutation is accompanied by the lack of detectable type IIb myofibers in human skeletal muscles (40).

Mettl21e is located in cytoplasm. This cytoplasmic localization indicates that *Mettl21e* is not a histone methyltransferase that functions to regulate gene transcription. Our mass spectrometry data show that *Mettl21e* physically interacts with subunits of 26S proteasome and its associated chaperone proteins. *Mettl21e*^{KO} or *Mettl21e*^{KD} increases the proteasome activity, therefore reducing myofiber size. Moreover, inhibition of proteasome activity rescues *Mettl21e*^{KD}-induced myofiber atrophy, demonstrating that *Mettl21e* regulates myofiber size through proteasome.

Our data shows that *Mettl21e* methylates Vcp, which is reported as a substrate of human *METTL21C* and *METTL21D* (18, 19, 23). Vcp is an important element of the ubiquitin-proteasome system (35, 41). It acts as a molecular chaperone to bring ubiquitinated substrates to the 26S proteasome for degradation (33, 41). It also maintains the activity of 26S proteasome through binding and antagonizing the inhibitory function of the proteasome inhibitor PI31 (42). However, the effect of methylation of Vcp on proteasome has yet to be determined (19). Current knowledge indicates that nonhistone methylation regulates protein functions through the crosstalk with other post-translational modification or through changing protein-protein interactions (17, 43). It is possible that *Mettl21e* methylates Vcp to affect its binding with the

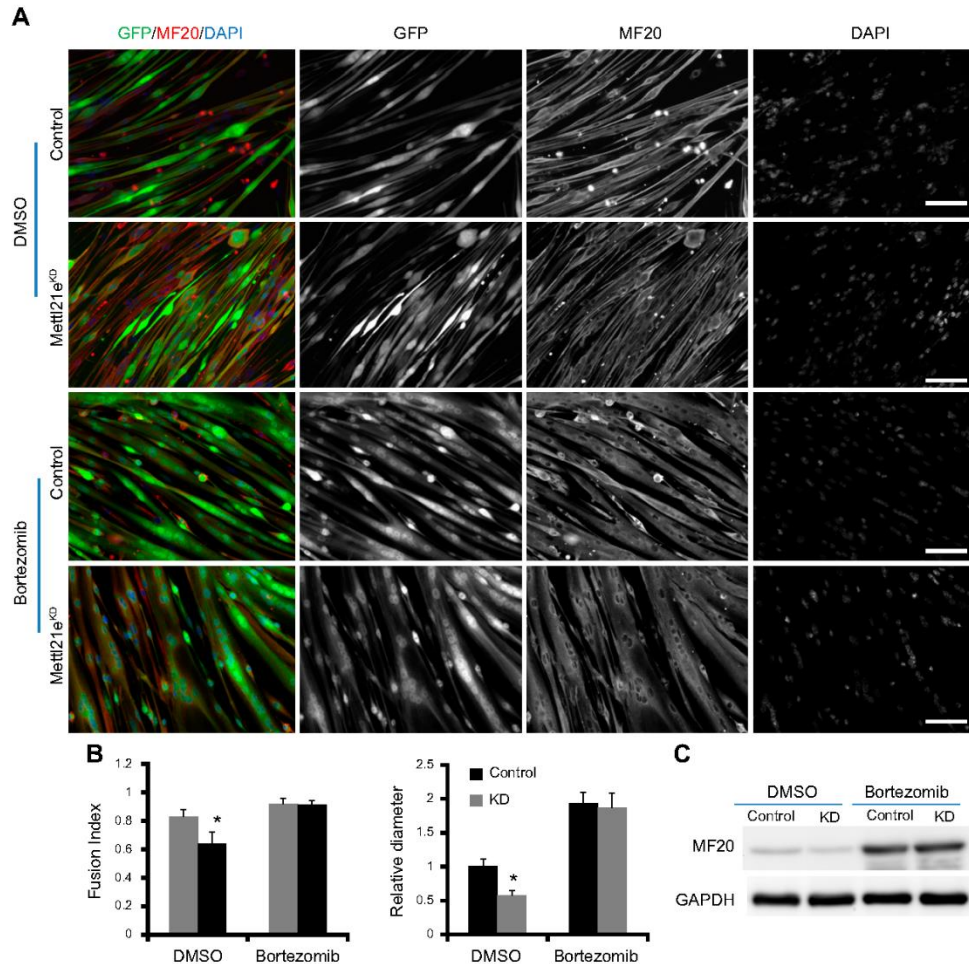


Figure 6. Bortezomib rescues myofiber atrophy induced by Mettl21e KD. **A**) Control and Mettl21e^{KD} cell were treated with DMSO or Bortezomib for 2 d. The differentiated myoblasts were stained by MF20 (Red), and nuclei were counterstained by DAPI (Blue). Scale bars, 100 μ m. **B**) Fusion index and relative diameters of myotubes counted according to the staining in **A**. Only GFP⁺ cells were used for quantification. **C**) Western blots analysis of Myh. Error bars represent mean + SD of 5 independent biologic experiments. * $P < 0.05$ (Student's *t* test).


other proteins and thus affects proteasome activity. It is also possible that Mettl21e directly methylates one of the 26S proteasome subunits to regulate proteasome activity. Understanding the detailed mechanisms will extend our knowledge about molecular regulation of proteasome activity.

Mettl21e KO muscle showed a relatively milder phenotype in comparison with Mettl21e KD myotubes. We mutated Mettl21e by a frame-shifting indel, resulting in a premature termination codon and a truncated protein that does not contain the methyltransferase domain. It is possible that the truncated protein retains some methyltransferase-independent function in regulating myofiber size. In addition, the most recent studies

report a nonsense-induced transcriptional compensation mechanism, in which mRNA containing premature termination codon is degraded through nonsense-mediated RNA decay to induce transcriptional up-regulation of genes similar to itself (44, 45). In this regard, other members of the Mettl21 family proteins may have partially compensated for the Mettl21e loss-of-function.

Myofiber size may be regulated at both cellular (myogenic differentiation and myoblast fusion) (46) and molecular (protein synthesis and degradation) level. Our data do not support a role of Mettl21e in myogenic differentiation. First, Mettl21e expression increases by 35-fold at 3 d after differentiation, suggesting that Mettl21e mainly functions after myogenic differentiation. Second, Mettl21e

KD did not inhibit myoblast differentiation. However, we could not completely rule out a role of Mettl21e in myoblast fusion based on the observation that Mettl21e KD reduces myoblast fusion. As myoblast fusion was also increased with inhibition of proteasome, it is possible that the reduced myoblast fusion after Mettl21e KD is a secondary effect of reduced proteasome activity.

Our finding helps to explain the myofiber-size paradox, in which oxidative myofibers are smaller than glycolytic myofibers despite the higher myonuclei content (per volume) and higher rates of protein synthesis in the oxidative myofibers (3, 47, 48). One explanation for this paradox is due to differential rate of protein degradation, which is mainly mediated by the ubiquitin-proteasome system. This system includes 2 discrete and successive processes: conjugating substrate proteins with multiple ubiquitin molecules and the subsequent degradation of the tagged proteins by the 26S proteasome (49). The conjugation of ubiquitin to lysine residues of substrate proteins requires 3 ATP-dependent enzymatic steps (49). E3 is the key enzyme in the conjugation process. Two muscle-specific E3, muscle atrophy Fbox and muscle RING finger 1, are found to have a higher expression in oxidative muscles than in glycolytic muscles (3, 50). Thus, the higher rates of protein degradation in the oxidative muscles may have underlined their smaller size. Here, we provide evidence that type IIb myofibers also employ an active mechanism to maintain their larger size through Mettl21e-mediated inhibition of 26S proteasome activity. 

ACKNOWLEDGMENTS

The authors thank Fengfeng Zhang (ViewSolid Biotech, Beijing, China) for generating the methyltransferase-like 21e knockout mice, Chris Bidwell (Purdue University, West Lafayette, IN, USA) for discussion and sharing Callipyge sheep data, Jessie Ellis (Purdue University) for technical assistance, Jun Wu (Purdue University) and Mary Larimore (Purdue University) for mouse colony maintenance, and members of the Kuang laboratory for valuable comments. This work was supported by a grant from the U.S. National Institutes of Health, National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) (R01AR071649) and Innovation Fund for Medical Sciences (CIFMS) from Chinese Academy of Medical Sciences (2016-I2M-1-012). The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

C. Wang and S. Kuang conceived the project, designed the experiments, and prepared the manuscript; C. Wang, B. Zhang, A. C. Ratliff, J. Arrington, J. Chen, Y. Xiong, F. Yue., Y. Nie, and W. Jin performed the experiments and analyzed the data; and K. Hu, W. A. Tao, C. A. Hrycyna, and X. Sun provided key reagents and technical assistance.

REFERENCES

- Schiaffino, S., Gorza, L., Sartore, S., Saggin, L., Ausoni, S., Vianello, M., Gundersen, K., and Lomo, T. (1989) Three myosin heavy chain

- isoforms in type 2 skeletal muscle fibres. *J. Muscle Res. Cell Motil.* **10**, 197–205
- Schiaffino, S., and Reggiani, C. (2011) Fiber types in mammalian skeletal muscles. *Physiol. Rev.* **91**, 1447–1531
- Van Wessel, T., de Haan, A., van der Laarse, W. J., and Jaspers, R. T. (2010) The muscle fiber type-fiber size paradox: hypertrophy or oxidative metabolism? *Eur. J. Appl. Physiol.* **110**, 665–694
- Sanes, J. R., and Lichtman, J. W. (1999) Development of the vertebrate neuromuscular junction. *Annu. Rev. Neurosci.* **22**, 389–442
- Dolmetsch, R. E., Lewis, R. S., Goodnow, C. C., and Healy, J. I. (1997) Differential activation of transcription factors induced by Ca²⁺ response amplitude and duration. *Nature* **386**, 855–858
- Naya, F. J., Mercer, B., Shelton, J., Richardson, J. A., Williams, R. S., and Olson, E. N. (2000) Stimulation of slow skeletal muscle fiber gene expression by calcineurin in vivo. *J. Biol. Chem.* **275**, 4545–4548
- Wu, H., Rothermel, B., Kanatous, S., Rosenberg, P., Naya, F. J., Shelton, J. M., Hutcheson, K. A., DiMaio, J. M., Olson, E. N., Bassel-Duby, R., and Williams, R. S. (2001) Activation of MEF2 by muscle activity is mediated through a calcineurin-dependent pathway. *EMBO J.* **20**, 6414–6423
- Potthoff, M. J., Wu, H., Arnold, M. A., Shelton, J. M., Backs, J., McAnally, J., Richardson, J. A., Bassel-Duby, R., and Olson, E. N. (2007) Histone deacetylase degradation and MEF2 activation promote the formation of slow-twitch myofibers. *J. Clin. Invest.* **117**, 2459–2467
- Hogan, P. G., Chen, L., Nardone, J., and Rao, A. (2003) Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes Dev.* **17**, 2205–2232
- Liu, N., Garry, G. A., Li, S., Bezprozvannaya, S., Sanchez-Ortiz, E., Chen, B., Shelton, J. M., Jaichander, P., Bassel-Duby, R., and Olson, E. N. (2017) A Twist2-dependent progenitor cell contributes to adult skeletal muscle. *Nat. Cell Biol.* **19**, 202–213
- Stark, D. A., Coffey, N. J., Pancoast, H. R., Arnold, L. L., Walker, J. P., Vallée, J., Robitaille, R., Garcia, M. L., and Cornelison, D. D. (2015) Ephrin-A3 promotes and maintains slow muscle fiber identity during postnatal development and reinnervation. *J. Cell Biol.* **211**, 1077–1091
- Van Rooij, E., Quiat, D., Johnson, B. A., Sutherland, L. B., Qi, X., Richardson, J. A., Kelm, R. J., Jr., and Olson, E. N. (2009) A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Dev. Cell* **17**, 662–673
- Lin, J., Wu, H., Tarr, P. T., Zhang, C. Y., Wu, Z., Boss, O., Michael, L. F., Puigserver, P., Isotani, E., Olson, E. N., Lowell, B. B., Bassel-Duby, R., and Spiegelman, B. M. (2002) Transcriptional co-activator PGC-1 alpha drives the formation of slow-twitch muscle fibres. *Nature* **418**, 797–801
- Lee, K. Y., Singh, M. K., Ussar, S., Wetzel, P., Hirshman, M. F., Goodyear, L. J., Kispert, A., and Kahn, C. R. (2015) Tbx15 controls skeletal muscle fibre-type determination and muscle metabolism. *Nat. Commun.* **6**, 8054
- Quiat, D., Voelker, K. A., Pei, J., Grishin, N. V., Grange, R. W., Bassel-Duby, R., and Olson, E. N. (2011) Concerted regulation of myofiber-specific gene expression and muscle performance by the transcriptional repressor Sox6. *Proc. Natl. Acad. Sci. USA* **108**, 10196–10201
- Copeland, R. A., Solomon, M. E., and Richon, V. M. (2009) Protein methyltransferases as a target class for drug discovery. *Nat. Rev. Drug Discov.* **8**, 724–732
- Biggar, K. K., and Li, S. S. (2015) Non-histone protein methylation as a regulator of cellular signalling and function. *Nat. Rev. Mol. Cell Biol.* **16**, 5–17
- Cloutier, P., Lavallée-Adam, M., Faubert, D., Blanchette, M., and Coulombe, B. (2013) A newly uncovered group of distantly related lysine methyltransferases preferentially interact with molecular chaperones to regulate their activity. *PLoS Genet.* **9**, e1003210
- Kernstock, S., Davydova, E., Jakobsson, M., Moen, A., Pettersen, S., Mælandsmo, G. M., Egge-Jacobsen, W., and Føltnes, P. O. (2012) Lysine methylation of VCP by a member of a novel human protein methyltransferase family. *Nat. Commun.* **3**, 1038
- Jakobsson, M. E., Moen, A., Bousset, L., Egge-Jacobsen, W., Kernstock, S., Melki, R., and Føltnes, P. O. (2013) Identification and characterization of a novel human methyltransferase modulating Hsp70 protein function through lysine methylation. *J. Biol. Chem.* **288**, 27752–27763
- Hamey, J. J., Wienert, B., Quinlan, K. G. R., and Wilkins, M. R. (2017) METTL21B is a novel human lysine methyltransferase of translation elongation factor 1A: discovery by CRISPR/Cas9 knockout. *Mol. Cell. Proteomics* **16**, 2229–2242

