Role of the STAS domain of the *E. coli* anion transporter YchM

by

Rediet Taddese

A thesis submitted in conformity with the requirements for the degree of Masters in Science
Graduate Department of Biochemistry
University of Toronto

© Copyright by Rediet Taddese (2013)
Role of the STAS domain of the *E. coli* anion transporter YchM

Rediet Taddese

Masters of Science

Graduate Department of Biochemistry
University of Toronto

2013

Abstract

YchM is the sole *E. coli* member of the SLC26 superfamily of anion transporters, which are characterized by an N-terminal transport domain and a C-terminal cytosolic STAS (Sulphate Transporter and Anti-Sigma factor antagonist) domain. In a previous study, the STAS domain of YchM co-purified and crystallized with acyl carrier protein (ACP). In this study, analysis of the ACP-STAS interaction using isothermal titration calorimetry (ITC) showed that the 4’phosphopantetheine of ACP and R523 and R527 of the STAS are crucial for binding. The binding constant for the ACP-STAS interaction was found to be 0.7 +/- 0.1 µM. The potential role of YchM for pH regulation and fatty acid degradation studied *in vivo* indicated that a) YchM does not provide selective advantage for growth in alkaline pH and b) YchM was not essential for cell growth, even when fatty acids were the sole carbon source.
Acknowledgments

This work would not have been possible without the support and encouragement of several individuals. I’m extremely grateful to my supervisor, Dr. Trevor Moraes, for providing me with the opportunity to work in his laboratory, as well as his time and unreserved energy to help me through out my project. I appreciate his enthusiasm for science and genuine care for his graduate students. I thank my co-supervisor, Dr. Reinhart Reithmeier, for his intellectual insights and optimistic outlook that made everything seem feasible. I am grateful to my committee members, Dr. Walid Houry and Dr. Julie Forman-Kay, for their helpful inputs and for challenging me to perfect my critical thinking skills.

I thank past and present members of the Moraes Lab, especially Dr. Charles Calmettes, Christine Lai, Anastassia Pogoutse, Andrew Judd, David Shen, David Guo, Brandon Sit and Anastasia Bosc for their help during my project and their role in creating a friendly atmosphere where we could share ideas (scientific or otherwise) freely.

I would like to acknowledge all members of the Reithmeier, Houry and Rini laboratories for their technical assistance.

I extend my deep gratitude to my parents for their unconditional love and unwavering support. They not only poured everything they had in me, but continue to believe in my ideas and dreams.

To my sister, Yamrot Taddese, cousin Hiruth Seifu and friends (esp Irsa Ademi, Mahlet Getachew, Sofia Huroy, Hanin Issa, Saron Marcos and Hiwot Telaye) - thank you for lending your time, ears and shoulders when needed, being my alarm clocks, editing my work, helping me find the humor during bad days and studying/writing with me. I certainly could not have done it without you.

Lastly, this thesis is dedicated to the beautiful soul I had the privilege of having as a second mother, cheerleader and advisor-in-chief, Aunt Amsale.
# Table of Contents

Acknowledgments........................................................................................................................... iii
List of abbreviations ............................................................................................................................ vi
List of Figures .......................................................................................................................................... viii
List of Tables ........................................................................................................................................ ix
List of Appendices ............................................................................................................................... x

Chapter 1 Introduction .......................................................................................................................... 1
  1.1 Membrane Transport Proteins ...................................................................................................... 2
  1.2 Solute Carriers (SLC) .................................................................................................................... 3
  1.3 The Solute Carrier 26 and Sulfate Permease (SLC26/SulP) Super Family .................................... 4
    1.3.1 The human Solute Carrier 26 Transporters ........................................................................... 4
    1.3.2 Medical significance of SLC26 transporters .......................................................................... 7
    1.3.3 Defining SLC26 features using a prokaryotic model: BicA ..................................................... 9
  1.4 The Sulfate Transporter and Anti-Sigma factor antagonist (STAS) domain ......................... 12
    1.4.1 Bacterial sigma factors, anti-sigma factors and anti-sigma factor antagonists .................... 12
    1.4.2 Functions of the STAS domains of different SLC26 transporters ...................................... 13
    1.4.4 Structural comparisons of STAS domains of mammalian and bacterial STAS domains .... 18
  1.5 YchM: E.coli’s sole member of the SLC26 family ..................................................................... 21
    1.5.1 The STAS (YchM)-ACP interaction: A potential link between bicarbonate transport and fatty acid synthesis .......................................................................................................................... 23
  1.6 Hypothesis and Project Objective .............................................................................................. 31

Chapter 2 Methods and Materials ..................................................................................................... 33
  2.1 Materials ...................................................................................................................................... 34
  2.2 In vitro analysis of the ACP-STAS interaction .......................................................................... 35
    2.2.1 Plasmids and Cloning ............................................................................................................ 35
    2.2.2 Protein expression and purification ...................................................................................... 36
    2.2.3 Malonyl-ACP Synthesis ...................................................................................................... 38
    2.2.4 SDS-PAGE, Native-PAGE and Western Blot ....................................................................... 38
    2.2.5 Circular dichroism spectroscopy and thermal denaturation assay .................................. 39
    2.2.6 Guanidine denaturation assay ............................................................................................. 39
    2.2.7 Surface Plasmon Resonance .................................................................................................. 40
    2.2.8 Affinity Pull-down ................................................................................................................. 40
    2.2.9 Bio-layer interferometry ....................................................................................................... 41
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Acetyl-CoA carboxylase</td>
</tr>
<tr>
<td>ACP</td>
<td>Acyl carrier protein</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BLI</td>
<td>Bio-Layer Interferometry</td>
</tr>
<tr>
<td>CO₂</td>
<td>Carbon dioxide</td>
</tr>
<tr>
<td>CCM</td>
<td>Carbon dioxide concentrating mechanism</td>
</tr>
<tr>
<td>CD</td>
<td>Circular Dichroism</td>
</tr>
<tr>
<td>CoA</td>
<td>Co-enzyme A</td>
</tr>
<tr>
<td>CFTR</td>
<td>Cystic fibrosis transmembrane conductance regulator</td>
</tr>
<tr>
<td>CLD</td>
<td>Chloride losing diarrhea</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTD</td>
<td>Dystrophic dysplasia</td>
</tr>
<tr>
<td>DRA</td>
<td>Down regulated in adenoma</td>
</tr>
<tr>
<td>ECL</td>
<td>Enhanced chemiluminescence</td>
</tr>
<tr>
<td>EDC</td>
<td>N-ethyl-N(dimethylaminopropyl) carbodiimide</td>
</tr>
<tr>
<td>FAB</td>
<td>Fatty acid biosynthesis</td>
</tr>
<tr>
<td>FAD</td>
<td>Fatty acid degradation</td>
</tr>
<tr>
<td>FMN</td>
<td>Flavin mononucleotide</td>
</tr>
<tr>
<td>GDP</td>
<td>Guanosine-5’-diphosphate</td>
</tr>
<tr>
<td>GPCR</td>
<td>G-protein coupled receptors</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione S-transferase</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine-5’-triphosphate</td>
</tr>
<tr>
<td>HUGO</td>
<td>Human Genome Organisation</td>
</tr>
<tr>
<td>HGNC</td>
<td>The Human Genome Organisation Nomenclature Committee</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>ITC</td>
<td>Isothermal titration calorimetry</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Broth</td>
</tr>
<tr>
<td>LOV</td>
<td>Light, Oxygen and Voltage</td>
</tr>
<tr>
<td>NHS</td>
<td>N-hydroxysuccinimide</td>
</tr>
<tr>
<td>OAS</td>
<td>O-acetyl serine</td>
</tr>
<tr>
<td>OASTL</td>
<td>O-acetyl serine (thiol) lyase</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffer saline</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinyl difluoride</td>
</tr>
<tr>
<td>eSGA</td>
<td><em>E.coli</em> synthetic genetic array</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SHST</td>
<td><em>Stylosanthes hamata</em> sulfate transporter</td>
</tr>
<tr>
<td>SLC</td>
<td>Solute carriers</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface Plasmon Resonance</td>
</tr>
<tr>
<td>STAS</td>
<td>Sulfate Transporter and Anti-Sigma factor antagonist</td>
</tr>
<tr>
<td>SulP</td>
<td>Sulphate Permease</td>
</tr>
<tr>
<td>TC</td>
<td>Transport Classification</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris (2-carboxyethyl) phosphine</td>
</tr>
</tbody>
</table>
**List of Figures**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Structural alignments of bacterial and mammalian STAS domains</td>
</tr>
<tr>
<td>2</td>
<td>Predicted topology of <em>E.coli</em>’s SLC26 transporter, YchM</td>
</tr>
<tr>
<td>3</td>
<td>The crystal structure of the YchM-STAS and ACP complex</td>
</tr>
<tr>
<td>4</td>
<td>Structural alignment of holo-ACP and dodecanoyl-ACP with the STAS(YchM)-ACP complex</td>
</tr>
<tr>
<td>5</td>
<td>Bacterial fatty acid biosynthesis initiation and condensation</td>
</tr>
<tr>
<td>6</td>
<td>Sequence alignment of <em>B. subtilis</em> SpoIIAA and <em>E. coli</em> YchM.</td>
</tr>
<tr>
<td>7</td>
<td>Purification of GST-STAS and STAS domain by size exclusion chromatography</td>
</tr>
<tr>
<td>8</td>
<td>Secondary structure analysis of the STAS domain</td>
</tr>
<tr>
<td>9</td>
<td>The STAS domain exhibits co-operative folding.</td>
</tr>
<tr>
<td>10</td>
<td>Native-PAGE of apo-ACP, holo-ACP dimer and holo-ACP monomer.</td>
</tr>
<tr>
<td>11</td>
<td>The binding constant for the ACP-STAS interaction is predicted to be in the micro-molar range.</td>
</tr>
<tr>
<td>12</td>
<td>Co-purification of GST-STAS with ACP</td>
</tr>
<tr>
<td>13</td>
<td>Holo-ACP does not alter the tryptophan emission spectra of the STAS domain.</td>
</tr>
<tr>
<td>14</td>
<td>Isothermal titration calorimetry indicates a $K_d = 0.6 \mu M$ for the STAS-holoACP interaction.</td>
</tr>
<tr>
<td>15</td>
<td>Isothermal titration calorimetry shows that the R523E STAS mutant does not interact with holo-ACP.</td>
</tr>
<tr>
<td>16</td>
<td>The role of YchM in <em>E. coli</em> growth under acidic and alkaline conditions.</td>
</tr>
<tr>
<td>17</td>
<td>The role of YchM in <em>E. coli</em> growth with fatty acids as the sole carbon source.</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table 1</th>
<th>The human SLC26 family members</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2</td>
<td>The roles of functionally characterized STAS domains of SLC26 transporters</td>
</tr>
<tr>
<td>Table 3</td>
<td>The sequence and structural similarities of STAS domains of known structures</td>
</tr>
<tr>
<td>Table 4</td>
<td>Topology of <em>E. coli</em> YchM using different prediction programs</td>
</tr>
</tbody>
</table>
## List of Appendices

| Appendix 1 | List of cloning and sequencing primers |
| Appendix 2 | Size exclusion elution profile of STAS\textsubscript{R527E} and STAS\textsubscript{R523E} |
| Appendix 3 | Thermal denaturation of the STAS domain |
| Appendix 4 | Purification of ACP by size exclusion chromatography |
| Appendix 5 | Mass spectrometry results of malonyl-ACP synthesis |
| Appendix 6 | YchM does not provide selective advantage for growth in alkaline pH |
| Appendix 7 | The role of YchM under alkaline conditions |
Chapter 1 Introduction
1.1 Membrane Transport Proteins

The maintenance of normal physiological processes within cells requires regulated entry of nutrients and ions as well as exit of metabolic end products across biological membranes. Membrane proteins accomplish this task by acting as passages or gateways for these molecules. The importance of membrane transporters is further underscored by virtue of the fact that 10% of characterized genes encode transport proteins \(^1\). Based on their mode of transport, these proteins belong to either of the two classes: channels or transporters. Channels allow the passive transport of molecules down electrochemical gradients at rates approaching the diffusion limit. Transporters, on the other hand, bind substrates that induce major conformational changes to allow transport of their substrates using electrochemical gradients as the driving force. The transporter classification (TC) system categorizes known membrane transport proteins into families, classes and subclasses based on functional and phylogenetic information (http://www.tcdb.org/). Based on this classification, membrane transport systems are further categorized into (1) pores/channels, (2) secondary carrier-type facilitators that couple proton or sodium motive force to solute transport, (3) primary active transporters depend on primary source of energy (ATP) to transport substrates against their concentration gradient, (4) group translocators that chemically alter their substrates during transport, and (5) transmembrane electron carriers \(^2\).