22. Malecki, J., Aileni, V. K., Ho, A. Y. Y., Schwarz, J., Moen, A., Sørensen, V., Nilges, B. S., Jakobsson, M. E., Leidel, S. A., and Falnes, P. O. (2017) The novel lysine specific methyltransferase METTL21B affects mRNA translation through inducible and dynamic methylation of Lys-165 in human eukaryotic elongation factor 1 alpha (eEF1A). *Nucleic Acids Res.* **45**, 4370–4389
23. Wiederstein, J. L., Nolte, H., Günther, S., Piller, T., Baraldo, M., Kostin, S., Bloch, W., Schindler, N., Sandri, M., Blaauw, B., Braun, T., Höpfer, S., and Krüger, M. (2018) Skeletal muscle-specific methyltransferase METTL21C trimethylates p97 and regulates autophagy-associated protein breakdown. *Cell Rep.* **23**, 1342–1356
24. Jakobsson, M. E., Davydova, E., Malecki, J., Moen, A., and Falnes, P. O. (2015) Saccharomyces cerevisiae eukaryotic elongation factor 1A (eEF1A) is methylated at Lys-390 by a METTL21-like methyltransferase. *PLoS One* **10**, e0131426
25. Welle, S., Cardillo, A., Zanche, M., and Tawil, R. (2009) Skeletal muscle gene expression after myostatin knockout in mature mice. *Physiol. Genomics* **38**, 342–350
26. Caetano-Anollés, K., Mishra, S., and Rodriguez-Zas, S. L. (2015) Synergistic and antagonistic interplay between myostatin gene expression and physical activity levels on gene expression patterns in triceps Brachii muscles of C57/BL6 mice. *PLoS One* **10**, e0116828
27. Fleming-Waddell, J. N., Olbricht, G. R., Taxis, T. M., White, J. D., Vuocolo, T., Craig, B. A., Tellam, R. L., Neary, M. K., Cockett, N. E., and Bidwell, C. A. (2009) Effect of DLK1 and RTL1 but not MEG3 or MEG8 on muscle gene expression in Callipyge lambs. *PLoS One* **4**, e7399
28. Kubik, R. M., Tietze, S. M., Schmidt, T. B., Yates, D. T., and Petersen, J. L. (2018) Investigation of the skeletal muscle transcriptome in lambs fed β adrenergic agonists and subjected to heat stress for 21 d. *Transl. Anim. Sci.* **2**, S53–S56
29. McPherron, A. C., Lawler, A. M., and Lee, S. J. (1997) Regulation of skeletal muscle mass in mice by a new TGF- β superfamily member. *Nature* **387**, 83–90
30. Wang, C., Yue, F., and Kuang, S. (2017) Muscle histology characterization using H&E staining and muscle fiber type classification using immunofluorescence staining. *Bio Protoc.* **7**, e2279
31. Wang, C., Wang, M., Arrington, J., Shan, T., Yue, F., Nie, Y., Tao, W. A., and Kuang, S. (2017) Ascl2 inhibits myogenesis by antagonizing the transcriptional activity of myogenic regulatory factors. *Development* **144**, 235–247
32. Welle, S., Burgess, K., and Mehta, S. (2009) Stimulation of skeletal muscle myofibrillar protein synthesis, p70 S6 kinase phosphorylation, and ribosomal protein S6 phosphorylation by inhibition of myostatin in mature mice. *Am. J. Physiol. Endocrinol. Metab.* **296**, E567–E572
33. Dai, R. M., and Li, C. C. (2001) Valosin-containing protein is a multi-ubiquitin chain-targeting factor required in ubiquitin-proteasome degradation. *Nat. Cell Biol.* **3**, 740–744
34. Imai, J., Maruya, M., Yashiroda, H., Yahara, I., and Tanaka, K. (2003) The molecular chaperone Hsp90 plays a role in the assembly and maintenance of the 26S proteasome. *EMBO J.* **22**, 3557–3567
35. Klopsteck, P., Ewens, C. A., Förster, A., Zhang, X., and Freemont, P. S. (2012) Regulation of p97 in the ubiquitin-proteasome system by the UBX protein-family. *Biochim. Biophys. Acta* **1823**, 125–129
36. Amthor, H., Macharia, R., Navarrete, R., Schuelke, M., Brown, S. C., Otto, A., Voit, T., Muntoni, F., Vrbova, G., Partridge, T., Zammit, P., Bunger, L., and Patel, K. (2007) Lack of myostatin results in excessive muscle growth but impaired force generation. *Proc. Natl. Acad. Sci. USA* **104**, 1835–1840; erratum: 104, 4240
37. Carpenter, C. E., Rice, O. D., Cockett, N. E., and Snowder, G. D. (1996) Histology and composition of muscles from normal and callipyge lambs. *J. Anim. Sci.* **74**, 388–393
38. Bidwell, C. A., Waddell, J. N., Taxis, T. M., Yu, H., Tellam, R. L., Neary, M. K., and Cockett, N. E. (2014) New insights into polar overdominance in callipyge sheep. *Anim. Genet.* **45**(Suppl 1), 51–61
39. Zhang, Z. D., Frankish, A., Hunt, T., Harrow, J., and Gerstein, M. (2010) Identification and analysis of unitary pseudogenes: historic and contemporary gene losses in humans and other primates. *Genome Biol.* **11**, R26
40. Smerdu, V., Karsch-Mizrachi, L., Campione, M., Leinwand, L., and Schiaffino, S. (1994) Type IIx myosin heavy chain transcripts are expressed in type IIb fibers of human skeletal muscle. *Am. J. Physiol.* **267**, C1723–C1728
41. Meyer, H., and Wehl, C. C. (2014) The VCP/p97 system at a glance: connecting cellular function to disease pathogenesis. *J. Cell Sci.* **127**, 3877–3883
42. Clemen, C. S., Marko, M., Strucksberg, K. H., Behrens, J., Wittig, I., Gärtner, L., Winter, L., Chevessier, F., Matthias, J., Türk, M., Tangavelou, K., Schütz, J., Arhzaouy, K., Klopffleisch, K., Hanisch, F. G., Rottbauer, W., Blümcke, I., Just, S., Eichinger, L., Hofmann, A., and Schröder, R. (2015) VCP and PSMF1: antagonistic regulators of proteasome activity. *Biochem. Biophys. Res. Commun.* **463**, 1210–1217
43. Zhang, X., Wen, H., and Shi, X. (2012) Lysine methylation: beyond histones. *Acta Biochim. Biophys. Sin. (Shanghai)* **44**, 14–27
44. Ma, Z., Zhu, P., Shi, H., Guo, L., Zhang, Q., Chen, Y., Chen, S., Zhang, Z., Peng, J., and Chen, J. (2019) PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components. *Nature* **568**, 259–263
45. El-Brolosy, M. A., Kontarakis, Z., Rossi, A., Kuenne, C., Günther, S., Fukuda, N., Kikhi, K., Boezio, G. L. M., Takacs, C. M., Lai, S. L., Fukuda, R., Gerri, C., Giraldez, A. J., and Stainier, D. Y. R. (2019) Genetic compensation triggered by mutant mRNA degradation. *Nature* **568**, 193–197
46. Reza, M. M., Subramaniam, N., Sim, C. M., Ge, X., Sathikumar, D., McFarlane, C., Sharma, M., and Kambadur, R. (2017) Irisin is a pro-myogenic factor that induces skeletal muscle hypertrophy and rescues denervation-induced atrophy. *Nat. Commun.* **8**, 1104
47. Tseng, B. S., Kasper, C. E., and Edgerton, V. R. (1994) Cytoplasm-to-myonucleus ratios and succinate dehydrogenase activities in adult rat slow and fast muscle fibers. *Cell Tissue Res.* **275**, 39–49
48. Goldberg, A. L. (1967) Protein synthesis in tonic and phasic skeletal muscles. *Nature* **216**, 1219–1220
49. Lecker, S. H., Goldberg, A. L., and Mitch, W. E. (2006) Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *J. Am. Soc. Nephrol.* **17**, 1807–1819
50. Bodine, S. C., Latres, E., Baumhueter, S., Lai, V. K., Nunez, L., Clarke, B. A., Poueymirou, W. T., Panaro, F. J., Na, E., Dharmarajan, K., Pan, Z. Q., Valenzuela, D. M., DeChiara, T. M., Stitt, T. N., Yancopoulos, G. D., and Glass, D. J. (2001) Identification of ubiquitin ligases required for skeletal muscle atrophy. *Science* **294**, 1704–1708

Received for publication March 4, 2019.

Accepted for publication April 29, 2019.

Methyltransferase-like 21c methylates and stabilizes the heat shock protein HSPA8 in type I myofibers in mice

Chao Wang^{1*}, Justine Arrington², Anna C. Ratliff², Jingjuan Chen¹, Hannah E. Horton³, Yaohui Nie¹, Feng Yue¹, Christine A. Hrycyna^{2,5}, W. Andy Tao^{4,5}, Shihuan Kuang^{1,5*}

From the ¹ Department of Animal Sciences, Purdue University, West Lafayette 47907, Indiana, USA; ² Department of Chemistry, Purdue University, West Lafayette 47907, Indiana, USA; ³ Department of Biological Sciences, Purdue University, West Lafayette 47907, Indiana, USA; ⁴ Department of Biochemistry, Purdue University, West Lafayette 47907, Indiana, USA; ⁵ Center for Cancer Research, Purdue University, West Lafayette 47907, Indiana, USA

Running title: *Mettl21c methylates and stabilizes Hspa8*

*Authors for correspondence (Email: skuang@purdue.edu and wangc129@gmail.com)

Keywords: methyltransferase-like 21c (Mettl21c); protein methylation; chaperone-mediated autophagy (CMA); skeletal muscle; myocyte enhancer factor 2 (Mef2); post-translational modification; heat shock protein family A (Hsp70) member 8 (HSPA8); Hsc70; myogenesis

ABSTRACT

Protein methyltransferases mediate post-translational modifications of both histone and non-histone proteins. Whereas histone methylation is well known to regulate gene expression, the biological significance of non-histone methylation is poorly understood. Methyltransferase-like 21c (METTL21c) is a newly classified non-histone lysine methyltransferase whose *in vivo* function has remained elusive. Using a *Mettl21c*^{LacZ} knockin mouse model, we show here that *Mettl21c* expression is absent during myogenesis and restricted to mature type I (slow) myofibers in the muscle. Using co-immunoprecipitation, mass spectrometry and methylation assays, we demonstrate that METTL21c tri-methylates heat shock protein 8 (HSPA8) at Lys-561 and enhances its stability. As such, *Mettl21c* knockout reduced HSPA8 trimethylation and protein levels in slow muscles, and *Mettl21c* overexpression in myoblasts increased HSPA8 trimethylation and protein levels. We further show that METTL21c-mediated stabilization of HSPA8 enhances its function in chaperone-mediated autophagy, leading to degradation of

client proteins such as the transcription factors myocyte enhancer factor 2A (MEF2A) and MEF2D. In contrast, *Mettl21c* knockout increased MEF2 protein levels in slow muscles. These results identify HSPA8 as a METTL21c substrate and reveal that non-histone methylation has a physiological function in protein stabilization.