To date, the TC system has identified 5,600 unique proteins classified into 600 transporter families. The current study focuses on a family of anion transporters known as SLC26 that belong to the solute carrier (SLC) superfamily.
1.2 Solute Carriers (SLC)

The human solute carrier (SLC) genes encode membrane proteins that are responsible for the uptake and efflux of various solutes including sugars, amino acids and ions. They can be passive transporters, exchangers or ion-coupled transporters. The SLC family represents the second largest family of human membrane proteins next to G-protein coupled receptors (GPCRs). The Human Genome Organisation (HUGO) nomenclature committee (HGNC) online database, lists 52 families of solute carriers genes that encode more than 375 transporters. A given transporter is assigned to a family if it shares 20-25 % amino acid sequence identity with other members of the family. Each family is identified by the abbreviation, SLC for solute carrier, followed by a number corresponding to a group of genes (e.g. SLC4, solute carrier family 4). To refer to a specific gene with in a family, the family name is followed by the letter “A” and a number (e.g. SLC6A1). The 52 families identified to date are involved in diverse physiological roles such as bicarbonate transporters (SLC4), iron transporters (SLC40), sodium-glucose co-transporters (SLC5) and many others. More transporters and families are still being identified. For example, the solute carrier 26 (SLC26) was first defined in 1994 by expression cloning of Na⁺-independent sulfate transporter from rat liver. Although the Human Genome Organisation (HUGO) nomenclature used “SLC26A” to denote a human gene/protein, in the literature, the name has been used to refer to the SLC26/SulP superfamily of transporters; SulP referring to bacterial Sulfate Permeases.
1.3 The Solute Carrier 26 and Sulfate Permease (SLC26/SulP) Super Family

The solute carrier 26 (SLC26) family includes 10 anion transporters found in polarized epithelial cells in humans and their homologous sulfate permeases (SulP) found in archaea, bacteria, fungi and plants. A phylogenetic analysis of 33 representative members of the SLC26/SulP family from Bacillus subtilis, Synechocystis sp., Saccharomyces cerevisiae, Arabidopsis thaliana, Caenorhabditis elegans and Homo sapiens initially characterized all of them as sulfate transporters, and they fell under the sulfate permease (SulP) family. It is now clear that these transporters are also involved in the transport of monovalent and divalent anions such as chloride (Cl\(^{-}\)), bicarbonate (HCO\(_3^{-}\)), sulfate (SO\(_4^{2-}\)), formate and oxalate. SLC26/SulP members typically contain a hydrophobic transmembrane domain that is predicted to span the membrane 10-14 times. Most transporters of the family also possess a C-terminal cytosolic domain known as STAS (Sulfate Transporter and Anti-Sigma factor antagonist) that shares a weak but significant homology with the B. subtilis SpoIIAA.

1.3.1 The human Solute Carrier 26 Transporters

The mammalian SLC26 transporters represent a subgroup of the superfamily that have been functionally characterized. The human solute carrier 26 proteins are encoded by eleven genes but SLC26A10 is predicted to be a pseudogene. These transporters are large proteins that operate as anion exchangers or channels except SLC26A5 (Prestin), a membrane-bound motor protein expressed in the outer hair cells. Members of the family share 21-45% sequence identity with each other. They are found in polarized epithelial cells of the kidney, pancreas, stomach, liver and other tissues, where they mediate transport across the apical or basolateral membranes. As shown in Table 1, some transporters are expressed ubiquitously, while others
display tissue specific expression. Individual paralogs exhibit different substrate specificities such that proteins like SLC26A4 transport only monovalent anions while SLC26A6 transports both monovalent and divalent anions. The variation in substrate selectivity accounts for the transporter’s role in diverse physiological processes including bicarbonate secretion, thyroid hormone synthesis, skeletal development, and transepithelial Na\(^+\)/Cl\(^-\) transport.

Based on their functions, human SLC26 members have been categorized into three major groups. The first group includes SLC26A1 and its closest paralog, SLC26A2, which are both sulfate transporters. Group 2 encompasses the Cl\(^-\)/HCO\(_3\)\(^-\) exchangers SLC26A3, SLC26A4 and SLC26A6 that operate with different stoichiometries. SLC26A3 functions as a 2:1 Cl\(^-\)/HCO\(_3\)\(^-\) exchanger, SLC26A4 as a 1:1 Cl\(^-\)/HCO\(_3\)\(^-\) exchanger, while SLC26A6 exchanges 1 Cl\(^-\) for 2 HCO\(_3\)\(^-\). The third group is made up of the ion channels, SLC26A7 and SLC26A9. There is not a sufficient amount of information on the transport mechanism of SLC26A8 and SLC26A11 to categorize them into any of the aforementioned classes. SLC26A5 functions as a motor protein in cochlea although there are recent studies that reported anion transport activity of this protein.
Table 1: The human SLC26 family members

<table>
<thead>
<tr>
<th>Transporter Name (alias)</th>
<th>Expression</th>
<th>Substrates</th>
<th>Associated pathology</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLC26A1 (Sat-1)</td>
<td>Liver, kidney, pancreas, testis</td>
<td>SO₄²⁻, oxalate</td>
<td></td>
</tr>
<tr>
<td>SLC26A2 (DTDST)</td>
<td>Rib cartilage, small intestine</td>
<td>SO₄²⁻/Cl⁻ exchanger</td>
<td>Dystrophic dysplasia</td>
</tr>
<tr>
<td>SLC26A3 (DRA)</td>
<td>Intestine, pancreas, prostate, sweat glands</td>
<td>Cl⁻/HCO₃⁻</td>
<td>Congenital chloride diarrhea</td>
</tr>
<tr>
<td>SLC26A4 (Pendrin)</td>
<td>Kidney, cochlea, thyroid, salivary gland</td>
<td>I⁻, Cl⁻/HCO₃⁻, formate</td>
<td>Pendred syndrome, hearing loss</td>
</tr>
<tr>
<td>SLC26A5 (Prestin)</td>
<td>Ear (cochlea)</td>
<td></td>
<td>Hearing loss</td>
</tr>
<tr>
<td>SLC26A6</td>
<td>Ubiquitous</td>
<td>Cl⁻/HCO₃⁻, Cl⁻/Oxalate, Cl⁻/formate</td>
<td>Mice: Impaired Cl⁻ absorption from intestinal and pancreatic duct</td>
</tr>
<tr>
<td>SLC26A7</td>
<td>Kidney, stomach, nasal epithelium, endothelial venules, epididymal ducts</td>
<td>Cl⁻</td>
<td></td>
</tr>
<tr>
<td>SLC26A8 (Tat1)</td>
<td>Testes</td>
<td>SO₄²⁻, Cl⁻, Oxalate</td>
<td></td>
</tr>
<tr>
<td>SLC26A9</td>
<td>Bronchial and alveolar epithelia, gastric surface</td>
<td>SO₄²⁻, Cl⁻, Oxalate</td>
<td></td>
</tr>
<tr>
<td>SLC26A11</td>
<td>Ubiquitous</td>
<td>SO₄²⁻</td>
<td></td>
</tr>
</tbody>
</table>

The table summarizes sites of expression, substrate specificities and associated pathologies of the human SLC26 transporters. SLC26A10, a predicted pseudogene, is not included. The table was created based on information gathered from the reviews by Dorwart et al.⁶ and Mount and Romero⁸.
The low sequence identity among the SLC26 members has made it difficult to accurately predict the number of transmembrane segments. An average hydropathy plot for 33 members of the SLC26 family showed 12 peaks of hydrophobicity\(^7\). Immunolocalization studies in SLC26A5 and SLC26A6 indicate the amino- and carboxy-terminal domains are intracellular suggesting even number of transmembrane helices (10-14 TM segments)\(^{18, 19}\). Furthermore, topology studies of BicA, an SLC26 transporter in cyanobacteria, indicated 12 transmembrane helices\(^9\) (details below).

Blue-native PAGE analysis and chemical crosslinking studies of bacterial, teleost and two mammalian SLC26 paralogues denote an evolutionarily conserved dimeric quaternary structure of the transporters\(^{20}\). This observation was supported by small angle neutron scattering (SANS) of a bacterial \textit{Yersinia enterocolitica} SLC26 transporter that indicated the transporter forms a dimer that is stabilized by the transmembrane core\(^{21}\). Currently, little is known about the structure-function relationship in this family of transporters and the molecular basis for their functional diversity remains a mystery. However, functional characterization of disease causing mutations in SLC26 family have provided some clues and have emphasized their physiological roles.

### 1.3.2 Medical significance of SLC26 transporters

The importance of SLC26 transporters is particularly highlighted by the discovery that mutations in three members of the family are associated with different human pathologies. Mutations in SLC26A2 leads to chondrodysplasia\(^{22}\); mutant forms of SLC26A3 are associated with congenital chloride losing diarrhea (CLD)\(^{23}\); and SLC26A4 is defective in Pendred’s syndrome and non-syndromic deafness\(^{14}\).
In mammals, sulfate is crucial for various cellular and metabolic processes, including the development of bones and cartilages \(^{24}\). The sulfation of proteoglycan, a major structural component of the extracellular matrix growth-plate of chondrocytes, is crucial for proper cartilage development \(^{25}\). In chondrocytes, the sulphate/chloride antiporter, SLC26A2, is the major importer of sulphate. Under-sulfation of proteoglycans, as a result of a defective SLC26A2 transporter at the membrane, has been implicated in four types of chondrodysplasias: a recessive form of multiple epiphyseal dysplasia (rMED)\(^{26}\), diastrophic dysplasia (DTD)\(^{22}\), atelosteogenesis type 2 \(^{27}\), and achondrogenesis 1B \(^{28}\). Patients with chondrodysplasia often display disproportionate short stature, generalized joint dysplasia and spinal deformation \(^{15}\).

Mutations in the Cl\(^-\)/HCO\(_3^-\) exchanger SLC26A3 cause congenital chloride losing diarrhea \(^{29}\), a rare autosomal recessive disease characterized by chronic diarrhea with excess chloride loss \(^{30}\). Under normal conditions, SLC26A3, also known as DRA (Down Regulated in Adenoma) functions in parallel with Na\(^+\)/H\(^+\) exchanger to absorb NaCl in intestinal epithelial cells. Defect in SLC26A3 mediated Cl\(^-\)/HCO\(_3^-\) exchange activity in the ileum/colon followed by disruption of Na\(^+\)/H\(^+\) transport leads to loss of NaCl and fluid \(^{13}\).

SLC26A4 (Pendrin) plays an important role for proper hearing. In 1997, the pendrin gene was identified as the gene mutated in Pendred syndrome, a pathology associated with sensoneural hearing loss and thyroid dysfunction \(^{14}\). In the inner ear, the Cl\(^-\)/HCO\(_3^-\) exchanger, Pendrin, is involved in pH regulation of the endolymphatics fluids \(^{31, 32}\). Lack of Pendrin leads to acidification of endolymph \(^{33}\) and enlargement of vestibular aqueduct, which are associated with hearing loss \(^{34}\).
1.3.3 Defining SLC26 features using a prokaryotic model: BicA

The lack of a high-resolution structure has made it difficult to determine the domain architecture of the SLC26 family of membrane proteins. Since human membrane proteins are often challenging to study, prokaryotic homologues have been crucial in gaining structural and functional insights. An important prokaryotic model for the SLC26 family has been the cyanobacteria *Synechococcus* member, BicA.