INTRODUCTION

The skeletal muscle provides body support, powers body movements, and regulates systemic energy homeostasis. Mammalian skeletal muscles are heterogeneous, composing two general types of muscle cells (myofibers): the slow-twitch (type I) and fast-twitch (type II) myofibers (1). The myocyte enhancer factor-2 (Mef2) family transcription factors are indispensable for muscle development and for the establishment of type-I myofibers (2-4). However, the level and activity of Mef2 proteins must be controlled in the adult during muscle maintenance, as aberrantly enhanced Mef2 activity is associated with myotonia and compromised mitochondria function (5,6). Current understanding about the regulation of Mef2 is mainly on its transcriptional activity.

Mef2 is activated by Ca^{2+} /calmodulin kinases (7), calcium-activated serine/threonine phosphatase calcineurin (8), and mitogen-activated protein kinases, such as p38 and ERK5 (9,10). By contrast, the activity of Mef2 is repressed by class II histone deacetylases (11,12). Recent studies have also pointed to a role of post-transcriptional regulation of *Mef2* by microRNAs (6).

At the post-translational level, heat shock protein 8 (Hspa8/Hsc70) plays a role in regulating Mef2 proteins through chaperone-mediated autophagy (CMA). In CMA, cytoplasmic client proteins with CMA-targeting motif are recognized and delivered by heat shock protein 8 (Hspa8/Hsc70) to lysosome through the lysosomal receptor, lysosomal-associated membrane protein 2A (Lamp2a) (13). Hspa8 has been shown to mediate the degradation of two major isoforms of the Mef2 proteins, Mef2A and Mef2D, in neuronal cells (14,15). However, there is lack of a physiological context linking Hspa8 and Mef2 proteins in skeletal muscle.

Protein lysine methyltransferases (PKMTs) transfer a methyl group (CH₃) from S-adenosyl-L-methionine (SAM) to lysine residuals of substrates (16). Methylation alters the hydrophobic and steric properties of a lysine residue, leading to changes in protein stability, protein-protein interactions and protein-DNA interactions (17). Histone lysine methylation regulates chromatin structures and thus modulates gene expression (18), but the biological significance of non-histone methylation is poorly understood (19). Recently, a group of 10 non-histone PKMTs is characterized as methyltransferase family 16 (MTF16), among this family METTL21A-D forms a subgroup based on phylogenetic analysis (20,21). Intriguingly, *METTL21C* is identified as a pleiotropic gene for both skeletal muscle and bone through a genome-wide association study (22). Knockdown of *Mettl21c* impairs myogenesis of C2C12 myoblasts (22). A recent study identifies that *Mettl21c* is a type I myofiber specific protein, and functions to trimethylate valosin-containing protein (Vcp/p97), a protein required for autophagy (23,24).

Mettl21c^{-/-} mice have normal muscle patterning, but dysregulated autophagy and muscle weakness (23). In human cells, METTL21C physically interacts with heat shock 70 kDa (HSPA) proteins (21), but it is unknown whether the HSPA proteins are *bona fide* substrates of METTL21C.

In this study, we used the *Mettl21c*^{LacZ} knockin/knockout mouse model to track the developmental expression of *Mettl21c*. We found that *Mettl21c* was absent during myogenesis and was only expressed in mature type I myofibers. We then demonstrated that *Mettl21c* methylates and stabilizes Hspa8, thereby reducing the level of Hspa8 client proteins, Mef2A and Mef2D. These data demonstrate an *in vivo* physiological function of *Mettl21c* and reveal the biological significance of non-histone methylation in the skeletal muscle.

RESULTS

Mettl21c is not expressed during myogenesis and only expressed in mature slow myofibers

We tracked the expression of *Mettl21c* using *Mettl21c*^{LacZ/+} mice, which contain the β -Gal reporter targeted to replace exon 2–5 of *Mettl21c* (Fig. 1A). Immunofluorescence staining of the slow (type I) myofiber specific myosin heavy chain (Myh7) and X-Gal staining in serial sections of adult soleus (SOL) muscle indicate type I myofiber-specific expression of *Mettl21c* (Fig. 1B). We also isolated single myofibers from WT mice and detected the expression of *Mettl21c* specifically in type I myofibers expressing *Myh7* in a separate study (25), which confirms X-Gal signal as a reporter of *Mettl21c* expression. We then investigated the developmental expression of *Mettl21c*. Whole-mount X-Gal staining showed no discernible positive signals in limb muscles of E18.5 embryos (data not shown). Additionally, there were no X-Gal signals in SOL muscle of postnatal day 3 (P3) mice, though *Myh7*⁺ (type I) myofibers were readily detectable at the same stage (Fig. 1C), indicating that the expression of *Myh7* precedes that of *Mettl21c* and excluding a

possible role of Mettl21c in type I myofiber specification. X-Gal signals were first shown in a group of Myh7⁺ myofibers in SOL muscles of P7 mice (Fig. 1D). The X-Gal signals became stronger and nearly all Myh7⁺ myofibers had positive X-Gal staining in SOL muscle of P10 mice (Fig. 1E). Consistent with the X-Gal staining patterns, mRNA levels of *Mettl21c* were low from E18.5 to P3, then upregulated at P7, reaching the plateau levels at P10 in hindlimb muscles (Fig. 1F). Hence, *Mettl21c* is expressed after the establishment of Myh7⁺ myofiber and the expression level is increased with the maturation of Myh7⁺ myofibers.

To further confirm the absence of *Mettl21c* expression during myogenesis, we cultured primary myoblasts isolated from the *Mettl21c*^{LacZ/+} (HET) or *Mettl21c*^{LacZ/LacZ} (KO) mice, and induced the cells to differentiate. Quantitative PCR analyses showed that the mRNA levels of *Mettl21c* in the HET myotubes were identical to those of KO myotubes (both were nearly undetectable), indicative of the lack of *Mettl21c* expression in the HET myotubes (Fig. 1G). In contrast, the level of *Mettl21c* in the adult SOL muscles of HET was over 100 times higher than that in the KO muscles (Fig. 1G). In addition, there were no morphological differences between HET and KO myotubes (Fig. 1H), suggesting that Mettl21c is dispensable for myogenic differentiation. These results suggest that Mettl21c mainly functions in mature type I myofibers.

Mettl21c is indispensable for proper muscle function

We next studied the function of Mettl21c *in vivo* using KO mice models. To exclude the potential influence of β -Gal, we compared KO mice with *Mettl21c*^{LacZ/+} (HET) mice. KO mice were indistinguishable from HET and WT mice in gross morphology, body weight and SOL muscle weight (Fig. 2A and B). Total numbers and sizes of myofibers in SOL muscles were comparable between HET and KO mice (Fig. 2C and D). Immunostaining of Myh isoforms showed that the percentage of type I myofibers in SOL muscle was comparable between HET and KO mice (Fig. 2E and F), suggesting that

Mettl21c KO does not affect slow myofiber development or maintenance. However, KO mice run shorter distance and had lower speed than HET mice (Fig. 2G), indicating that Mettl21c is necessary for proper functioning of skeletal muscles.

Hspa8 is a substrate of Mettl21c

To understand the function of Mettl21c, we first attempted to identify Mettl21c associated proteins. We used FLAG antibody to immunoprecipitate (IP) protein complexes from myocytes transduced with *Mettl21c-FLAG* or *GFP-FLAG* adenoviral vectors. SDS-PAGE analysis of the proteins obtained by IP revealed a ~70 kDa protein co-purified with Mettl21c but not with GFP (Fig. 3A). Using affinity purification-mass spectrometry (AP-MS), we identified several chaperone proteins specifically in Mettl21c-associated complexes, but not in GFP-associated complexes (Fig. 3B). Among these chaperone proteins, Hspa5 and Hspa8 have a molecular weight at ~70 kDa. IP analysis confirmed that Mettl21c binds with Hspa8 (Fig. 3C), but not with Hspa5 in myocytes (Fig. 3C). These results indicate a physical association between Mettl21c and Hspa8.

We further performed *in vitro* methyltransferase assay to determine if Hspa8 is methylated by Mettl21c. Various concentrations of recombinant Mettl21c and Hspa8 proteins were reacted in the presence of [¹⁴C] SAM, and the incorporation of methyl groups into substrates was measured as TCA-insoluble radioactivity. To establish the validity of the assay, we used recombinant Vcp as a positive control, which was recently shown to be methylated by Mettl21c (23). We found that Vcp was dose-dependently methylated by Mettl21c, with a K_M at around 0.0075 (Fig. 3D). We next held the Mettl21c at a constant concentration (1 μ M) with varying concentration (0, 0.001, 0.005, 0.0075, 0.01, 0.05, 0.1, 0.25, 0.5 and 1.0 μ M) of Hspa8, and found that Hspa8 was dose-dependently methylated by Mettl21c, with a K_M at around 0.03 (Fig. 3E). The K_M of Mettl21c on Hspa8 was slightly higher than that on Vcp, indicating that Mettl21c-Hspa8 had a relatively lower reaction rate than that of Mettl21c-Vcp.

These data provide compelling biochemical evidence that Hspa8 is a new substrate of Mettl21c.

Mettl21c methylates Hspa8 at K561

Then we utilized AP-MS to characterize methylated peptides enriched by IP using a pan-methyl lysine antibody in myocytes transduced with *Mettl21c* (*Mettl21c^{OE}*) or *GFP* (as control) adenoviral vectors (Fig. 4A). This led to the identification of lysine (K) 561 in Hspa8 that was highly trimethylated in primary myoblasts transduced by *Mettl21c*-expressing adenovirus in relative to cells transduced by *GFP* control (Fig. 4B). Specifically, the abundance of the trimethylated Hspa8 peptide was 5-fold higher in *Mettl21c^{OE}* myocytes than in control myocytes (Fig. 4B), indicating that *Mettl21c* is responsible for the K561 trimethylation. To verify the methylation site, we mutated the K561 to non-methylatable alanine (A) and used the recombinant K561A-Hspa8 mutant as a substrate with increasing concentrations of *Mettl21c* in the methyltransferase assay. Whereas the WT Hspa8 was robustly methylated by *Mettl21c*, the K561A mutation abolished the dose-dependent incorporation of radioactivity and methylation by *Mettl21c* (Fig. 4C-D), indicating that K561 was the only site in Hspa8 that can be methylated by *Mettl21c*. We further utilized the same AP-MS assay to investigate methylated peptides in SOL muscles of HET and KO mice. We did not detect any K561 monomethylation in Hspa8, and only detected either dimethylation or trimethylation of K561 (Fig. 4E). Importantly, loss of *Mettl21c* decreased the abundance of trimethylated Hspa8 peptide by 51%, compensated by 11% increase in the abundance of dimethylated Hspa8 peptide (Fig. 4E). Our data demonstrate that *Mettl21c* trimethylates Hspa8 at K561.

K561 methylation stabilizes Hspa8

We next investigated the biological consequences of K561 methylation. It has been reported that the methylation perturbs interaction between Hspa8 and an E3 ubiquitin ligase C-terminus of Hsp70 Interacting Protein (CHIP), leading to decreased ubiquitination on Hspa8 (26). Thus, we hypothesize that the K561

methylation prevents Hspa8 from degradation. We transfected 293A cells with K561R or K561A mutant pcDNA5-*GFP-Hspa8* plasmids to mimic methylated or unmethylated K561, and treated the 293A cells with a protein synthesis inhibitor cycloheximide (CHX) to observe the degradation of the two mutants (Fig. 5A). Whereas the protein level of K561A mutant was decreased within 8 hours after CHX treatment, the protein level of K561R mutant was unchanged within 24 hours after CHX treatment (Fig. 5A), indicating that K561 methylation increased the stability of Hspa8.

It has been reported that Hspa8 targets Mef2A and Mef2D to lysosome for degradation through chaperone-mediated autophagy (CMA) in neurons (14,15). We therefore examined how Hspa8 affects Mef2 proteins in myotubes. We transduced myotubes by adenovirus expressing *Hspa8-GFP* (*Hspa8^{OE}*) or *GFP* only (control), and found that *Hspa8^{OE}* reduced the levels of Mef2A and Mef2D in myotubes (Fig. 5B). Consistently, *Hspa8^{OE}* decreased the transcriptional activities of Mef2A and Mef2D, but not Mef2C (Fig. 5C). Notably, K561A-Hspa8 similarly inhibited the activity of Mef2A and Mef2D (Fig. 5C), indicating that K561 methylation was not required for the inhibition of Mef2A and Mef2D.

We further investigated the effect of K561 trimethylation on Hspa8 stability. We detected the level of Hspa8 and its target proteins Mef2A and Mef2D in SOL muscles of *Mettl21c^{KO}* mice. *Mettl21c^{KO}* reduced the K561 trimethylation and the level of Hspa8 in sarcoplasm and increased the protein levels of Mef2A and Mef2D in myonuclei (Fig. 4E and 5D). To examine the effect of *Mettl21c*-mediated K561 trimethylation on Hspa8 level and function, differentiated myoblasts were transduced with *Mettl21c* (*Mettl21c^{OE}*) or *GFP* (as control) adenoviral vectors. The expression of *Mettl21c* was over 2000-fold higher in *Mettl21c^{OE}* myotubes than in control myotubes (Fig. 5E). *Mettl21c^{OE}* increased the trimethylation and the level of Hspa8, concomitant with a decrease in the level of Mef2A and Mef2D (Fig. 4B and 5F). These results provide compelling evidence that K561 methylation stabilizes Hspa8.

As Mettl21c trimethylates and maintains the level of Hspa8, which was shown to function as carrier to facilitate degradation of Mef2A and Mef2D through the CMA pathway, we hypothesize that the regulation of Mef2A and Mef2D by Mettl21c is CMA-dependent. To test this hypothesis, we selectively blocked CMA by silencing *Lamp2a* using lentiviral shRNA (27). *Lamp2a* shRNA efficiently reduced the level of Lamp2a, accompanied by an accumulation of Mef2A and Mef2D (Fig. 5G). Importantly, Lamp2a silencing prevented Mettl21c^{OE}-mediated reduction of Mef2A and Mef2D (Fig. 5F and 5G), indicating that the CMA pathway mediates the degradation of Mef2A and Mef2D in Mettl21c^{OE} myotubes. These results together establish a model in which Mettl21c methylates Hspa8, to facilitate degradation of its client proteins such as Mef2A and Mef2D (Fig. 5H).

The expression of Mettl21c is absent during muscle regeneration

Although high levels of Mef2 proteins are detrimental to mature muscles (6), Mef2 transcription factors are required for muscle regeneration (3). As Mettl21c regulates the levels of Mef2A and Mef2D, we predict that Mettl21c expression should be down-regulated during muscle regeneration to allow accumulation of Mef2 proteins. We thus examined the expression of *Mettl21c* using the *Mettl21c*^{LacZ/+} reporter mice during SOL muscle regeneration after Cardiotoxin (CTX)-induced injuries. We performed immunofluorescence and X-Gal staining on sections of regenerated SOL muscles. At day 7 post CTX injection, only a small number of regenerating Myh7⁺ myofibers had weak X-Gal signals (Fig. 6A), indicating that Mettl21c is not expressed in newly regenerated type I myofibers. At day 14 post CTX injection, a substantial number of regenerated myofibers expressed Myh7, accompanied by increased levels of X-gal signal specifically in Myh7⁺ myofibers (Fig. 6B). In parallel, *Mettl21c* mRNA levels was sharply decreased by more than 90% within 3 days after CTX-induced muscle degeneration, and gradually increased in the following 2 weeks, reaching the maximal levels at day 14 post CTX-injury (Fig. 6C). These results indicate

that Mettl21c expression is downregulated during muscle regeneration, and suggests that Mettl21c is dispensable for muscle regeneration. We also investigated the regeneration of SOL muscles that normally express Mettl21c, in HET and KO mice after CTX-injury. Consistently, the regeneration of KO muscles was comparable to that of HET muscles at day 7 post CTX-injury (Fig. 6D). These results assure the notion that Mettl21c mainly functions in mature type I myofibers.

DISCUSSION

In this study we dissect the dynamic expression and in vivo function of a type I myofiber-specific methyltransferase Mettl21c. The expression of *Mettl21c* is absent during myogenesis, indicating that Mettl21c is not required for the development of skeletal muscle, in consistent with the normal muscle morphology in KO mice. Despite the normal muscular morphology, KO mice have decreased motor activity. Wiederstein *et al.* report a similar weak muscle phenotype and they additionally show that *Mettl21c* KO leads to dysregulated autophagy program potentially through Vcp (23).

We identified Hspa8 as a novel substrate of Mettl21c and pinpointed K561 as the specific residue being trimethylated. Intriguingly, Jakobsson *et al.* reported that human HSPA8, which shares similar primary structure with murine Hspa8, is trimethylated at K561 by METTL21A, a paralog of METTL21C (28). A potential compensatory role of murine Mettl21a may explain why Mettl21c KO only partially blocks the trimethylation of Hspa8. Our AP-MS results indicate that Hspa8 is also dimethylated at K561 in skeletal muscles. However, the K561 dimethylation is not mediated by Mettl21c as deletion of Mettl21c leads to increased levels of dimethylated Hspa8. Although what dimethylates Hspa8 is unclear, SETD1A has been reported to dimethylate K561 of Hspa1, a paralog of Hspa8 (29).

At the functional level, K561 methylation of Hspa8 decreases its affinity for the E3 ligase, CHIP, and thus reduces the ubiquitination of Hspa8 (26). The latter event prevents ubiquitin-

mediated proteasomal degradation of Hspa8 and stabilizes Hspa8. Consistently, we show that the mimetic of methylated Hspa8 has better stability than unmethylated Hspa8. In addition, we found that Mettl21c^{OE} or Mettl21c^{KO} increases or decreases the protein level of Hspa8, respectively.

Hspa8 plays several different roles in the autophagy program (30). For example, Hspa8 cooperates with BAG3, which mediates macroautophagy and maintains mitochondria homeostasis (31,32). Intriguingly, impaired Hspa8-BAG3-mediated autophagy results in muscle weakness (33). Hspa8 inhibitors, such as P140 and 2-phenylethanesulfonamide, have been shown to disrupt the binding of Hspa8 with the other chaperones and downregulate autophagy program (30,34,35). Thus, the decreased level of Hspa8 may contribute to the impaired autophagy and weak muscle in the KO mice (23).

In addition, Hspa8 is a key regulator in the CMA pathway (36). It selectively recognizes KFERQ-like motif-containing proteins, including Mef2A and Mef2D, and shuttles them to lysosome for degradation. The degradation of Mef2A and Mef2D is blocked by interrupting their interactions with Hspa8 in neuronal cells (14,15). Here, we identify a conserved role of Hspa8 in skeletal muscles by demonstrating that Hspa8^{OE} reduced the protein levels of Mef2A and Mef2D. Thus, the expression of Mettl21c could harness the level of Mef2 proteins through Hspa8. The expression of *Mettl21c* is absent during myogenesis and muscle regeneration, allowing Mef2 transcription factors to accumulate and function during muscle establishment. Upon completion of muscle development or regeneration, Mettl21c is then increased to lower the level of Mef2 transcription factors. A proper level of Mef2 is key for normal muscle function, as high level of Mef2A has been shown to compromise mitochondria function in the skeletal muscle (6). Collectively, our study establishes Mettl21c as a regulator of Hspa8 to maintain its level in mature type I myofibers.

EXPERIMENTAL PROCEDURES

Mice

The Mettl21c^{LacZ/+} mice (stock#025271) were from Jackson lab (23). Unless otherwise indicated, we used 2-month old male mice for experiments. All procedures involving mice were approved by Purdue University's Animal Care and Use Committee under protocol # 1112000440.

Primary myoblast isolation, culture and differentiation

Primary myoblasts were isolated from 5-week old WT female mice. The hind limb skeletal muscles were minced and digested in type I collagenase and dispase B mixture (Roche Applied Science). The digestion was stopped with F-10 Ham's medium containing 17% fetal bovine serum (FBS), and the cells were filtered from debris, centrifuged and cultured in growth media (F-10 Ham's media supplemented with 17% FBS, 4 ng ml⁻¹ basic fibroblast growth factor and 1% penicillin-streptomycin) on uncoated dishes for three days when 5 ml growth media were added each day. Then the supernatant were collected, centrifuged and trypsinized with 0.25% trypsin. After washing off the trypsin, primary myoblasts were seeded on collagen-coated dishes, and the growth medium was changed every two days. Myoblasts were induced to differentiate on matrigel-coated dishes and cultured in differentiation media (DMEM supplemented with 2% horse serum and 1% penicillin-streptomycin). Differentiation media was replaced every day. Primary myoblasts were cultured in normal humidified tissue culture incubators with 5% CO₂.

Immunostaining, X-gal staining and image acquisition

For immunostaining of muscle samples, they were processed following the protocol described by Wang et al. (37). Generally, muscle slides were directly blocked with blocking buffer (5% goat serum, 2% BSA, 0.2% triton X-100, and 0.1% sodium azide in PBS) for at least 30 min. Dilute primary antibodies with blocking buffer. Myh2 (2F7 from Developmental Studies Hybridoma Bank (DSHB)) and Myh7 (BA-F8 from DSHB) antibodies are diluted at a ratio

1:300. Dystrophin (ab15277 from Abcam) antibody is diluted at a ratio 1:1,000. Remove blocking buffer from the sections and add diluted primary antibodies on sections overnight at 4 °C. After washing with PBS, the samples were incubated with respective secondary antibodies and DAPI for 45 min at room temperature. Myotubes were fixed with 4% paraformaldehyde, and then were blocked with blocking buffer for at least 30 min. Then the samples were incubated with primary antibodies (1:200 in blocking buffer) [MF20 was from DSHB] overnight. After washing with PBS, the samples were incubated with respective secondary antibodies and DAPI for 45 min at room temperature. Fluorescent images were captured using a Leica DM 6000B fluorescent microscope. For X-gal staining, tissue sections were fixed with 4% paraformaldehyde for 5 min. Then samples were washed with washing buffer (0.1 M phosphate buffer, pH 7.3, 2 mM MgCl₂, and 0.02% NP40) for at least 3 times. Then samples were stained with staining buffer (washing buffer containing 1 mg ml⁻¹ X-gal in dimethyl formamid, 2.12 mg ml⁻¹ potassium ferrocyanide and 1.64 mg ml⁻¹ potassium ferricyanide) overnight at 37°C. After staining, pour off the staining buffer and replace with washing buffer.

RNA extraction and Real-time qPCR

Total RNA of muscles or myoblasts were extracted using Trizol Reagent. RNA was treated with RNase-free DNase I to remove genomic DNA. The purity and concentration of total RNA were measured by Nanodrop 3000 (Thermo Fisher). Random primers and Moloney murine leukemia virus reverse transcriptase were used to convert RNA into cDNA. Real-time PCR was performed using Roche Lightcycler 480 PCR System with SYBR Green Master Mix. Forward and reverse primers for *Mettl21c* were 5'-AGGAGCTCAAGTCACAGCAACAGA and 5'-AGAGGCCAGCACGTAGTCATAACA, respectively. Ct value of *18S* rRNA was used as internal control and 2^{-ΔΔCT} method was used to analyze the relative mRNA expression of various genes.

Protein extraction and western blot analysis

Muscle samples and cultured myoblasts were washed with PBS and homogenized with radioimmune precipitation assay buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS). Protein concentrations were determined using Pierce BCA Protein Assay Reagent (Pierce Biotechnology). Proteins (100 ug) were separated by 10% SDS-PAGE, electrotransferred onto PVDF membrane (Millipore Corporation, Billerica, MA), and incubated with specific primary antibodies. Mef2A (B-4), Mef2D (H-11), GFP (B-2), and Gapdh (6C5) antibodies (1:1000 in 5% w/v nonfat dry milk) were from Santa Cruz, and Hspa8/Hsc70 (ab2788), NOQ/Myh7 (ab11083) and Lamp2a (ab18528) were from Abcam. Immunodetection was performed using enhanced chemiluminescence (ECL) Western blotting substrate (Pierce Biotechnology) and detected with FluoChem R imaging system (ProteinSimple).