In their 2.5 billion years of existence, aquatic cyanobacteria have had to meet the demands of a changing gaseous environment and they have developed an effective carbon dioxide (CO$_2$) concentrating mechanism (CCM) in response to high carbon dioxide to oxygen ratio. The CCM is not only a growth requirement for all species of cyanobacteria, but it also contributes to an estimated 25% of the annual global primary productivity. One of the main functional units of the CCM is a protein micro-compartment within the cell known as the carboxysome, which contains carbonic anhydrase and the CO$_2$ fixing enzyme, Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase). The carboxysome accumulates inorganic carbon in close proximity to the photosynthetic enzyme. Inorganic carbon enters the cell in the form of HCO$_3^-$, a substrate of the carboxysome, or CO$_2$. Inside the carboxysome, carbonic anhydrase converts the HCO$_3^-$ to carbon dioxide and water making the former product available for photosynthesis. Since the rate of CO$_2$ diffusion in liquid is lower than it is in air, aquatic cyanobacteria have evolved CO$_2$ and HCO$_3^-$ transporters that allow them to import and accumulate these substrates in the cytosol. In 2004, Price *et al.* identified BicA, a member of the SLC26 family, as a class of Na$^+$-dependent HCO$_3^-$ transporter and a component of the CCM in marine cyanobacteria *Synechococcus* PCC7002. Even though BicA is one of six transporters that contribute to the CCM, a gain of function study showed that BicA was
sufficient to confer HCO\textsubscript{3}\textsuperscript{-} uptake activity. Another ortholog was identified in another oceanic strain, Synechococcus WH8102.\textsuperscript{40}

BicA is the first prokaryotic SLC26 member to have its topology mapped. Using alkaline phosphatase and beta-galactosidase fusion genes, BicA was predicted to contain 12 transmembrane segments with cytosolic N- and C-termini.\textsuperscript{9} Several BicA-PhoA/LacZ fusion encoding plasmids were expressed in E. coli and a ratio of alkaline phosphatase to beta-galactosidase activity was used as a topology marker. Since alkaline phosphatase and beta-galactosidase are only active in the periplasm and cytoplasm, respectively, a ratio greater than 1 indicated periplasmic location. The result predicted 12 transmembrane segments and mapped two conserved regions that might be defining features of the SLC26 family.

First, the first three helices are connected by short loops that are conserved in both spacing and sequence.\textsuperscript{9} This result was supported by a previous evolutionary trace analysis of the human SLC26 member, Prestin, and bioinformatics analysis of the legume Stylosanthes hamata member sulfate transporter (SHST, Stylosanthes hamata sulfate transporter) that had identified conserved residues in helices 1 and 2.\textsuperscript{41,42} Furthermore, the spacing between the first two helices was conserved in Prestin.\textsuperscript{41,42} Mutations of these conserved polar residues in the first two helices of SHST resulted in reduced sulfate transport activity when the mutants were expressed in yeast deficient in sulfate transport\textsuperscript{42} suggesting that conserved regions in the first two helices might be critical for transport function.

The second interesting feature of BicA was the presence of a cytoplasmic loop between helices 8 and 9 that includes a previously defined conserved motif: NSNKE\textsubscript{L}IGQGLGN (highly conserved residues underlined). Some of the Chondrodysplasia and CLD causing mutations in SLC26A2 and SLC26A3, respectively, have also been mapped to residues E282
and N291 of this loop\textsuperscript{43,44}. Since mutations that change residue charges in this cytosolic loop region affect proper folding and trafficking of the transporter, it is hypothesized to be in an intracellular regulatory region that mediates interaction with either the membrane or the STAS domain\textsuperscript{9,45}.

Another feature common to most members of the SLC26 transporters is the cytosolic domain at the carboxyl end of the transporters known as the STAS domain (see below). Some of the disease causing mutations of SLC26A3 and SLC26A4 are found in the STAS domain, underscoring their importance for proper functioning of the proteins. For example, 18% of CLD causing SLC26A3 mutations identified until 2011 are found within the STAS domain\textsuperscript{29}; some of which affect the folding and trafficking of the transporter to the plasma membrane\textsuperscript{23}. Interestingly, some SLC26 members such as the \textit{H. pylori} YchM do not possess a STAS domain raising a question of why some SLC26 members need the domain while others do not. \textit{H. pylori} exists under acidic conditions in the stomach where bicarbonate levels are low, suggesting that its SLC26 member may not be involved in bicarbonate transport. Even though the role of the STAS domain of BicA and most SLC26 transporters is not known, there have been a few functional and structural characterizations of STAS domains conducted in bacteria, plant and human SLC26 transporters.
1.4 The Sulfate Transporter and Anti-Sigma factor antagonist (STAS) domain

1.4.1 Bacterial sigma factors, anti-sigma factors and anti-sigma factor antagonists

The C-terminal cytosolic domain of SulP/SLC26 transporters shares a distant sequence homology with bacterial anti-σ factor antagonist, proteins involved in transcription regulation. Gene expression in bacteria is primarily regulated at the level of transcription initiation. While the evolutionarily conserved ~400 kDa catalytic core of the bacterial RNA polymerase (subunit composition αββ’ω) is capable of synthesizing RNA on its own, it is insufficient to initiate transcription. In order to begin promoter specific transcription, RNA polymerase reversibly associates with an additional subunit, sigma (σ), to form a holoenzyme. This sigma factor directs the recognition of promoter sequences located 10 and 35 base pairs upstream of the transcription initiation site. All bacteria have at least one essential σ factor and one or more alternative σ factors to transcribe specific classes of genes. In *E. coli*, for example, σ70 is the house keeping sigma factor that dictates the transcription of genes important for cell viability. Sigma factors themselves are regulated by anti-sigma factors which in turn can be inhibited by anti-anti-σ factors (anti-σ antagonists) which disinhibit σ factors by promoting their release from the anti-sigma factors.

The link between the SLC26 transporters and bacterial anti-σ antagonists became apparent upon the discovery that SpoIIAA, an anti-σ antagonist in *Bacillus subtilis*, has a weak sequence homology with C-terminal ends of SLC26 transporters. In *B. subtilis*, sporulation involves asymmetric division of the cell into two compartments called the forespore and the mother cell. Gene transcription in the forespore is achieved by controlled activation of the sporulation specific sigma factor, σF. Prior to cell division, the serine kinase and anti-σ, SpoIIAB, forms a
complex with $\sigma^F$ inhibiting the sigma factor from binding to the core RNA-polymerase$^{53}$. SpoIIAA counteracts this inhibition by interacting with SpoIAB and inducing the dissociation of $\sigma^F$ from the inhibitory complex$^{54, 55}$. SpoIIAA itself can be inhibited by the phosphorylation of its conserved serine (Ser58) residue$^{54}$.

It is reasonable to enquire why membrane transport proteins share homology with anti-sigma factor antagonists. The comparison of the STAS domain structure for E.coli against other structures deposited on the PDB using DALI server$^{56}$ hypothesized that the STAS domains SLC26 transporters also serve as sites of protein-protein interaction. Considering that mutations in the STAS domains of human SLC26 transporters have been associated with several pathologies, some studies have aimed at shedding light into the roles of these domains.

1.4.2 Functions of the STAS domains of different SLC26 transporters

The exact function of the STAS domain associated with most members of the SLC26 family has not been elucidated. However, studies in mammalian, plant and bacterial members of the family suggest different functions including protein folding and trafficking$^{23, 57}$, regulation of transport$^{58, 59}$ and nucleotide binding$^{60}$ (Table 2).

1.4.2.1 The role of STAS in protein folding and trafficking

Glycosylation and immunofluorescence experiments in the human Cl$^-$/HCO$_3^-$ exchanger, SLC26A3$^{23}$ and random mutagenesis studies in sultr1;2, the major sulfate transporter in the roots of Arabidopsis thaliana$^{57, 61}$ indicate that STAS domains are important for proper trafficking of the transporters.
Four chloride-losing diarrhea (CLD) mutations (ΔY526/7, I544N, I675/6Ins and G702T) are located within the STAS domain of SLC26A3. These mutant SLC26A3 transporters show reduced or no Cl⁻/HCO₃⁻ exchange activity. In order to understand the mechanism by which STAS domain mutations lead to CLD, the glycosylation state and cellular location of the wild-type and mutant transporters were monitored. Both these experiments indicated that STAS domain mutations result in ER retention and loss of functional protein at the plasma membrane. Deletion of the entire STAS domain in human SLC26A3 did not impair its trafficking to the plasma membrane in transfected HEK cells (Reithmeier et al., unpublished observations), showing that it is not essential for ER exit of this membrane protein.

Shibagaki and Grossman examined the activity of wild-type Sultr1;2 and modified versions of the transporter in yeast cells null for sulfate transport activity. In this heterologous system, truncated Sultr1;2 devoid of the STAS domain was unable to rescue the mutant phenotype. Different classes of STAS domain mutations in Sultr1;2 were later identified based on their effect on transport activity and localization to plasma membrane. Mutations of some residues within the STAS domain of Sult1;2 were shown to both diminish Sultr1;2 accumulation at the plasma membrane (<20% compared to wild-type) and decrease transport activity. On the other hand, other mutations within the STAS domain as well as the linker region between the N-terminal transmembrane domain and STAS domain had no effect on protein accumulation but significantly decreased transport activity suggesting that in addition to proper transporter localization, the STAS domain of Sultr1;2 might also have a transporter regulatory role.
1.4.2.2 Regulatory role of the STAS domain

A yeast 2-hybrid screen has shown that the STAS domain of Sultr1;2 physically interacts with an enzyme involved in cysteine synthesis, O-acetyl serine(thiol) lyase (OASTL)\(^59\). Once SO\(_4^{2-}\) is internalized into plant root cells, it is reduced to sulfide (S\(^2-\)). O-acetyl serine (thiol) lyase or cysteine synthase then catalyzes the conjugation of S\(^2-\) to O-acetyl serine (OAS) to form cysteine. Co-expression of OASTL and Sultr1;2 in yeast cells decreased the SO\(_4^{2-}\) transport activity of Sultr1;2\(^59\) while the activity of purified OASTL increased when co-incubated with the STAS domain \textit{in vitro}. These results implied that the STAS domain plays a regulatory role linking sulfate transport and metabolism. Accordingly, when the level of SO\(_4^{2-}\) is high, OASTL associates with the STAS domain of Sultr1;2 to inhibit its transport activity.

Interestingly, accumulating evidence suggests that the STAS domains of SLC26A3, SLC26A6, SLC26A8 and SLC26A9 transporters interact with the R-region of Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) to regulate Cl\(^-\) absorption and HCO\(_3^-\) secretion in secretory epithelia\(^62\)\(^-\)\(^64\). A single channel analysis demonstrated a reciprocal activation between SLC26A3 or SLC26A6 and CFTR as a result of the interaction between the STAS and R domains of SLC26A3 and CFTR, respectively\(^63\). Similarly, the Cl\(^-\)/OH transport activity of SLC26A3, SLC26A4 and SLC26A6 was enhanced in HEK293 cells co-expressing both the SLC26 transporters and CFTR\(^58\). Moreover, cAMP-activated Cl\(^-\) current via CFTR was stimulated by SLC26A3 and SLC26A6, and the isolated STAS domain was sufficient for this activation. The mechanism by which CFTR and SLC26 transporters mutually activate each other is not fully understood. However, the PDZ domain binding motifs found in the C-terminal parts of SLC26 transporters might be important since the deletion of these motifs results in
dissociation of the CFTR-SLC26 transporter complex and loss of mutual activation \(^{58}\). The current model proposes that CFTR and SLC26 transporters initially bind to a scaffolding protein via their PDZ domain binding motifs \(^{58}\). The binding recruits the CFTR and SLC26 transporters to form a Cl\(^-\)/HCO\(_3\)^- transporting complex. The interaction of the R domain of CFTR with the STAS domain of SLC26 transporters results in reciprocal activation of the two proteins, which in turn stimulates Cl\(^-\) absorption and HCO\(_3\)^- secretion. In contrast, the transport activity of SLC26A9, another human Cl\(^-\)/HCO\(_3\)^- exchanger, is inhibited by its interaction with the R-region of CFTR \(^{65}\).