Myonuclei and sarcoplasm protein preparation

SOL muscle were homogenized in the cell lysis buffer (10 mM HEPES, pH 7.5, 10 mM KCl, 0.1 mM EDTA, 1 mM DTT, 0.5% NP-40, 0.5 mM PMSF and 1x protease inhibitor cocktail) and allowed to swell on ice for 15-20 min with intermittent mixing. Tubes were vortexed to disrupt cell membranes and then centrifuged at 12,000 g at 4°C for 10 min. The supernatant is the sarcoplasm extract. The pelleted nuclei were washed twice with the cell lysis buffer and resuspended in the nuclear extraction buffer (20 mM HEPES, pH 7.5, 400 mM NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF and 1x protease inhibitor cocktail) and incubated in ice for 30 min. Myonuclei extract was collected by centrifugation at 12,000 g for 15 min at 4°C.

Adenovirus generation

The adenoviruses with Hspa8, *Mettl21c* or *Mettl21c-FLAG* were generated using the AdEasy system. Hspa8 ORF, *Mettl21c* ORF and *Mettl21c-FLAG* ORF were cloned with primers (5'-

AAGCGGCCGCATGTCTAAGGGACCTGCA
GTTGG/5'-
CGTCTAGATTAATCCACCTCTTCAATGG
for *Hspa8*; 5'-
AAGCGGCCGCATGGATCAGCATCTCCAC
ATAG/5'-
CGTCTAGATCACTCCCACCTTTAATATCCC
for *Mettl21c*; 5'-
AAGCGGCCGCATGGATCAGCATCTCCAC
ATAG/5'-
CGTCTAGATCACTTGTCGTCATCGTCTTT
GTAGTCCTCCACTTTAATATCCC for
Mettl21c-FLAG). These cloned DNA sequences
were inserted into pAdTrack-CMV plasmid (The
cloned *Mettl21c-FLAG* or *Mettl21c* ORF was
also inserted into pcDNA3.1), and then were
digested by PmeI and transfected the DH5a
competent cell with pAdEasy-1. The following
steps were exactly following the methods
described by Wang et al. (38).

Immunoprecipitation to pull down Mettl21c binding complex

Vectors containing *Mettl21c-FLAG*, or *GFP-FLAG* were transduced into primary myoblasts (70-80% confluent) by adenovirus. Two days post-transduction, cells were induced to differentiation for 3 days and then were scraped with ice-cold PBS (from 10 × 100mm plates for each plasmid), and were centrifuged. The cell pellet was completely resuspended with 1 ml lysis buffer (50mM Tris-HCl, pH 7.5, 150mM NaCl, and 1% NP40 with 1x Protease inhibitor cocktail), on ice, and sonicated with 1 sec pulse and 5 sec interval for 10 times. After 5 min on ice, cell lysate was centrifuged with 14000 rpm at 4°C, and then the supernatant (containing 5 mg proteins determined by the Pierce bicinchoninic acid assay) was transferred to 50 μ l pre-cleaned anti-FLAG magnetic beads slurry (Sigma, M8823). After incubation on rotator for 3 hours in cold room (4°C), the magnetic beads were washed three times with lysis buffer. Then, added 500 μ l ddH₂O, up and down several times to remove salt or solvent remainder (twice). Beads captured proteins were eluted with 100 μ l of 50 mM triethyl amine (TEA) and 5 mM DTT on thermal shaker (99 °C, 5 min). The protein elution was centrifuged with CentriVap Concentrator

(Labconco) to partially remove TEA, and then was adjusted to 100 μ l, pH 8.0 with 1% acetic acid and 15mM Iodoacetamide and placed in dark for 1 hour. Before desalting the protein elution with C18 Zip tips (Glygen # NT3C18.96), 1 μ g trypsin was reacted with the 100 μ l adjusted protein elution for 16 hours at 37°C. The following steps for mass spectrometric data acquisition and analysis were according to the descriptions by Wang et al. (38).

Mass spectrometry analysis of peptides with methylated lysine

Peptides with methylated lysine were enriched by immunoprecipitation using the PTMScan Pan-Methyl Lysine Kit (#14809, Cell Signaling) according to the manufacturer's instructions. In brief, myocytes or SOL muscles lysates were prepared using urea lysis buffer (20 mM HEPES pH 8.0, 9 M urea, 2.5 mM sodium pyrophosphate). Protein concentrations were determined and adjusted to be the same. DTT (4.5 mM) and Iodoacetamide (10 mM) were used to introduce reduction and alkylation of proteins. The urea was then diluted with 20 mM HEPES pH 8.0 to a final concentration of approximately 2 M. Before desalting the protein elution with C18 Sep-Paks, 1 μ g trypsin was reacted with the 100 μ l adjusted protein lysates for 16 hours at 37 °C. Freeze the protein elution (-80 °C freezer) for 2 hr and lyophilize frozen peptide solution for 2 days. Lyophilized peptides were resuspended using the IAP buffer provided in the kit. Then the IAP-peptides solution was mixed with antibody-bead slurry and incubated on a rotator for 2 hr at 4 °C. The antibody-beads were washed 3 times with IAP buffer and 3 more wash with ddH₂O. Beads captured peptides were eluted with 100 μ l of 0.15% trifluoroacetic acid (TFA). Eluted peptides were purified with StageTips. The following steps for mass spectrometric data acquisition and analysis were according to the descriptions by Wang et al. (38).

Generation of mutated Hspa8

The pcDNA5-GFP-*Hspa8* (K561A) and pcDNA5-GFP-*Hspa8* (K561R) were generated using Q5 Site-Directed Mutagenesis Kit (E0554,

NEB) according to the manufacturer's protocol. In brief, the pcDNA5-GFP-Hspa8 was amplified by PCR primers (Hspa8-A-f: GCGATTAACGATGAGGACAAAC or Hspa8-R-f: AGGATTAACGATGAGGACAAAC and Hspa8-Mu-r: GCCTTGAAGTTTCTCATCTTC). Note the plasmid concentrated should be less than 25 ng. PCR products were then reacted with KLD enzyme mix (provided in the kit). KLD mix was incubated with chemically-competent cells. Plasmids were extracted from transfected cells and sequenced to confirm the correct mutation. To detect the degradation of mutated Hspa8, 293A cells were transfected with indicated plasmid. Two days after the transfection, 293A cells were treated with 100 $\mu\text{g ml}^{-1}$ cycloheximide (Sigma, C7698) for up to 8 hours to block protein translation.

Recombinant protein production and purification

pETDuet-1-derived plasmids incorporating His6-tagged proteins were transformed into the *Escherichia coli* expression strain BL21 (DE3) (Invitrogen). Cells were cultured in LB media with 0.1 mg/mL ampicillin at 37 °C in a shaking incubator at 220 rpm until the absorbance at 600 nm reached 0.8 OD. The culture was induced with 100 μM Isopropyl β -D-thiogalactoside (Gold Biotechnology) and the temperature was lowered to 18 °C for 18 h. Cells were harvested by centrifugation at 7000xg. Cell pellets were resuspended in lysis buffer (50 mM Tris (pH 7.5), 500 mM NaCl, 10% (w/v) glycerol, 30 mM imidazole, 3 mM 2-mercaptoethanol, 0.5% Triton X-100, 1 tablet cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche), and 2 mM AEBSF (Gold Biotechnology). Cells were lysed via probe sonication and underwent centrifugation at 100,000xg for 1 h. The supernatant was rocked with Ni-NTA resin (Thermo Scientific) at 4 °C. After 1h, resin was washed with buffer (50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10% glycerol, and 30 mM imidazole), followed by an additional wash with the addition of 0.5 M KCl. Recombinant proteins were removed from resin with elution buffer (50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 300 mM imidazole). The eluted protein

was concentrated in Amicon Ultra MWCO concentrators columns (Millipore) with appropriate molecular mass cutoffs. Proteins were aliquoted and stored at -80 °C and were thawed on ice prior to use in assays. Concentrations of each protein were determined via Bradford protein assay (Thermo Scientific).

In vitro methyltransferase reaction

Methyltransferase reactions were performed in 50- μl volumes for 1h at 37 °C in methyltransferase reaction buffer (50 mM Tris (pH 7.5), 50 mM KCl, 5 mM MgCl_2 , 1mM ATP), 13 μM [^{14}C]SAM (2 μCi), 1 μM methyltransferase enzyme, and varying concentrations of substrate. The reactions were stopped by precipitating proteins with 50- μl 10% (v:v) TCA at 4° for 1 hour. The reactions were spotted onto glass fiber filters (Whatman, 45um) and the acid-insoluble material was retained during vacuum filtration. The reaction tubes were rinsed with an additional 50 μl 10% (v:v) TCA and applied to the filters. The filters were then washed with 1mL of 10% (v:v) TCA, followed by 1mL 100% ethanol, and left to dry for 10 mins at room temperature. The dried filters were placed in scintillation vials with 10mL Bio-Safe II™ biodegradable scintillation cocktail (Research Products International) and radioactivity was measured by scintillation counting.

Luciferase assay

Transient transfections were performed with Lipofectamine2000 (Thermo) according to manufacturer's instructions. HEK293 cells were grown in Dulbecco's modified Eagle's media (DMEM) supplemented with 10% FBS. Briefly, 100ng of reporter 3xMEF2-luc, 10ng *Renilla* and 300 ng of each other plasmids (pCGN-MEF2A, pcDNA3.1-MEF2C-HA, pCGN-MEF2D, pcDNA5-GFP-Hspa8 (K561A) and pcDNA5-GFP-Hspa8) was mixed with 2 μl of Lipofectamine2000 and added to 293 cells in 24-well plates. After 48 hours, cells were harvested and analyzed with the Dual-Luciferase Reporter Assay System (Promega). The total amount of DNA added in each transfection was kept constant by referring to the *Renilla* signal. The 3xMEF2-luc (Addgene plasmid # 32967),

pCGN-MEF2A (Addgene plasmid # 32958) and pCGN-MEF2D (Addgene plasmid # 32963) were gifts from Ron Prywes; the pcDNA3.1-MEF2C-HA was a gift from Andrew Lassar (Addgene plasmid# 32515); and the pcDNA5-HSPA8 was a gift from Harm Kampinga (Addgene plasmid # 19460).

Lentivirus generation

To deplete endogenous Lamp2a, shRNAs target 5'-GACTGCAGTGCAGATGAAG, 5'-CTGCAATCTGATTGATTA, and 5'-TAAACACTGCTTGACCACC, corresponding to the exon 8 of *Lamp2a* were chosen according to the description by Massey et al. (27). The hairpin (sense-loop-antisense) for these

sequences was cloned in the Plko.1 plasmid. Lentiviral particles containing Plko.1-Lamp2a-shRNAs or Plko.1-Scramble (a gift from David Sabatini, Addgene plasmid # 1864) (39) were produced according to the protocol described by Wang et al. (40).

Statistical analysis

The data were presented with mean and standard deviation (SD). *P*-values were calculated using unpaired two-tailed Student's *t*-test for two groups' comparison and one-way ANOVA for multiple groups' comparison. *P*-values <0.05 were considered to be statistically significant.

Acknowledgements:

This work was supported by grants from the US National Institutes of Health (R01AR071649) and the National Institute of Food and Agriculture (NC-1184). We thank Jessie Ellis for technical assistance; Jun Wu and Mary Larimore for mouse colony maintenance and members of the Kuang laboratory for valuable comments.

Competing Interests:

The authors declare that they have no conflicts of interest with the contents of this article.

Author contributions:

C.W. and S.K. conceived the project, designed the experiments and prepared the manuscript. C.W., A.R., J.A., J.C., H.H., Y.N., and F.Y. performed the experiments and analyzed the data. W.T., and C.H. provided key reagents and technical assistance.