1.4.2.3 The STAS domain as nucleotide binding protein

The first record of a STAS domain as a nucleotide-binding unit came upon the observation that the anti-sigma factor antagonist SpoIIAA of *Bacillus subtilis* bound GTP and ATP with dissociation constants of 0.25µM and 2.5µM, respectively \(^{66}\). Furthermore, SpoIIAA exhibits GTPase and ATPase activity \(^{66}\). Since then, the STAS domains of an SLC26 transporter in *Mycobacterium tuberculosis* and a blue-light responsive photoreceptor in *Bacillus subtilis* have also been shown to bind guanine nucleotides \(^{67},^{68}\). Rv1739c is one of the three SLC26 transporters in *Mycobacterium tuberculosis* and the only one that has been purified and characterized. Over expression of Rv1739c is associated with increased sulfate uptake in *E.coli* \(^{69}\). Since sulfolipids present in the outer membrane of *M. tuberculosis* have been correlated with pathogenesis, Rv1739c mediated sulfate transport may be important for virulence. Intrinsic fluorescence quench data and \(^1\)H-\(^{15}\)N chemical shift perturbation (CSP) studies demonstrated that monomeric STAS domain of Rv1739c binds GDP and GTP with dissociation constants of 146µM and 152µM, respectively \(^{70}\). Purified STAS (Rv1739c) also displays modest GTPase
activity. However, further studies are required to decipher the relationship between GTPase activity and sulfate transport for Rv1739c.

In *Bacillus subtilis*, the general stress transcriptional factor, $\sigma^B$, controls the expression of over 200 genes in response to environmental stress. $\sigma^B$ itself is regulated by cascades of signal transductions involving Ser/Thr kinases and phosphatases. The blue-light responsive photoreceptor, YtvA, is one of these proteins that are involved in regulation of $\sigma^B$. It is a 261 amino acid protein with an N-terminal flavin mononucleotide (FMN) binding LOV (Light, Oxygen and Voltage) domain and a C-terminal STAS domain linked by a helical amino acid sequence. In response to blue light, the YtvA positively regulates $\sigma^B$. Even though the role of the STAS domain in YtvA’s photosensitivity has not been elucidated, a fluorescence assay demonstrated that the STAS domain of this photoreceptor binds the fluorescent analogue of GTP, BODIPY-GTP, with a dissociation constant of 38 $\mu$M. This observation implicated the STAS domain of YtvA as a transducer of the photosignal from the FMN binding LOV domain. It was proposed that light received by the LOV domain induces a conformational change in the linker region and the NTP binding STAS domain. However, in a recent study, isothermal titration calorimetry (ITC), fluorescence and NMR spectroscopy experiments showed YtvA bound to BODIPY-GTP and not unlabeled GTP suggesting a non-specific interaction between the fluorescent dye and STAS(YtvA). Hence, the role of the STAS domain of YtvA in photosignal transduction remains unclear.
Table 2 The roles of functionally characterized STAS domains of SLC26 transporters

<table>
<thead>
<tr>
<th>Transporter, Organism</th>
<th>Characterized function of the STAS domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sultr1;2, <em>A. thaliana</em></td>
<td>Protein folding and trafficking(^{59}) Interacts with cysteine synthases of the sulfur metabolism pathway</td>
</tr>
<tr>
<td>SLC26A3, <em>H. sapiens</em></td>
<td>Proper folding and trafficking of the transporter</td>
</tr>
<tr>
<td>YchM, <em>E. coli</em></td>
<td>Binds Acyl Carrier Protein (ACP)(^{77})</td>
</tr>
<tr>
<td>Rv1739c, <em>M. tuberculosis</em></td>
<td>Binds GTP(^{70,78})</td>
</tr>
<tr>
<td>SLC26A3, SLC26A4, SLC26A6 and SLC26A8</td>
<td>Interact with the R-region of Cystic fibrosis transmembrane conductance regulator (CFTR) to regulate Cl(^-) absorption and HCO(_3^-) secretion in secretory epithelia(^{62-64})</td>
</tr>
</tbody>
</table>

1.4.4 Structural comparisons of mammalian and bacterial STAS domains

The structure of SpoIIAA features interspersed \(\beta\)-strands and \(\alpha\)-helices and has been used as a model for other SLC26 STAS domain structures. SpoIIAA contains five \(\beta\) strands and four \(\alpha\) helices arranged in the order: \(\beta1\)- \(\beta2\)- \(\alpha1\)- \(\beta3\)- \(\alpha2\)- \(\beta4\)- \(\alpha3\)- \(\beta5\)- \(\alpha4\). To date, there are 5 bacterial STAS domain structures from *M. tuberculosis*, *V. cholera*, *R. sphearoides*, *W. succinogenes* and *E. coli*, and one mammalian STAS domain structure of rat SLC26a5 deposited in the Protein Data Bank (PDB). As illustrated in Table 3, despite the low primary sequence identity among these proteins, they share similar tertiary folds. For example, the STAS domains of *V. cholera* and *E.coli* have only 15% sequence identity, but structural alignments of their STAS domains
show an RMSD of 2.1 Å. Structural alignments of known STAS structures indicate characteristic β-pleated sheet made of at least three parallel strands flanked by three α helices (Figure 1).

<table>
<thead>
<tr>
<th>Sequence Identity (%)</th>
<th>RMSD (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STAS (E.coli)</td>
</tr>
<tr>
<td><strong>E. coli</strong></td>
<td>100</td>
</tr>
<tr>
<td><strong>R. sphaeroides</strong></td>
<td>20</td>
</tr>
<tr>
<td><strong>M. tuberculosis</strong></td>
<td>17</td>
</tr>
<tr>
<td><strong>W. succinogenes</strong></td>
<td>21</td>
</tr>
<tr>
<td><strong>Prestin (SLC26A5)</strong></td>
<td>15</td>
</tr>
<tr>
<td><strong>V. cholera</strong></td>
<td>15</td>
</tr>
</tbody>
</table>

Table 3 Sequence and structural similarities of STAS domains with known structures. The sequences for STAS domains of *E.coli* were aligned against those of *V. cholera, M. tuberculosis, R. sphaeroides, W. succinogenes* and Prestin (SLC26A5). RMSD calculations were estimated based on secondary structure matching generated by COOT.

There are some divergences that can be observed at the N-terminal boundaries of the STAS domains. Compared to other STAS domains, *E.coli* YchM and rat Prestin STAS domains contain an additional anti-parallel β strand (β₀) at their N-termini. In contrast, the STAS domain of *V. cholera* has an N-terminal β turn. *V. cholera* and *W. succinogenes* STAS domains were crystalized as dimers (Figure 1c), while the *E.coli* STAS co-crystalized with a binding partner, acyl-carrier protein (ACP), which is discussed in detail below. Further analysis of the STAS domain structures shows no overlap among the dimerization interfaces in *V. cholera* and *W. succinogenes* STAS domains and protein-protein interaction site of *E. coli* STAS. It is possible for these STAS domains to interact with potential binding partners as dimers or monomers. These observations emphasize the view that STAS domains act as hubs of protein-protein interactions.
Figure 1 Structural alignments of bacterial and mammalian STAS domains. The STAS domains of SLC26 transporters from *E. coli* (dark blue, PDB ID: 3NY7), *V. cholera* (purple PDB ID: 4DGH), *M. tuberculosis* (green, PDB ID: 2KLN) *R. sphaeroides* (orange, PDB ID: 3OIZ), *W. succinogenes* (violet, PDB ID: 4DGH) and Prestin (red, PDB ID: 3LLO) were overlayed using SSM in COOT. (a) The structural alignments of STAS domains are shown as monomers (without their binding partners). *E. coli* STAS domain is represented in cartoon. (b) The overlaid STAS structures together with the binding partner of *E. coli* STAS, Acyl carrier protein, is shown. (c) The homodimeric STAS domains of *V. cholera* and *W. succinogenes* illustrate that the dimerization interfaces do not overlap.
1.5 YchM: E. coli’s sole member of the SLC26 family

In order to gain structural and functional insight into the SLC26 family, a prokaryotic homologue, YchM from E. coli, was targeted for structural and functional characterization. Unlike humans and plants that possess multiple SLC26/SulP transporters, E. coli possesses a single member of the SLC26 family, known as YchM. It is a 550 amino acid, inner membrane protein, with an N-terminal transmembrane domain (residues 1-435) and a C-terminal STAS domain (residues 436-550). Topology prediction programs suggest YchM spans the membrane 10-13 times (Table 4). These prediction results vary in both the orientation of the N and C-termini as well as the number of transmembrane segments. A topology map based on a sequence alignment with BicA predicts 12 transmembrane alpha-helices with cytosolic N and C-termini (Figure 2), the most likely topology for this membrane protein.

Little is known about the function of YchM or the role of its STAS domain. A whole cell [\textsuperscript{14}C] bicarbonate incorporation assay on wild-type and ychM\Delta E. coli cells showed a modest bicarbonate uptake increase in cells expressing YchM \textsuperscript{77}. Moreover, bicarbonate uptake was enhanced in the presence of NaCl suggesting that YchM might be a Na\textsuperscript{+} -dependent bicarbonate transporter. While YchM mediated bicarbonate transport might not be essential at neutral pH, earlier experiments have shown growth defects in ychm\Delta HfrC E. coli cells that are grown at pH 8.3 \textsuperscript{77}. As show in the reaction below, under alkaline pH, the reaction favours the formation of HCO\textsubscript{3}\textsuperscript{-}, a negatively-charged substrate that requires a transporter to move across a membrane.

\[
\text{H}_2\text{CO}_3 \leftrightarrow \text{HCO}_3^- + \text{H}^+ \quad \text{pKa = 6.35}
\]

On the contrary, during acidic pH, the reverse reaction produces carbonic acid, which is quickly converted to CO\textsubscript{2} and H\textsubscript{2}O. CO\textsubscript{2} can diffuse across membranes without the aid of a
transporter. Further study is required to better understand the role of YchM in mediating bicarbonate transport at different pH’s.

Table 4 Topology of *E. coli* YchM using different prediction programs

<table>
<thead>
<tr>
<th>Prediction Program</th>
<th>Number of TM helices predicted</th>
<th>Location of N-terminus</th>
<th>Location of carboxyl STAS domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCAMPI-seq</td>
<td>11</td>
<td>Cytoplasmic</td>
<td>Periplasmic</td>
</tr>
<tr>
<td>SCAMPI-msa</td>
<td>12</td>
<td>Cytoplasmic</td>
<td>Cytoplasmic</td>
</tr>
<tr>
<td>PRODIV</td>
<td>13</td>
<td>Cytoplasmic</td>
<td>Periplasmic</td>
</tr>
<tr>
<td>PRO</td>
<td>11</td>
<td>Cytoplasmic</td>
<td>Periplasmic</td>
</tr>
<tr>
<td>OCTOPUS</td>
<td>12</td>
<td>Cytoplasmic</td>
<td>Cytoplasmic</td>
</tr>
<tr>
<td>TOPCONS</td>
<td>13</td>
<td>Cytoplasmic</td>
<td>Periplasmic</td>
</tr>
<tr>
<td>TMHMM</td>
<td>11</td>
<td>Periplasmic</td>
<td>Cytoplasmic</td>
</tr>
</tbody>
</table>

![Figure 2 Predicted topology of *E. coli*’s SLC26 transporter, YchM. YchM transmembrane helices were modeled based on the topology of Bic A which spans the membrane 12 times with cytosolic N- and C-terminal. The topology map was generated using TrasMembrane Protein Re-presentation in 2 Dimensions (TMRPres2D) software tool.](image-url)
1.5.1 The STAS (YchM)-ACP interaction: A potential link between bicarbonate transport and fatty acid synthesis

1.5.1.1 The STAS domain interacts with acyl carrier protein (ACP)

In an attempt to structurally characterize the STAS domain of YchM, STAS expressed in *E. coli* co-purified and co-crystallized with endogenous acyl carrier protein (ACP) \(^{77}\). ACP is a small \((Mw \sim 9kDa)\) and highly conserved, soluble protein that acts as a cofactor in all stages of fatty acid biosynthesis pathway \(^{80}\), lipid A biosynthesis \(^{81}\), and membrane phospholipid synthesis \(^{82}\). It is initially synthesized as an apoprotein and in a post-translational modification, ACP synthase (AcpS) adds a 4'-phosphopantetheine (4'-PPa) group onto a conserved serine residue (Ser36) of apo-ACP to form holo-ACP \(^{83}\). *In vivo*, almost all of the ACP is maintained as holo-ACP \(^{84}\). During fatty acid biosynthesis, the terminal sulfhydryl (SH) group of the 4'-PPa prosthetic group allows holo-ACP to shuttle acyl chains of various lengths via a thiol ester linkage.

High-resolution NMR and crystal structures of apo, holo and different acylated derivatives of ACP have been determined \(^{85-88}\). The ACP structure consists of a four \(\alpha\)-helical bundle. The second helix of ACP possesses highly conserved negatively charged amino acids that are important for ACP’s interaction with different enzymes \(^{89}\). Even though the over all structures of apo- and holo- ACP are similar \(^{85}\), solution structures of *M. tuberculosis* holo-ACP illustrates the protein exists in conformational equilibrium in which the 4'-PPa group of ACP oscillates between two states \(^{86}\). In the first state, the prosthetic group is bound within a hydrophobic cavity created by helices 2 and 3 of ACP. In its second, solvent-exposed state, the prosthetic group protrudes away from the ACP moiety making itself available for protein/enzyme interactions. Once acyl chains are conjugated to the prosthetic groups of ACP,
the hydrocarbons are tucked within the hydrophobic cavity\textsuperscript{88}. As the acyl-chain gets longer, the cavity expands to accommodate the growing hydrocarbon chain\textsuperscript{90}.

As illustrated in Figure 3a, the crystal structure of the ACP-STAS complex depicts ACP with 4'- phosphopantethiene conjugated to a malonyl group. There are no lone structures of either the E.coli STAS domain or malonyl-ACP. The STAS domain of YchM within the ACP-STAS complex shows five beta-strands and four $\alpha$- helices with the pattern: $\beta 1$-$\beta 2$- $\alpha 1$- $\beta 3$- $\alpha 2$- $\beta 4$- $\alpha 3$- $\beta 5$- $\alpha 4$ (Figure 3). A five stranded $\beta$ sheet, with one anti-parallel strand ($\beta 1$) and four parallel strands make up the core of the STAS structure. The four parallel strands are flanked by two parallel helices. Several residues on the $\alpha 3$ helix of the STAS domain interact with residues on the $\alpha 2$ and $\alpha 3$ helices of ACP. Arg 523 on the $\alpha 3$ helix of the STAS domain makes hydrogen bonds with carboxyl groups of Glu 48 and Asp 57, and the carbonyl backbone of Ile 55 of ACP. In addition, Arg527 of STAS makes a Van der Waals contact with Met45 of ACP. Furthermore, the malonyl and phosphopanthethiene groups of ACP protrude away from the ACP structure and sit in a hydrophobic groove created between the $\alpha 3$ and $\beta 5$ of the STAS domain of YchM. This hydrophobic groove is mainly formed by the STAS domain residues Phe 539, Phe 540, Pro 532, Pro 541 and Ile 530, which make direct contact with the 4’-PPa group\textsuperscript{77}. Gln531 of the STAS domain forms hydrogen bond with the malonyl group. Acylated derivatives of ACP with their hydrocarbons hidden in pockets formed by ACP helices would not be expected to interact the same way. Structural alignments of holo-ACP or decanoyl-ACP with the STAS-malonyl-ACP (Figure 4) show the decanoyl- acyl chain tucked away from the ACP-STAS interaction interface.
Figure 3 The crystal structure of the YchM-STAS and ACP complex (a) The structure of the ACP-STAS complex illustrates the phosphopantetheine group of ACP (light blue) interacts with The STAS domain of YchM (aa 436-550) (deep blue). (b) Structural analysis of the ACP-STAS interaction interface shows Arg523 and Arg 527 of the STAS domain interacting with Asp 57 and Glu 48 of ACP.
Figure 4 Structural alignment of holo-ACP and dodecanoyl-ACP with the STAS(YchM)-ACP complex. Holo-ACP (light blue, PDB accession code: 2F2Q) and decanoyl-ACP (orange, PDB accession code: 2FAE) are overlaid onto Malonyl-ACP (light green) of the STAS(YchM)-ACP complex structure (PDB accession code: 3NY7). The malonyl group of malonyl-ACP is shown interacting with the STAS domain of YchM while the 4’PPa and acyl chain of holo-ACP and decanoyl-ACP, respectively, are tucked in between the helices of ACP.
ACP is a highly interactive protein that binds a number of proteins involved in various biochemical pathways \(^91\). Therefore, it is important to show biological relevance of the ACP-STAS interaction. In addition, ACP has been shown to bind proteins with unrelated functions to fatty acid or lipid synthesis. For example, ACP interacts with SpoT \(^92,\,^93\), an enzyme that catalyzes the synthesis and hydrolysis of the transcriptional regulator ppGpp \(^94\) and MukB \(^95\), a protein required for correct chromosome partitioning in *E. coli*. In order to evaluate the biological relevance of ACP-YchM interaction and its link with fatty acid synthesis, the functional relationship between the two proteins was independently analyzed using a high-throughput *E.coli* synthetic genetic array (eSGA) screen. This technique allows identification of genes that operate on the same or parallel pathway by comparing the growth phenotypes of cells harboring single gene deletion with those carrying double gene deletions. Growth defects in double mutants compared to the single mutants imply the two deleted genes work in parallel pathway where knockout of either of the gene is not lethal but knockout of both is. eSGA analysis of YchM, in which *YchMΔ* cells were conjugated with cells harboring non-essential and hypomorphic alleles of essential *E.coli* genes identified interactions with genes encoding components of the fatty acid biosynthesis pathway – FabA and FabR \(^77\).

The crystal structure of the ACP-STAS complex together with the genetic interactions reported by the eSGA analysis indicate YchM’s potential role in fatty acid biosynthesis.

**1.5.1.2 YchM-ACP interaction and Fatty acid biosynthesis**

*De novo* fatty acid biosynthesis is an essential process that supplies substrates for membrane biogenesis, bacterial quorum sensing and post-translational protein modification \(^80\). Bacteria and plants carry out fatty acid synthesis using a set of conserved individual proteins
expressed from separate genes (called Type II)\textsuperscript{97}. A crucial player in Type II fatty acid synthesis is the acyl carrier protein (ACP), essential for all stages of fatty acid biosynthesis (FAB) including initiation, elongation and transfer to a membrane bilayer \textsuperscript{80}.

In the first committed stage of fatty acid biosynthesis, acetyl-CoA is converted to malonyl-CoA by two partial reactions catalyzed by acetyl-CoA carboxylase (AccABCD) \textsuperscript{98-100} (Figure 5, step 1). In the first reaction, bicarbonate is used to carboxylate biotin in an ATP-dependent reaction to form carboxybiotin. The carboxyl group of carboxybiotin is transferred to acetyl-CoA to create malonyl-CoA. Then, malonyl-CoA:ACP transacylase (FabD) transfers the malonyl group to ACP forming malonyl-ACP: an activated, acetyl group donor in the condensation step of fatty acid synthesis \textsuperscript{101} (Figure 5, step 2). It is this intermediate that binds the STAS domain of YchM. The condensation of malonyl-ACP with acetyl-CoA by \(\beta\)-ketoacyl-ACP synthase III (FabH) initiates fatty acid elongation. In the next two steps of elongation, FabG reduces \(\beta\)-ketoacyl-ACP to \(\beta\)-hydroxyacyl-ACP which in turn is dehydrated by FabA or FabZ to form trans-2-enoylACP. This product is then reduced in a NADH-dependent reaction by enoylreductase (FabI) \textsuperscript{102}. The acyl-ACP formed in this last stage undergoes repeated condensation reaction with malonyl-ACP (produced by AccABCD above), growing its acyl chain 2 carbons per cycle. These condensation reactions are catalyzed \(\beta\)-oxoacyl synthases I or II (FabB or FabF) and repeated until the acyl chain is 16-18 carbons long.

The crystal structure of the ACP-STAS complex showed that YchM interacts with malonyl-ACP in which both the acyl and 4’-phosphopanthetiene group interact with the STAS domain residues. The interaction of ACP with the STAS domain suggests that YchM might play a role in fatty acid biosynthesis. The requirement of bicarbonate in the first step of FAB coupled with the interaction of the STAS domain of YchM with malonyl-ACP implies a potential link.
between bicarbonate transport and fatty acid biosynthesis. Since, fatty acid synthesis is an energetically costly process, the rate of the fatty acid production needs to be tightly regulated. Indeed, acylated derivatives of ACP can govern fatty acid synthesis by product feedback inhibition of the catalytic enzymes AccABCD and FabH\textsuperscript{103, 104}. Fatty acid synthesis is also controlled at the level of transcription where, for instance, malonyl-ACP can bind to and disinhibit the transcriptional repressor FabR\textsuperscript{105}.

As the first step of fatty acid synthesis consumes ATP, the carboxylation of acetyl-CoA to malonyl-CoA is one of the main sites of regulation. A possible mode of fatty acid regulation is via the inhibition of YchM, a potential bicarbonate supplier for the carboxylation of acetyl-CoA to malonyl-CoA. Accordingly, YchM could be regulated via STAS mediated interaction with malonyl-ACP depending on the level of malonyl-CoA inside the cell, a high level indicating an ample supply of bicarbonate. There have not been any studies that show bicarbonate-mediated regulation of the reaction. Furthermore, functional associations between transporters and metabolic enzymes that act on the substrates of the transporters have previously been reported for different systems including \textit{E.coli}\textsuperscript{106}. In fact, this association is reminiscent of the aforementioned regulatory association between the plant SLC26 member, Sultr1;2, and cysteine synthase\textsuperscript{59}. 

**Figure 5** Bacterial fatty acid biosynthesis initiation and condensation. Fatty acid biosynthesis begins with the carboxylation of acetyl-CoA to malonyl-CoA (step 1). Then, malonyl-CoA:ACP trasacylase (FabD) transfers the acyl group from malonly-CoA to ACP to form malonyl-ACP (step 2). The condensation of malonly-ACP with acetyl-CoA by β-ketoacyl-ACP synthase III (FabH) initiates fatty acid elongation (Step 3). Reproduced with permission from Chan D.I and Vogel H. J. (2010) *Biochem J.* 430 1-19 © the Biochemical Society

In addition to fatty acid biosynthesis, YchM has also been linked with fatty acid degradation (FAD) through its interaction with FAD components, FadE and FadB. Genetic analysis and affinity pull-downs have demonstrated that YchM interacts with FadE, the enzyme involved in the first step of β-oxidation. Furthermore, the deletion of *ychm* together with genes encoding components of FAD were found to be lethal in *E. coli*, implying that these metabolic
enzymes work in parallel pathways with YchM. This suggests that YchM might be involved in fatty acid degradation.