REFERENCES

1. Schiaffino, S., and Reggiani, C. (2011) Fiber types in mammalian skeletal muscles. *Physiol Rev* **91**, 1447-1531
2. Potthoff, M. J., Wu, H., Arnold, M. A., Shelton, J. M., Backs, J., McAnally, J., Richardson, J. A., Bassel-Duby, R., and Olson, E. N. (2007) Histone deacetylase degradation and MEF2 activation promote the formation of slow-twitch myofibers. *J Clin Invest* **117**, 2459-2467
3. Liu, N., Nelson, B. R., Bezprozvannaya, S., Shelton, J. M., Richardson, J. A., Bassel-Duby, R., and Olson, E. N. (2014) Requirement of MEF2A, C, and D for skeletal muscle regeneration. *Proc Natl Acad Sci U S A* **111**, 4109-4114
4. Potthoff, M. J., Arnold, M. A., McAnally, J., Richardson, J. A., Bassel-Duby, R., and Olson, E. N. (2007) Regulation of skeletal muscle sarcomere integrity and postnatal muscle function by Mef2c. *Mol Cell Biol* **27**, 8143-8151
5. Wu, H., and Olson, E. N. (2002) Activation of the MEF2 transcription factor in skeletal muscles from myotonic mice. *J Clin Invest* **109**, 1327-1333
6. Wust, S., Droese, S., Heidler, J., Wittig, I., Klockner, I., Franko, A., Bonke, E., Gunther, S., Gartner, U., Boettger, T., and Braun, T. (2018) Metabolic Maturation during Muscle Stem Cell Differentiation Is Achieved by miR-1/133a-Mediated Inhibition of the Dlk1-Dio3 Mega Gene Cluster. *Cell metabolism* **27**, 1026-1039 e1026
7. Passier, R., Zeng, H., Frey, N., Naya, F. J., Nicol, R. L., McKinsey, T. A., Overbeek, P., Richardson, J. A., Grant, S. R., and Olson, E. N. (2000) CaM kinase signaling induces cardiac hypertrophy and activates the MEF2 transcription factor in vivo. *J Clin Invest* **105**, 1395-1406
8. Wu, H., Rothmel, B., Kanatous, S., Rosenberg, P., Naya, F. J., Shelton, J. M., Hutcheson, K. A., DiMaio, J. M., Olson, E. N., Bassel-Duby, R., and Williams, R. S. (2001) Activation of MEF2 by muscle activity is mediated through a calcineurin-dependent pathway. *EMBO J* **20**, 6414-6423
9. Ornatsky, O. I., Cox, D. M., Tangirala, P., Andreucci, J. J., Quinn, Z. A., Wrana, J. L., Prywes, R., Yu, Y. T., and McDermott, J. C. (1999) Post-translational control of the MEF2A transcriptional regulatory protein. *Nucleic Acids Res* **27**, 2646-2654
10. Yang, C. C., Ornatsky, O. I., McDermott, J. C., Cruz, T. F., and Prody, C. A. (1998) Interaction of myocyte enhancer factor 2 (MEF2) with a mitogen-activated protein kinase, ERK5/BMK1. *Nucleic Acids Res* **26**, 4771-4777
11. McKinsey, T. A., Zhang, C. L., Lu, J., and Olson, E. N. (2000) Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. *Nature* **408**, 106-111
12. Chang, S., McKinsey, T. A., Zhang, C. L., Richardson, J. A., Hill, J. A., and Olson, E. N. (2004) Histone deacetylases 5 and 9 govern responsiveness of the heart to a subset of stress signals and play redundant roles in heart development. *Mol Cell Biol* **24**, 8467-8476
13. Eskelinen, E. L., and Saftig, P. (2009) Autophagy: a lysosomal degradation pathway with a central role in health and disease. *Biochim Biophys Acta* **1793**, 664-673
14. Yang, Q., She, H., Gearing, M., Colla, E., Lee, M., Shacka, J. J., and Mao, Z. (2009) Regulation of neuronal survival factor MEF2D by chaperone-mediated autophagy. *Science* **323**, 124-127
15. Zhang, L., Sun, Y., Fei, M., Tan, C., Wu, J., Zheng, J., Tang, J., Sun, W., Lv, Z., Bao, J., Xu, Q., and Yu, H. (2014) Disruption of chaperone-mediated autophagy-dependent degradation of MEF2A by oxidative stress-induced lysosome destabilization. *Autophagy* **10**, 1015-1035
16. Lanouette, S., Mongeon, V., Figeys, D., and Couture, J. F. (2014) The functional diversity of protein lysine methylation. *Mol Syst Biol* **10**, 724
17. Hamamoto, R., Saloura, V., and Nakamura, Y. (2015) Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nat Rev Cancer* **15**, 110-124

18. Black, J. C., Van Rechem, C., and Whetstine, J. R. (2012) Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell* **48**, 491-507
19. Biggar, K. K., and Li, S. S. (2015) Non-histone protein methylation as a regulator of cellular signalling and function. *Nat Rev Mol Cell Biol* **16**, 5-17
20. Falnes, P. O., Jakobsson, M. E., Davydova, E., Ho, A., and Malecki, J. (2016) Protein lysine methylation by seven-beta-strand methyltransferases. *Biochem J* **473**, 1995-2009
21. Cloutier, P., Lavalley-Adam, M., Faubert, D., Blanchette, M., and Coulombe, B. (2013) A newly uncovered group of distantly related lysine methyltransferases preferentially interact with molecular chaperones to regulate their activity. *PLoS Genetics* **9**
22. Huang, J., Hsu, Y. H., Mo, C., Abreu, E., Kiel, D. P., Bonewald, L. F., Brotto, M., and Karasik, D. (2014) METTL21C is a potential pleiotropic gene for osteoporosis and sarcopenia acting through the modulation of the NF-kappaB signaling pathway. *J Bone Miner Res* **29**, 1531-1540
23. Wiederstein, J. L., Nolte, H., Gunther, S., Piller, T., Baraldo, M., Kostin, S., Bloch, W., Schindler, N., Sandri, M., Blaauw, B., Braun, T., Holper, S., and Kruger, M. (2018) Skeletal Muscle-Specific Methyltransferase METTL21C Trimethylates p97 and Regulates Autophagy-Associated Protein Breakdown. *Cell Rep* **23**, 1342-1356
24. Ju, J. S., Fuentealba, R. A., Miller, S. E., Jackson, E., Piwnicka-Worms, D., Baloh, R. H., and Weihl, C. C. (2009) Valosin-containing protein (VCP) is required for autophagy and is disrupted in VCP disease. *The Journal of cell biology* **187**, 875-888
25. Wang, C., Zhang, B., Ratliff, A. C., Arrington, J., Chen, J., Xiong, Y., Yue, F., Nie, Y., Hu, K., Jin, W., Tao, W. A., Hrycyna, C. A., Sun, X., and Kuang, S. (2019) Methyltransferase-like 21e inhibits 26S proteasome activity to facilitate hypertrophy of type IIb myofibers. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, fj201900582R
26. Zhang, H., Amick, J., Chakravarti, R., Santarriaga, S., Schlanger, S., McGlone, C., Dare, M., Nix, J. C., Scaglione, K. M., Stuehr, D. J., Misra, S., and Page, R. C. (2015) A bipartite interaction between Hsp70 and CHIP regulates ubiquitination of chaperoned client proteins. *Structure* **23**, 472-482
27. Massey, A. C., Kaushik, S., Sovak, G., Kiffin, R., and Cuervo, A. M. (2006) Consequences of the selective blockage of chaperone-mediated autophagy. *Proc Natl Acad Sci U S A* **103**, 5805-5810
28. Jakobsson, M. E., Moen, A., Bousset, L., Egge-Jacobsen, W., Kernstock, S., Melki, R., and Falnes, P. O. (2013) Identification and characterization of a novel human methyltransferase modulating Hsp70 protein function through lysine methylation. *J Biol Chem* **288**, 27752-27763
29. Cho, H. S., Shimazu, T., Toyokawa, G., Daigo, Y., Maehara, Y., Hayami, S., Ito, A., Masuda, K., Ikawa, N., Field, H. I., Tsuchiya, E., Ohnuma, S., Ponder, B. A., Yoshida, M., Nakamura, Y., and Hamamoto, R. (2012) Enhanced HSP70 lysine methylation promotes proliferation of cancer cells through activation of Aurora kinase B. *Nature communications* **3**, 1072
30. Stricher, F., Macri, C., Ruff, M., and Muller, S. (2013) HSPA8/HSC70 chaperone protein: structure, function, and chemical targeting. *Autophagy* **9**, 1937-1954
31. Gamerding, M., Hajieva, P., Kaya, A. M., Wolfrum, U., Hartl, F. U., and Behl, C. (2009) Protein quality control during aging involves recruitment of the macroautophagy pathway by BAG3. *EMBO J* **28**, 889-901
32. Tahrir, F. G., Knezevic, T., Gupta, M. K., Gordon, J., Cheung, J. Y., Feldman, A. M., and Khalili, K. (2017) Evidence for the Role of BAG3 in Mitochondrial Quality Control in Cardiomyocytes. *J Cell Physiol* **232**, 797-805
33. Arndt, V., Dick, N., Tawo, R., Dreiseidler, M., Wenzel, D., Hesse, M., Furst, D. O., Saftig, P., Saint, R., Fleischmann, B. K., Hoch, M., and Hohfeld, J. (2010) Chaperone-assisted selective autophagy is essential for muscle maintenance. *Curr Biol* **20**, 143-148

34. Page, N., Gros, F., Schall, N., Decossas, M., Bagnard, D., Briand, J. P., and Muller, S. (2011) HSC70 blockade by the therapeutic peptide P140 affects autophagic processes and endogenous MHCII presentation in murine lupus. *Ann Rheum Dis* **70**, 837-843
35. Leu, J. I., Pimkina, J., Frank, A., Murphy, M. E., and George, D. L. (2009) A small molecule inhibitor of inducible heat shock protein 70. *Mol Cell* **36**, 15-27
36. Cuervo, A. M., and Wong, E. (2014) Chaperone-mediated autophagy: roles in disease and aging. *Cell Res* **24**, 92-104
37. Wang, C., Yue, F., and Kuang, S. (2017) Muscle Histology Characterization Using H&E Staining and Muscle Fiber Type Classification Using Immunofluorescence Staining. *Bio-Protocol* **7**, e2279
38. Wang, C., Wang, M., Arrington, J., Shan, T., Yue, F., Nie, Y., Tao, W. A., and Kuang, S. (2017) Ascl2 inhibits myogenesis by antagonizing the transcriptional activity of myogenic regulatory factors. *Development* **144**, 235-247
39. Sarbassov, D. D., Guertin, D. A., Ali, S. M., and Sabatini, D. M. (2005) Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307**, 1098-1101
40. Wang, C., Liu, W., Liu, Z., Chen, L., Liu, X., and Kuang, S. (2015) Hypoxia inhibits myogenic differentiation through p53 protein-dependent induction of bhlhe40 protein. *Journal of Biological Chemistry* **290**, 29707-29716

Figure Legends

Figure 1. Dynamic expression of *Mettl21c* during muscle development and myogenesis

(A) Genetic targeting of *Mettl21c* through insertion of *LacZ* into exons 2-5. Grey box indicate the exon of *Mettl21c* gene, and blue box indicate *LacZ* gene. (B) Immunostaining of Myh7 (type I myofibers, Red) and Myh2 (type IIa myofibers, Green), and X-gal staining of serial sections of SOL muscles isolated from *Mettl21c^{LacZ/+}* mice. Scale bar: 100 μ m. (C-E) Immunostaining of Myh7 (Red) and Dystrophin (Green), and X-gal staining of serial sections of SOL muscles of P3 (C), P7 (D) and P10 (E) *Mettl21c^{LacZ/+}* mice. Scale bar: 100 μ m. (F) qPCR analysis of *Mettl21c* in different developmental stage. Error bars represent mean+s.d. of 6 mice with three technical repeats. Different letters indicates $p < 0.05$ (One-way ANOVA). (G) qPCR analysis of *Mettl21c* in myocytes differentiated for 3 days or SOL muscles isolated from *Mettl21c^{LacZ/+}* (HET) or *Mettl21c^{LacZ/LacZ}* (KO) mice. Error bars represent mean+s.d. of 5 independent biological experiments with three technical repeats. (H) Immunostaining of MF20 (Myosin heavy chain, Red) and DAPI (cyan) in differentiated myoblasts isolated from *Mettl21c^{LacZ/+}* mice at different days after differentiation induction. Scale bar: 50 μ m.

Figure 2. *Mettl21c* is required for proper muscle function

(A) Relative body weight and (B) SOL muscle weight of HET and KO mice. (C) Myofiber numbers in SOL muscles of HET and KO mice. (D) Myofiber size in SOL muscles of HET and KO mice. (E) Immunostaining of Myh7 (Red) and Myh2 (Green) in SOL muscles of HET or KO mice. Scale bar: 500 μ m. (F) Quantification of different types of myofibers in SOL muscles of HET or KO mice. (G) Running speed and distance of HET and KO mice in treadmill test. Error bars represent mean+s.d. of 7 mice. * indicates $p < 0.05$ (Student's t-test).

Figure 3. *Mettl21c* binds and methylates Hspa8

(A) Coomassie stained SDS-PAGE of proteins pulled down by the FLAG antibody from myocytes transduced with *Mettl21c-FLAG* or *GFP-FLAG* adenoviral vectors. (B) Mass spectrometry analysis of *Mettl21c* binding proteins. (C) Detection of Hspa8 but not Hspa5 in *Mettl21c-FLAG* protein complex after immunoprecipitation from myocytes. The lysates of myocytes was set as a positive control to indicate the position of Hspa8 and Hspa5. (D and E) *In vitro* methyltransferase activity of Vcp (D) and Hspa8 (E) as substrates of *Mettl21c*. *Mettl21c* was held constant at 1 μ M in the presence of 13 μ M [14 C] SAM and increasing concentrations of either Vcp or Hspa8 as the substrate. The data are represented as a percentage of CPM (counts per minute, representing methyl groups transferred) at 1 μ M of substrate. Error bars represent mean+s.d. of 3 experiments performed in duplicate.