Bacteria can utilize fatty acids of different acyl chains as their sole carbon and energy sources\(^\text{107}\). Degradation of these fatty acids begins with the transport of fatty acids into the cell by an outer membrane protein, FadL\(^\text{108}\). Then, an inner membrane associated protein, acyl-CoA synthase (FadD) transfers the acyl chain to coenzyme A to form acyl-CoA\(^\text{109}\). Once inside the cytosol, \(\beta\)-oxidation of fatty acids starts when acyl-coA dehydrogenase (FadE)\(^\text{110}\) converts acyl-CoA to enoyl-CoA, which in turn is converted to 3-ketoacyl-CoA by FadB (enoyl-CoA hydratase). In the last step, 3-ketoacyl-CoA is cleaved by FadA shortening the acyl chain by two carbons. This step releases acetyl-CoA that enters the Kreb’s (TCA) cycle.

YchM is not essential for growth in nutrient rich media and deletion of YchM is lethal when one other component of the FAD pathway is also absent. This observation indicates that YchM might be crucial when cells depend on FAD pathway for metabolism of fatty acids as their carbon source.

1.6 Hypothesis and Project Objective

The aim of this study is to elucidate the role of the STAS domain in YchM’s transport activity with particular focus on the ACP-YchM interaction. The interaction of YchM with different components of the fatty acid metabolism pathway suggests YchM plays a role in fatty biosynthesis and degradation, and the STAS domain may mediate protein interactions within FAB/FAD pathways through ACP. My hypothesis is that the STAS domain of YchM is
necessary for linking the bicarbonate transport activity of YchM to fatty acid metabolism via its interaction with ACP.

In order to test the proposed hypothesis, two independent but complementary approaches were carried out as follows:

(i) \textit{In vitro identification of residues critical for the ACP-STAS interaction}

In the first part of the study, an \textit{in vitro} system was designed in order to identify key residues that are crucial for the interaction of ACP and the STAS domain of YchM. Based on the crystal structure of the ACP-STAS complex, site-directed mutagenesis experiments were carried out against residues within the binding interface. The interactions of wild-type and mutant proteins were analyzed using different biochemical and biophysical assays to obtain binding constants and evaluate binding affinities. A mutation in the STAS domain that prevents binding to ACP may be useful in \textit{in vivo} studies of the functional consequences of disrupting the YchM-ACP interaction.

(ii) \textit{Development of an in vivo functional assay for YchM}

In order to characterize the role of the STAS domain or other portions of YchM, it is essential to develop a simple assay that qualitatively or quantitatively measures the function of YchM. The ability of YchM to provide a selective advantage for growth in alkaline pH or in media with fatty acids as sole carbon source was assessed.
Chapter 2 Methods and Materials
2.1 Materials

**Strains:** ychmΔ Hfr C and BW25113 cells were obtained from Dr. Andrew Emili’s (University of Toronto) and Dr. Alan Davidson’s (University of Toronto) laboratories. DY330 and MG1655 *E. coli* cells were acquired from Dr. Walid Houry’s laboratory (University of Toronto).

**Chemicals and Media:** All chemicals and Luria-Broth (LB) media components tryptone and yeast extract were purchased from Bioshop®. Oleic acid for growth assay was purchased from Sigma.

**PCR and cloning components:** Primers were purchased from Sigma. Pfu (Thermo Scientific®) and Phusion (New England BioLabs® Inc.) DNA polymerases were used for site-directed mutagenesis and Polymerase Chain Reaction PCR, respectively. Fermentas® plasmid purification kit and Sigma® DNA gel extraction kit were used. PCR was done within a BioRad® thermo-cycler.

**Instruments for Biophysical assays:** Surface Plasmon Resonance (SPR) was performed using a Biacore® X100. Bio-layer interferometry was performed on an Octet Red96 provided by ForteBio. Tryptophan fluorescence measurements were performed using a Fluorolog® 3 spectrofluorometer and a TECAN Infinite® 200 plate reader. A Jasco J-810 circular dichroism spectrometer was used for secondary structural analysis. Isothermal titration calorimetry (ITC) measurements were executed using MicroCal-VP-ITC. For growth assays, cells were grown in Bioscreen C® plate reader.

**Antibodies:** Anti-ACP antibodies were generously donated by Dr. Charles Rock’s laboratory (St. Jude’s Children’s Hospital, Memphis TN, USA). Mouse anti-His antibodies were purchased
from Cell Signaling®. Rabbit anti-mouse and goat anti-rabbit secondary antibodies were purchased from Novagen®.

Samples were prepared using 18 MΩ-cm resistance water. HCl or NaOH were used for pH adjustments unless noted.

### 2.2 In vitro analysis of the ACP-STAS interaction

#### 2.2.1 Plasmids and Cloning

Based on a sequence alignment of YchM and SpoIIAA (Figure 6), the STAS domain of YchM was previously defined to consist of amino acid residues 436-550 of YchM. pET-15b and pET-28a vectors containing N-terminal thrombin cleavable His-tagged ACP and STAS436 genes under the T7 RNA polymerase promoter were cloned as previously described.

![Figure 6 Sequence alignment of B. subtilis SpoIIAA and E. coli YchM. Based on the alignment, the STAS domain of YchM encompasses the last 115 amino acids of YchM.](image)

In order to stabilize the STAS domain, a second construct with N-terminal GST-tag was designed. Since the STAS domain is found at the C-terminus of YchM with the transmembrane domain of the protein at its N-terminus, an N-terminal GST tag should not occlude potential
protein-protein interaction sites. Forward primer (5’ACGGGGATCCATGACTCGCCT GGCA CCG,BamH1 site underlined) and reverse primer (5’CCGGGTGACTTTTATA AATCCGCCATCGCC, Sal1 site underlined) were used to PCR amplify the STAS_{436} encoding gene from the pET-28a construct. The PCR product was gel purified, digested with Sal1 and BamH1, and ligated to pGEX-6P-1 vector digested using the same restriction enzymes. The ligation product was transformed into DH5α cells, and the DNA was purified using Fermentas® Mini-prep kit. The constructs were sequenced by ACGT Corp (Toronto, Canada) using pGEX-6P-1 commercial forward and reverse primers pGEX5 and pGEX3 respectively. A third construct was designed with an N-terminal truncation to remove the first 11 hydrophobic amino acids (residues 436-446). The STAS_{447} construct was cloned using site-directed mutagenesis of STAS_{436} (pGEX-6P-1) construct using primers pGEX 477STAS F and pGEX 477STAS R (Appendices). Primers used for site-directed mutagenesis of STAS and ACP are presented in the Appendices.

2.2.2 Protein expression and purification

pET-15b and pET-28a vectors containing N-terminal thrombin cleavable His-tagged ACP and STAS genes (WT and mutants), respectively, were expressed in BL21(DE3) E. coli cells. Protein expression was induced according to the Studier auto-induction protocol\footnote{Studier, 2005}. Single colonies were used to inoculate 1-2 ml of LB media in the presence of ampicillin. The starter culture was grown for 6-7 hrs at 37°C and then transferred to 2.5 L of auto-induction media. Cells were grown overnight at 37 °C, harvested by centrifuging at 4,000 rpm for 20 min and frozen at – 80 °C

pET15b vector containing N-terminal hexahistidine-tagged ACP synthase was expressed in E. coli BL21(DE3) strain. The cells were inoculated overnight on ampicillin agar plate at 37
°C. A single colony of cells was transferred to 2 L LB media and grown for 3-4 hrs until OD$_{600}$~0.6 when it was induced with 0.5 mM IPTG. An overnight culture was harvested by centrifugation at 4,000 rpm for 20 minutes.

All His-tagged proteins (ACP and ACP synthase) were purified as follows: First, cells were resuspended in lysis buffer containing 50 mM NaH$_2$PO$_4$, pH 8, 300 mM NaCl and 10 mM imidazole. EDTA-free protease inhibitor cocktail tablet was added to the sample and the cells were sonicated for 2.5 minutes using a probe-type sonicator. Cells were pulsed for 30 seconds on and off at 4 °C. The whole cell extract was centrifuged at 16,000 rpm for 45 minutes. 1 mL of Ni-NTA resin (Qiagen) was equilibrated with the lysis buffer and added to the clarified lysate. The lysate-resin mixture was incubated for 1 hr at 4 °C and loaded onto a column. The resins were washed with lysis buffer followed by the same buffer containing 50 mM imidazole. Proteins were eluted with 300 mM imidazole. Peak fractions based on A$_{280}$ and SDS-PAGE analysis were collected and concentrated. For ACP Synthase, the imidazole-containing buffer was then exchanged for 10 mM HEPES, 50 mM NaCl and 15% glycerol buffer for stabilization and stored at -80°C.

N-terminal GST-tagged wild-type and mutant STAS domains (GST-STAS, GST-STAS$_{R527E}$, GST-STAS$_{R523E}$) were also expressed overnight in BL21(DE3) cells using auto-induction media that were harvested and frozen according to the protocol mentioned above. The cell pellets were thawed and re-suspended in phosphate buffer saline (PBS), then lysed by 2 passes of French Press at 1500 psi or 2 minutes of sonication. The lysate was centrifuged at 18,000 rpm for 45 minutes. The clear supernatant was filtered with 0.45µm membrane filter (Millipore®) and loaded on 1mL a GSTrap column. The column was washed with PBS and proteins were eluted using elution buffer containing 50 mM Tris, pH 8 and 10 mM glutathione.
In order to cleave the GST, fusion proteins were incubated with PreScission protease overnight while dialyzing against 50 mM Tris, pH 8, 150 mM NaCl, 1 mM DTT and 1 mM EDTA. The cleavage product was incubated with glutathione resins (GE Amersham®) to remove GST and uncleaved protein and then loaded onto a Superdex 75 size exclusion column.

### 2.2.3 Malonyl-ACP Synthesis

50 µM of apo-ACP was incubated with 100 µM malonyl-CoA, 1 µM ACP-synthase, 5 mM DTT, 1 mM MgCl₂ in 50 mM Tris, pH 7.5 or pH 8 and incubated for 3 hrs at 37 °C. The product was loaded on a Superdex75 column to remove ACP-synthase and exchanged into ITC running buffer. The formation of holo-ACP or malonyl-ACP was confirmed by MALDI-TOF mass spectrometry.

### 2.2.4 SDS-PAGE, Native-PAGE and Western Blot

2 x SDS sample buffer (0.5 mM Tris, pH 6.8, 15% glycerol, bromophenol blue) was added to each sample in 1:1 ratio. Gel electrophoresis was performed at 180V for 55 minutes. For SDS-PAGE analysis, the gels were stained with Coomassie Blue stain. For Western blots, the proteins were transferred to a polyvinyl difluoride (PVDF) membrane in transfer buffer (25 mM Tris, 192 mM glycine, 10% methanol) at 300 mA for 75 minutes. Prior to transfer, the PVDF membrane was activated by soaking in methanol for 15 seconds. After transfer, membranes were incubated in blocking buffer (5% skim milk in Tris buffered saline with 1% Tween, TBST) overnight or for 1 hr. Then, they were incubated with 1:10,000 or 1:5,000 of rabbit anti-ACP or mouse anti-His antibody, respectively. Anti-ACP antibodies were incubated overnight at 4°C. The membranes were then washed with TBST three times for a total of 20 min and incubated in 1:5000 dilution of HRP conjugated anti-rabbit or anti-mouse antibody for 1 hr.
Membranes were washed three times with TBST. Enhanced chemiluminescence (ECL) reagent (PerkinElmer) was used to probe the antibody. Protein bands were visualized using a BioRad® phosphorImager.