Figure 4. *Mettl21c* methylates Hspa8 at K561

(A) Flow chat of the analysis of methylated peptides. (B) Chromatogram view for the trimethylated Hspa8 peptide in myoblasts transduced with *Mettl21c-GFP* (*Mettl21c^{OE}*) or *GFP* only (as control) adenoviral vectors. The MS/MS fragmentation pattern of Hspa8 peptide is highlighted. The b- and y-ions supports trimethylation of Hspa8 at K561. (C and D) *In vitro* methyltransferase activity of *Mettl21c* and mutated Hspa8 (C) or WT Hspa8 (D). The substrate (K561A-Hspa8 or WT Hspa8) was held constant at 4 μ M in the presence of 13 μ M [14 C] SAM and increasing concentrations of *Mettl21c*. The data are represented as CPM. Error bars represent mean+s.d. of 3 experiments performed in duplicate. (E) Chromatogram view for the di- or trimethylated Hspa8 peptide in SOL muscles of HET or KO mice. MS/MS fragmentation pattern of Hspa8 peptide.

Figure 5. K561 methylation stabilizes Hspa8

(A) Western blots analysis of mutated Hspa8 after cycloheximide-induced inhibition of translation. (B) Western blots analysis showing the effect of Hspa8 overexpression (*Hspa8^{OE}*) on Mef2A and Mef2D proteins in myotubes. Primary myoblasts were induced to differentiate for 2 days and then were transduced with *Hspa8^{OE}* or control adenoviral vectors for 3 more days. (C) Luciferase assays showing

the effects of Hspa8 on the transcriptional activity of Mef2A, Mef2C and Mef2D using the 3xMEF2-luc reporter. Error bars represent mean+s.d. of 5 independent biological experiments. * indicates $p < 0.05$ (Student's *t*-test). (D) Western blots analysis of indicated proteins in sarcoplasm and myonuclei proteins of SOL muscles from HET and KO mice. (E) qPCR analysis of *Mettl21c* in differentiated myoblasts transduced with *Mettl21c-GFP* (*Mettl21c^{OE}*) or *GFP* only (as control) adenoviral vectors. Primary myoblasts were induced to differentiate for 2 days and then were transduced with *Mettl21c^{OE}* or control adenoviral vectors for 3 more days. Error bars represent mean+s.d. of 5 independent biological experiments. * indicates $p < 0.05$ (Student's *t*-test). (F) Western blots analysis of indicated proteins in control and *Mettl21c^{OE}* myotubes. (G) Western blots analysis of indicated proteins in control and *Mettl21c^{OE}* myotubes after Lamp2 knockdown. Primary myoblasts were treated with lentivirus for 2 days and selected in puromycin (1 μ g/ml) for 2 additional days. These primary myoblasts were then transduced with *Mettl21c^{OE}* or control adenoviral vectors. Two days after adenoviral transduction, the myoblasts were induced to differentiate for 3 days. (H) Diagram summarizing the role of *Mettl21c* in type I myofibers. *Mettl21c* methylates Hspa8 at K561 to interrupt its binding with an E3-ligase CHIP, and to prevent its ubiquitination and proteasome-dependent degradation. The surviving Hspa8 carries Mef2A and Mef2D to lysosome for degradation in chaperone mediated autophagy pathway.

Figure 6. Expression of *Mettl21c* during muscle regeneration

(A and B) Immunostaining of Myh7 (Red), Myh2 (Green) and Dystrophin (Blue), and X-gal staining of serial sections of SOL muscles at day 7 (A) and day 14 (B) after CTX injection. Scale bar: 100 μ m. (C) qPCR analysis of *Mettl21c* at different days during SOL muscle regeneration. Error bars represent mean+s.d. of 6 mice with three technical repeats. Different letters indicates $p < 0.05$ (One-way ANOVA). (D) H&E staining of SOL muscles from HET and KO mice 7 days after CTX injection. Scale bar: 100 μ m.

Fig.1

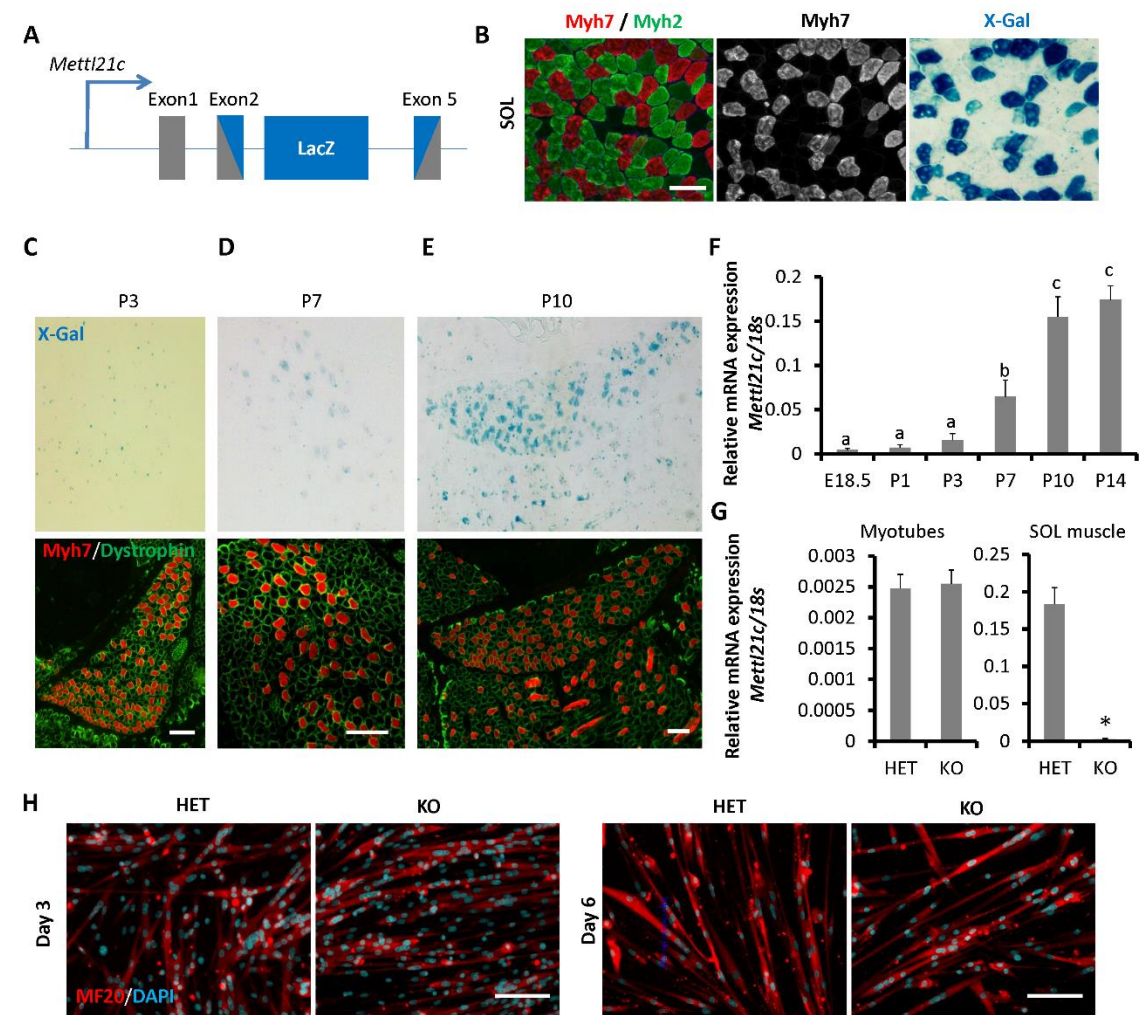
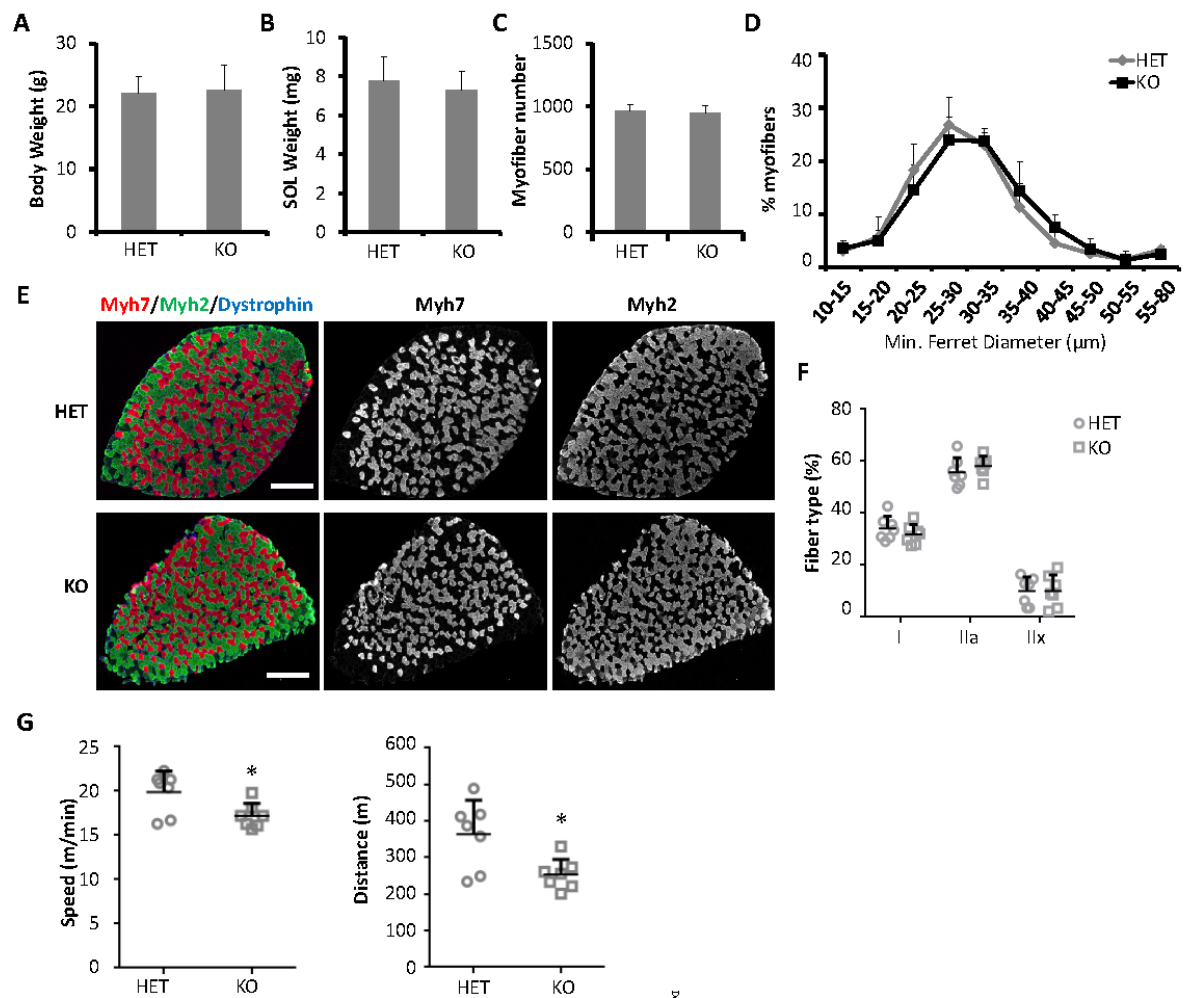