To visualize apo-and holo-ACP, 20% Tris-glycine native gels were used. The gels were run for 75 min and stained with Coomassie for visualization.

### 2.2.5 Circular dichroism spectroscopy and thermal denaturation assay

Prior to the experiment, the protein samples were dialyzed into 10 mM KCl and 100 mM K₃PO₄. CD measurements were performed between 190-270 nm at 10°C. For thermal denaturation assay, CD measurements of 0.5 mg/ml of the STAS domain were analyzed at a fixed wavelength of 222 nm between 20-70°C.

### 2.2.6 Guanidine denaturation assay

A 6 M guanidine-HCl stock solution was prepared. The concentration of the STAS domain was kept fixed at 5 µM in 20 mM HEPES, pH 8, 100 mM NaCl and increasing concentrations of Guanidine-HCl were added to make a final guanidne concentration of 0.5, 1, 2, 3, 4, 5 and 6 M. The mixtures were incubated for 1 hr at room temperature and tryptophan fluorescence of the STAS domain was measured at an excitation wavelength of 295 nm. Emission spectrum was collected at 1 nm/min for wavelengths 300-400 nm. Each measurement was performed in duplicate and an average of the two fluorescence intensities was used for graphical analysis.
2.2.7 Surface Plasmon Resonance

The STAS domain was immobilized onto a CM5 sensor chip using N-hydroxysuccinimide (NHS) and N-ethyl-N(dimethylaminopropyl) carbodiimide (EDC) mediated amine coupling. The carboxymethyl groups of the sensor chip were first activated using NHS/EDC and 2.8 ng/mL of STAS domain was injected for crosslinking. The coupling reaction occurred at pH 5 and was quenched using ethanolamine. Apo- or holo-ACP was diluted at various concentrations in 20 mM HEPES, pH 8, 150 mM NaCl and 0.005% surfactant P20. The coupled protein/chip was equilibrated with the same buffer, 20 μL of each sample was injected at a flow rate of 10 μL/min and the bound ACP was washed using the running buffer.

2.2.8 Affinity Pull-down

pGEX-6P-1 vectors containing genes that code for GST-STAS_{436}, empty pGEX-6P-1 vector that expresses GST (negative control) and pET-15b vector that expresses N-terminal Histagged ACP were transformed into BL21(DE3) cells. Single colonies were used to inoculate 25 mL of overnight expression media in the presence of 50 ng/mL ampicillin. Cells were grown at 37°C for 24 hrs, harvested by centrifugation at 4,000 rpm and re-suspended in 2 mL of PBS buffer. EDTA-free protease inhibitor tablets were added to each sample and the cells were lysed by probe-type sonicator. The whole cell extracts were centrifuged at 15,000 rpm for 20 min. Then, ACP cell lysates were mixed with either GST or GST-STAS_{436} cell lysates and incubated with 15 μL of glutathione resin for 2 hrs. The resins were washed four times with PBS. After the last wash, the resins were re-suspended in SDS loading buffer, boiled and loaded on SDS-PAGE for analysis. Anti-His Western blots were performed to confirm co-purification of ACP with each sample.
2.2.9 Bio-layer interferometry

All binding reactions were carried out in phosphate buffer saline (PBS), pH 7.4. In order to immobilize biotinylated STAS domain, all eight streptavidin probes were equilibrated in PBS buffer and then immersed into 50 ng/ml of biotinylated STAS solution. After saturating the probes with biotinylated STAS, the excess protein was washed off using PBS. A serial 1:2 dilution of holo-ACP was performed to give final concentrations ranging from 40 μM to 27 nM. Binding experiments were carried out by immersing the STAS bound probes in wells containing different concentrations of holo-ACP. After binding, the probes were washed for 2 min in PBS.

2.2.10 Tryptophan fluorescence Assay

Prior to the assay, all protein samples were dialyzed against a buffer containing 20 mM HEPES, pH 8, 100 mM NaCl and 1 mM TCEP. 5 μM of STAS was incubated with 2.5, 5, 10, 20 and 40 μM of holo-ACP for 1-2 h. Fluorescence measurements were performed at 25 °C. The samples were excited at 295nm and emission spectrum was collected for wavelengths 300 to 400 nm. For data analysis, the fluorescence intensities of the buffer and corresponding concentration of holo-ACP were subtracted from STAS and STAS-ACP measurements, respectively.

2.2.11 Isothermal Titration Calorimetry (ITC)

The thermodynamic parameters of the ACP-STAS interaction were measured using MicroCal ITC. Prior to ITC measurements, all samples were dialyzed against a buffer containing 10 mM HEPES, pH 8, 150 mM NaCl and 1 mM TCEP to reduce holo-ACP dimers. The reference cell was filled with water and the sample cell was loaded with 1.4 mL of buffer (negative control), 40 μM of wild-type, apo-ACP, holo-ACP, or ACP<sub>D57R</sub>. Aliquots of 10μL of
175 µM wild-type or mutant STAS were titrated into the sample cell at 5 minutes intervals. A total of 20 injections were made per experiment keeping a constant stirring speed of 300 rpm. Titration curves were fitted using the one binding site model in the MircoCal Origin software.

2.3 Development of an *in vivo* functional assay pH growth assay

2.3.1 Generation of *ychm*Δ *E.coli* strain

Sixty nucleotide-long primers that contain complementary regions to the genes flanking YchM and the chloramphenicol resistance cassette were designed (Appendix 1). The forward primer (P1) has 40bp homologous sequence to the region that is 40 nucleotides upstream of the *ychm* and its 3’ end (20 nts) is complementary to the chloramphenicol resistance (CmR) gene of pkD3 plasmid. The chloramphenicol cassette from the pKD3 plasmid was amplified using primers P1 and P2. The PCR product was gel purified (Qiagen kit) and electroporated into DY330 cells for homologous recombination. DY330 cells containing temperature inducible \( \lambda \) red recombinase were induced by incubating cells at 42 °C. The cells were recovered after 1hr and plated on LB-agar plates containing chloramphenicol. Colony PCR was performed to confirm the replacement of *ychm* with the CmR gene.

The knockout was transferred to MG1655 cells using P1 phage transduction. A single colony of *ychm*Δ:CmR DY330 *E.coli* cells was used to inoculated 5 mL of LB media containing 50 ng/mL of chloramphenicol. The next day, 4 tubes of 150 µL of the cultures were centrifuged and re-suspended in a buffer containing 5 mM of CaCl\(_2\) and 10 mM MgSO\(_4\). 80 µL, 150 µL or 300 µL of P1 phage was added to three of the tubes to test different phage:cell ratios. The phage-cell mixture was then incubated at 30°C for 40 min at which point 230 µL of 1 M sodium citrate...
was added to stop the transduction. The cells were centrifuged, re-suspended in 1 mL of LB containing 1 M sodium citrate and incubated at 30 °C for 1 hour. The recovered cells were then plated on LB agar containing chloramphenicol and 5 mM citrate. Colony-PCR was performed to confirm the ychm gene was knocked out.

2.3.2 pH growth assay

Three E.coli strains were used to analyze the role of YchM for growth under alkaline conditions: Hfr Cavali (obtained from Dr. Andrew Emili’s Laboratory); MG1655 and BW25113 E. coli (obtained from Dr. Alan Davidson’s Laboratory) cells. LB media in different alkaline pH’s (7, 7.5, 8, 8.5 and 9) were prepared by buffering the LB media with 40 mM MOPS/Tricine with the respective pH values. Single colonies of wild-type and ychm∆ MG1655 E coli cells were used to inoculated 5 mL of LB media and grown overnight at 37 °C. The next day, aliquots of overnight cultures were adjusted to an OD$_{600}$ = 0.1 and used to inoculate 100 µL of LB media pH 7, pH 8, pH 8.5 and pH 9 buffered with 40 mM MOPS/Tricine. A serial 1:10 dilution was performed in 100-well titer plates. OD$_{600}$ measurements were made every 15 minutes over 24 hrs in BioscreenC® growth plate reader.

2.3.3 Fatty acid growth assay

In order to test the ability of ychm∆:kan$^R$ BW25113 E. coli cells to metabolize fatty acids, a fatty acid media was prepared as follows: Minimal media containing 0.02 or 0.2 % glucose (positive control) or 0.02 or 0.2% oleic acid were prepared. Glycerol stocks of WT and ychm∆:kan$^R$ BW25113 E. coli cells were streaked on LB agar plates. Single colonies of WT and ychm Δ BW25113 E coli cells were used to inoculated 5 mL of LB media and grown overnight at 37 °C. The ychm Δ BW25113 E. coli cells were grown in the presence of kanamycin. The next
day, 1 mL of cells were centrifuged and re-suspended in M9 salt solutions. The cells were washed to remove residual LB and adjusted to an OD$_{600} = 0.1$ and used to inoculate 100 µL of fatty acid-containing media. A serial 1:10 dilution was performed in a 100-well titer plate and OD$_{600}$ measurements were made every 15 minutes for 24 hrs.
Chapter 3 Results
3.1 *In vitro* identification of residues that are critical for the ACP STAS interaction

3.1.1 Protein purification

In order to biochemically characterize the ACP-STAS interaction, a soluble and stable STAS domain is essential. The STAS domain of YchM was previously defined to encompass the last 115 amino acids (residue 436-550, STAS\textsubscript{436}). A construct with an N-terminal thrombin cleavable his-tagged STAS\textsubscript{436} was expressed in BL21(DE3) *E. coli* cells. (His)\textsubscript{6}-STAS\textsubscript{436} was purified using Ni\textsuperscript{2+} affinity chromatography. Although the protein expressed well, it was unstable and significant protein precipitation was observed during His-tag cleavage.

Wild type STAS and STAS mutants were expressed and purified with an N-terminal GST tag to aid in protein stability and solubility. The size exclusion elution profile of GST-STAS\textsubscript{436} is shown in Figure 7a. After the affinity purification step, the GST was cleaved by incubating GST-STAS with PreScission protease overnight. After removal of GST with a second round of glutathione resin incubation, the sample was loaded onto a Superdex 75 gel filtration column. The STAS\textsubscript{436} domain eluted at a volume of 12.7 mL, which corresponds to the molecular mass of a monomer (12.9 kDa) (Figure 7b) and confirmed by SDS-PAGE, where the apparent molecular weight of the STAS domain was \( \sim 12\text{kDa} \). The peak at \( \sim 10\text{mL} \) was determined to be residual GST (\( M_W=25\text{kDa} \)) dimer. Secondary structure analysis using circular dichroism (CD) spectroscopy showed negative bands at 222nm for purified GST-STAS\textsubscript{436} and STAS\textsubscript{436} typical of a protein with an \( \alpha \)-helical content (Figure 8). To confirm that the STAS domain was properly folded, the tryptophan emission spectrum of the STAS domain was measured with increasing concentrations of guanidinium. In the absence of the denaturant, a maximum fluorescence intensity
was observed at ~ 335 nm which red-shifted to 357 nm upon complete denaturation of the protein and exposure of the tryptophan to solvent (Figure 9).

The truncated STAS domain (STAS$_{447}$), missing the first β-strand was found to be unstable in solution after the GST was cleaved. Therefore, STAS$_{436}$ and GST-STAS$_{436}$ were used for all subsequent experiments.

ACP was purified using Ni$^{2+}$-affinity chromatography. The highly anionic ACP migrated slowly on SDS-PAGE with an apparent molecular weight of 17kDa. The gel filtration elution profile shows two peaks at elution volumes ~11 and ~12 ml (Appendix 4). These peaks correspond to 17kDa monomer and ~ 34kDa dimer protein on non-reducing SDS-PAGE (Appendix 4). Apo-ACP and holo-ACP migrate differently on native-PAGE with the latter running slower than the former (Figure 10). Moreover, holo-ACP forms a dimer via the sulfhydryl groups of the 4’-phosphopantethiene prosthetic group, indicating the presence of the prosthetic group. In vitro synthesis of malonyl-ACP resulted in holo-ACP with attached 4’-phosphopantethiene prosthetic group as a major species as confirmed by mass spectrometry (Appendix 5). This is likely due to the instability of the thioester bond that conjugates the malonyl to the –SH group of the phosphopantethiene group. Hence, holo-ACP was used for biophysical assays mentioned below.