Fig. 2



Downloaded from <http://www.jk-oxford.com/academic> at 2019

Fig.3

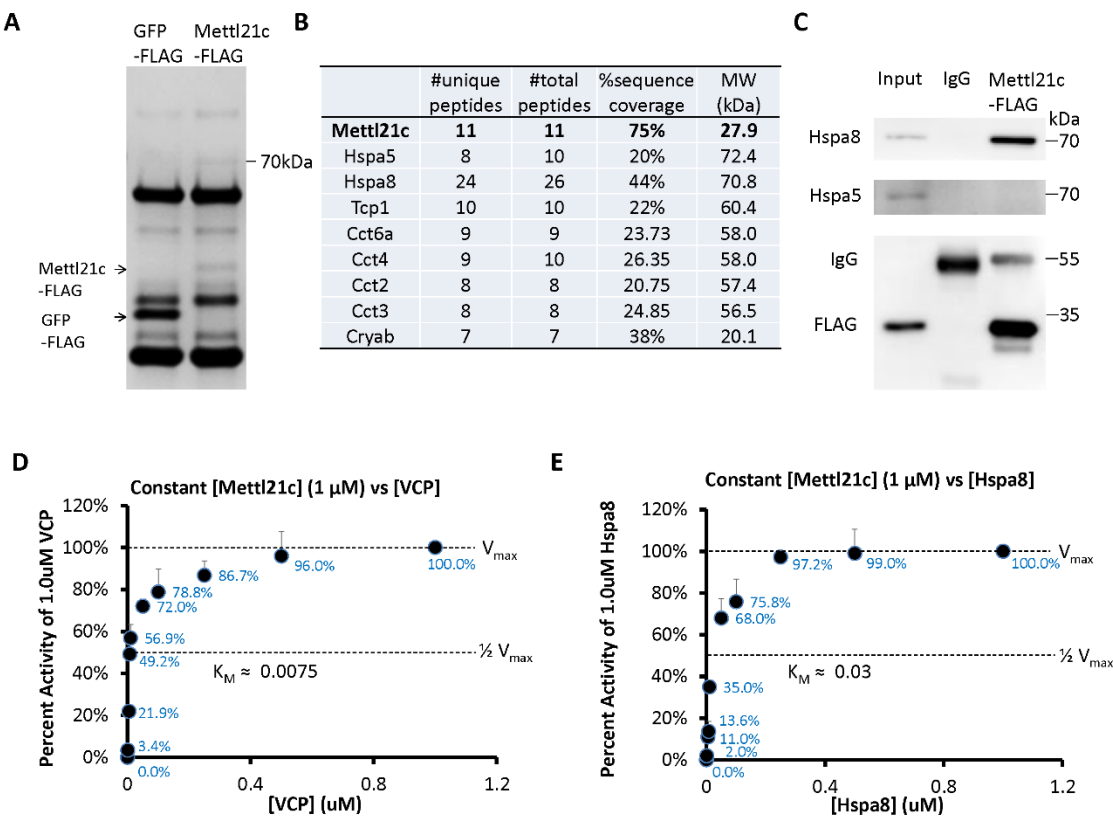


Fig.4

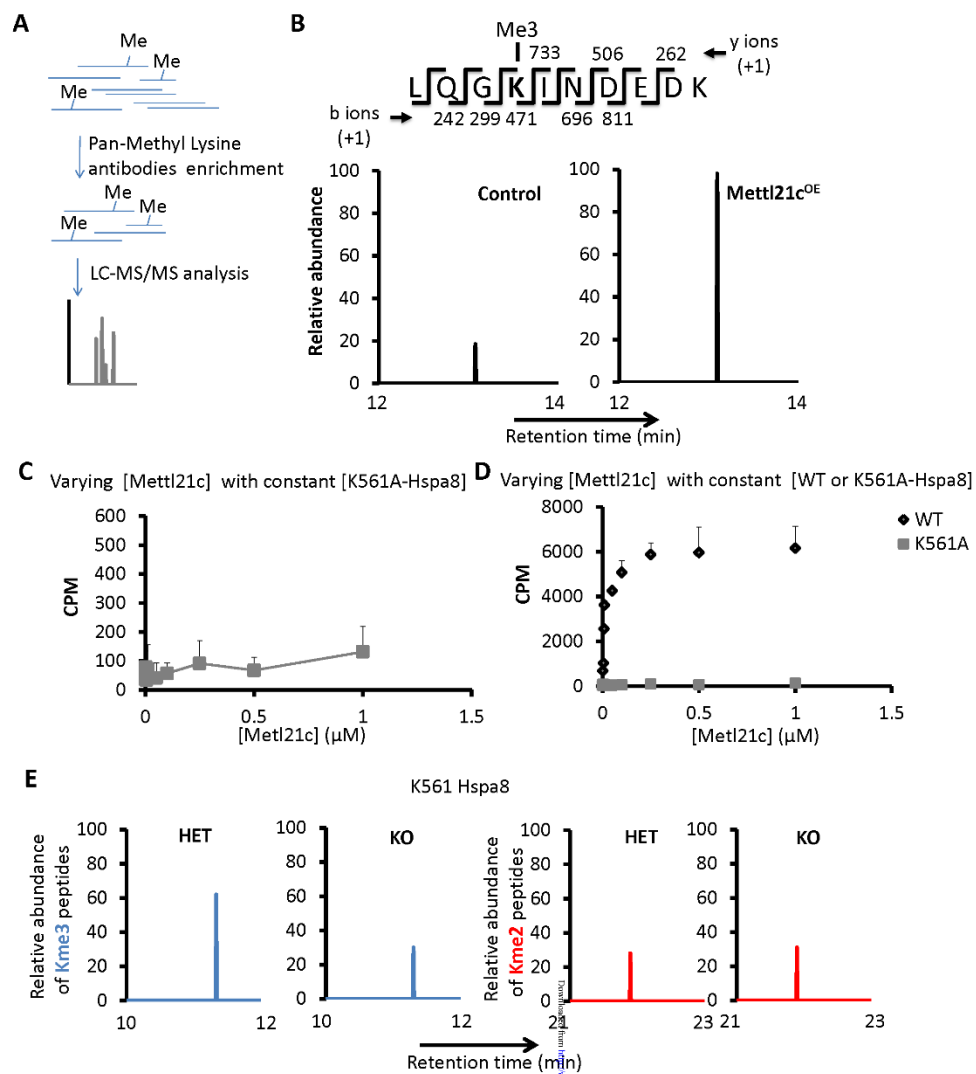


Fig.5

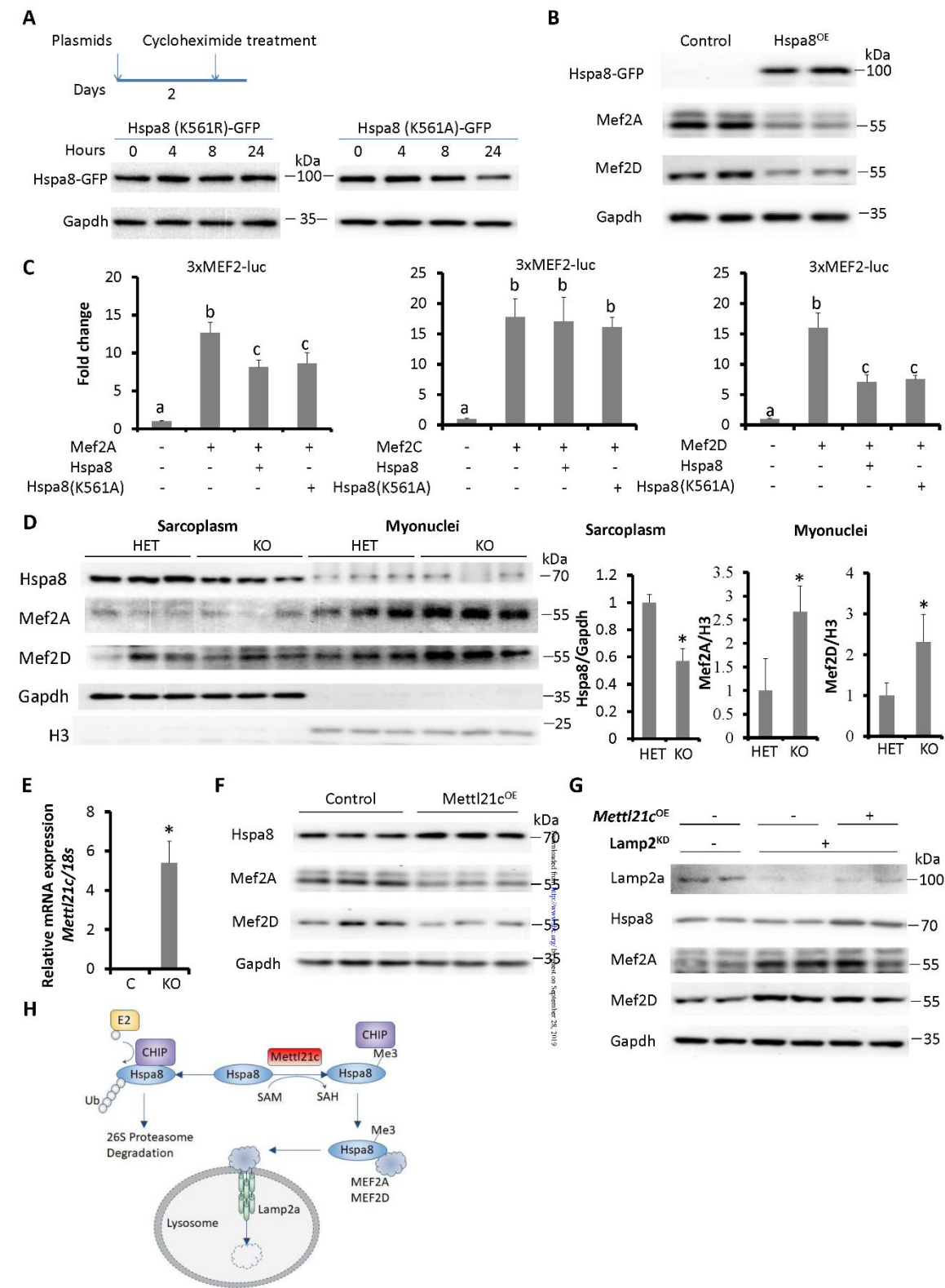
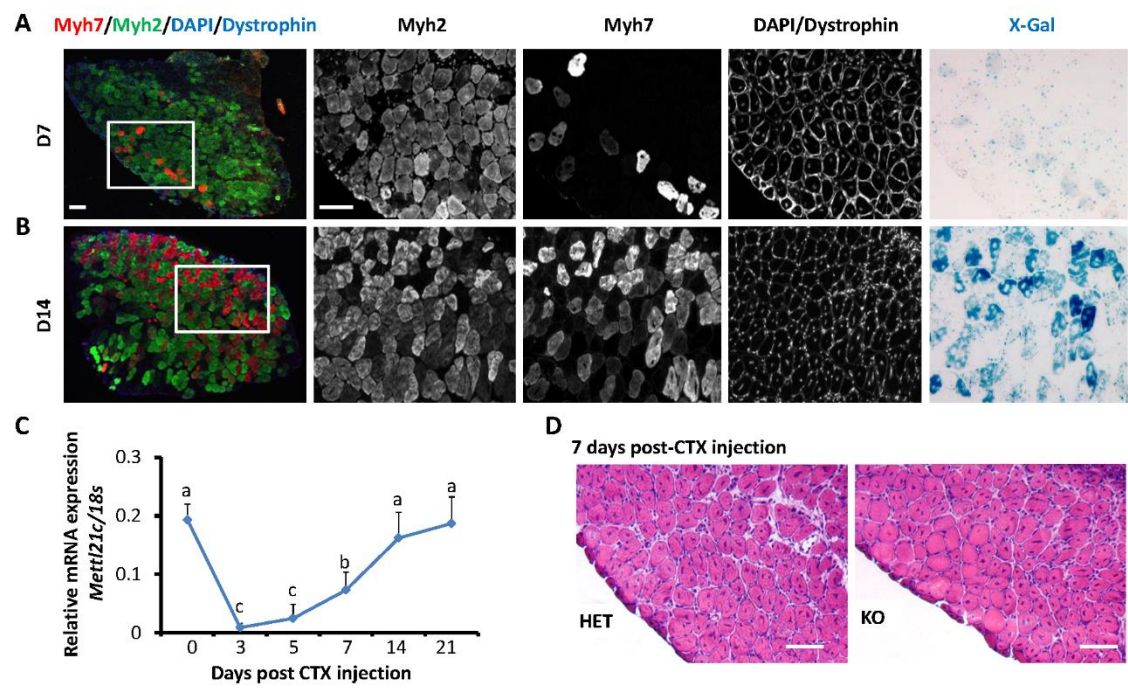


Fig.6



Methyltransferase-like 21c methylates and stabilizes the heat shock protein HSPA8 in type I myofibers in mice

Chao Wang, Justine Arrington, Anna C Ratliff, Jingjuan Chen, Hannah E. Horton, Yaohui Nie, Feng Yue, Christine A. Hrycyna, W. Andy Tao and Shihuan Kuang

J. Biol. Chem. published online July 25, 2019

Access the most updated version of this article at doi: [10.1074/jbc.RA119.008430](https://doi.org/10.1074/jbc.RA119.008430)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

Downloaded from <http://www.jbc.org/> by guest on September 28, 2019