3.1.2 Surface Plasmon Resonance (SPR)

The study of molecular interactions using surface plasmon resonance exploits an optical phenomenon observed during light-induced excitation of free electrons on a metal layer$^{112}$. The transfer of energy from photons to electrons of the metal surface creates oscillating electron density that is known as surface plasmon. Resonance occurs at an angle of incidence where the
energy of the photons equals the energy of the plasmons resulting in attenuated incidence light. 

The angle at which this minimum intensity of light is experienced is referred to as the SPR angle \(^{112}\). The binding study of molecules using SPR measures this change in SPR angle as a function of protein complex formation. An attempt to bind ACP onto Carboxy-methyl (CM) dextran surface failed due to repulsion of ACP by the negatively charged carboxyl groups. When the STAS domain was immobilized, binding studies were deterred by the aggregation of the protein on the chip.
Figure 7 Purification of GST-STAS and STAS domain by size exclusion chromatography. (a) Peak fractions from glutathione-affinity chromatography of GST-tagged STAS were pooled and loaded onto a Superdex200 gel filtration. (b) GST-tagged STAS domain incubated with PreScission protease was loaded onto a Superdex75 column. The STAS domain elutes at 13mL from a Superdex75 column as determined by coomassie-stained SDS-PAGE.
Figure 8 Secondary structure analysis of the STAS domain. Circular dichroism (CD) spectra of purified GST-tagged (a) and tag-free STAS domain (b). CD measurements from 200-260nm were performed in 10 mM KCl and 100 mM K₃PO₄ buffer at 10°C.
Figure 9 The STAS domain exhibits co-operative folding. Guanidine (abbreviated Gnd) denaturation assay of the STAS domain of YchM. The intrinsic fluorescence of 5 µM of STAS in the presence of 0.5-6 M guanidine was measured at an excitation wavelength of 295nM. (a) illustrates the tryptophan emission spectra of the STAS domain at different concentration of guanidine. (b) shows the wavelength of maximum tryptophan fluorescence as a function of guanidine concentration.
Figure 10 Native-PAGE of apo-ACP, holo-ACP dimer and holo-ACP monomer. After size-exclusion chromatography, peaks corresponding to dimer and monomer of ACP were run on Native-PAGE. Lane 1 shows the monomer fraction corresponding to ACP_{S36A}. Holo-ACP dimer is shown in lane 2. Addition of DTT reduced the dimer to monomer holo-ACP (lane 3).
3.1.3 Biolayer interferometry (BLI)

Bio-layer interferometry was used to determine the binding constant of the ACP-STAS interaction. BLI is a technique that measures change in an interference pattern generated from visible light reflected from a reference optical layer and a biolayer containing the immobilized protein of interest. The change in interference pattern that results upon interaction between the protein immobilized on a biolayer (probe) and a binding partner (analyte) is used to determine a binding constant. In this experiment, a biotinylated STAS domain bound to a streptavidin biosensor was used as a probe to analyze its interaction with ACP. Prior to the BLI experiment, ACP concentrations were estimated based on $A_{280}$ measurements. To find out the dissociation constant of the ACP-STAS interaction, the binding of increasing concentration of holo-ACP (27 nM to 20 µM) to the STAS bound sensors was analyzed. As shown in Figure 11, the binding signal increased with increasing concentration of ACP indicating a binding interaction. The data was globally fitted to a heterologous-binding model and gave a dissociation constant of 3 µM ($R^2 = 0.933$). The data quality could at best predict the $K_d$ to be in micro-molar range.

![Figure 11](image.png)

**Figure 11** The binding constant for the ACP-STAS interaction is determined to be in the micro-molar range. Bio-Layer interferometry raw data of biotinylated-STAS domain incubated with 27.4 nM -20 µM ACP. The response unit increased for the interaction increased with increasing ACP concentration.
3.1.4 Affinity Pull-down

To qualitatively determine the ACP-STAS interaction, an affinity pull-down was performed with GST-STAS as bait. GST-tagged wild-type STAS, His-tagged ACP and GST (negative control) were over-expressed in BL21(DE3) cells. ACP-(His)$_6$ lysate was mixed with either GST-STAS or GST-containing lysates. As shown in Figure 12, GST-STAS but not GST alone co-purified with ACP from the lysate mixture. ACP lysate sample was run as a positive control (Figure 12, lane 1).

![Co-purification of ACP with GST-STAS.](image)

3.1.5 Tryptophan Fluorescence Assay

The presence of a single tryptophan (Trp485) in the STAS domain of YchM and the absence of tryptophan in ACP makes a tryptophan fluorescence assay a potential method of analyzing the ACP-STAS interaction. In order to determine the role of the 4’-phosphopanthetiene group of holo-ACP, the intrinsic tryptophan fluorescence of the STAS
domain was studied in the presence of holo-ACP. In order to determine the dissociation constant ($K_d$) of the ACP-STAS interaction, 5 µM STAS was incubated with various concentrations of holo-ACP, and intrinsic tryptophan fluorescence of the STAS domain was measured by excitation at 295 nm.

As shown in Figure 13, the tryptophan fluorescence intensity of the STAS domain did not exhibit a concentration dependent decrease with increasing concentration of ACP suggesting that the manner in which this protocol was performed would not provide the kinetic information we sought.

**Figure 13** Holo-ACP does not alter the tryptophan emission spectra of the STAS domain. Intrinsic tryptophan fluorescence of the STAS domain in the presence of various concentrations of holo-ACP. 5 µM of STAS was incubated with 5-40 µM concentrations of ACP in 20 mM HEPES, pH 8, 100 mM NaCl and 1 mM TCEP for 1 hour. The samples were excited at 295 nm and emission spectrum was collected for wavelengths 300 to 400 nm. Fluorescence measurements were performed at 25 °C.
3.1.6 Isothermal Titration Calorimetry

In order to determine the dissociation constant of the ACP-STAS interaction, 175 µM STAS$_{436}$ was titrated into 40 µM holo-ACP. The resulting titration curve was fit to a 1 binding site model and a dissociation constant of 0.7 +/- 0.1 µM was obtained (Figure 14a). To assess the role of the phosphopantathiene group in the ACP-STAS interaction, ACP$_{S36A}$, a mutant that can not be post-translationally modified to holo-ACP, was used as analyte. As shown in Figure 14b, ACP$_{S36A}$ did not interact with the STAS domain. In addition to the interaction between 4’-phosphopantheniène and the STAS domain, the crystal structure of the ACP-STAS complex shows hydrogen bond interactions between R523 and R527 of the STAS domain with acidic residues (D57 and E48) of ACP. Hence, the interaction of two STAS domain mutants, STAS$_{R527E}$ and STAS$_{R523E}$, with holo-ACP were also assessed. As illustrated in Figure 15, both these STAS domain mutants did not interact with ACP.
Figure 14 Isothermal titration calorimetry indicates $K_d = 0.7 \, \mu$M for the STAS-holoACP interaction. ITC curves for holoACP-STAS (a) and apoACP-STAS (b) interaction. Aliquots of 10 µL of 175 µM STAS were titrated into 40 µM holo-ACP or apo-ACP with 5 minute intervals. A total of 20 injections were made per experiment keeping a constant stirring speed of 300 rpm. Titration curves were fitted using one binding site model in the MircoCal Origin software.
Figure 15 STAS\textsubscript{R523E} and STAS\textsubscript{R527E} do not interact with holo-ACP. ITC curves for holoACP-STAS interaction. Aliquots of 10 µL of 175µM STAS\textsubscript{R523E} (a) or STAS\textsubscript{R527E} (b) were titrated into 40 µM holo-ACP with 5 minutes intervals. A total of 20 injections were made per experiment keeping a constant stirring speed of 300 rpm. Titration curves were fitted using one binding site model in the MircoCal Origin software.
3.2 *In vivo* development of functional assay for YchM

3.2.1 pH Growth Assay

To investigate the potential role of YchM for growth in alkaline conditions, wild-type and ychmΔ BW25113 and MG1655 *E. coli* cells were grown in various LB media buffered with MOPS and Tricine to pH 6.5, 7, 8.5 and 9. As shown in Figure 16, deletion of *ychm* did not result in growth defect at these pH values. The same results were observed in an alkaline growth assay using HfrC cells (Appendices 6).

3.2.2 Fatty acid growth assay

Genetic and physical interaction of YchM with fatty acid degradation components such as FadE and FadJ implicate YchM’s role in fatty acid degradation. Since the outer membrane fatty acid transporter (FadL) is selective for long-chain fatty acids over medium chain, oleic acid was chosen as a carbon source. The ability of wild-type and ychmΔ cells to utilize oleic acid as their carbon source was tested. As shown in Figure 17 ychmΔ did not exhibit any growth defect.
Figure 16 The role of YchM under acidic and alkaline conditions. WT (blue) and ychm Δ (red) BW25113 E coli cells were grown in LB media of pH 5.5, pH 6.5, pH 7, pH 8.5 and pH 9. Overnight cultures of wild-type and ychm knock out cells were used to inoculate 100µL of LB media under the different pH conditions. Prior to inoculation, aliquots of overnight cultures were adjusted to an OD<sub>600</sub>=0.1 and a serial 1:10 dilution was performed in 100-well titer plates. OD<sub>600</sub> measurements were made every 15 minutes for 24hrs.
Figure 17 The role of YchM in fatty acid degradation. Wild-type and ychm Δ BW25113 E.coli cells were grown overnight in LB rich media and used to inoculate a 100-well titer plate containing 100 μL of minimal media supplemented with 0.2% oleate.
Chapter 4 Discussion
YchM is the only member of the SLC26 family of anion transporters found in *E.coli*. It is a putative bicarbonate transporter that is predicted to span the inner membrane 12 times, with both the N and C termini facing the cytosol. Akin to most members of the SLC26 family, YchM possesses a C-terminal cytosolic STAS domain. Previous genetic and structural studies of the STAS domain of YchM have identified the fatty acid biosynthesis co-factor, ACP, as a binding partner. However, the functional consequence of this interaction has not been fully characterized. This study sought to shed light into the structural and functional relationship between YchM and ACP. The first part of the study focused on identifying critical residues for the ACP-STAS interaction as indicated by the co-crystal structure.

First, a stable STAS domain with defined N-terminal boundary was required. Two STAS domain constructs, STAS<sub>436</sub> and STAS<sub>447</sub>, were purified as GST-tagged fusion proteins. Since STAS<sub>447</sub> was insoluble in solution, the N-terminal boundary of the STAS domain was determined to be Met436. The STAS domain was purified as a monomer with an elution volume of ~13mL on a Superdex 75 gel filtration column (*Figure 7*). In agreement with the crystal structure of the STAS domain of YchM, secondary structural analysis using circular dichroism spectroscopy illustrated a negative peak at 222nm that corresponds to a helical content. Furthermore, thermal (Appendix 3) and guanidine denaturation assays demonstrated cooperative folding of the STAS domain (*Figure 9*).

ACP was purified using Ni<sup>2+</sup> affinity chromatography followed by gel filtration. During this purification, both *holo*-and *apo*- forms of ACP were obtained as confirmed by native-PAGE (*Figure 10*). Compared to *apo*-ACP, *holo*-ACP (ACP with 4’-phosphapantethiene group) migrates slower on native-PAGE. Furthermore, holo-ACP forms a dimer via the sulfhydryl group of the prosthetic group. Since ACP does not have any intrinsic cysteine residue, native-
PAGE electrophoresis was proven to be a quick method of identifying holo-ACP dimers, holo-ACP and apo-ACP.

Once the STAS domain, ACP and their respective mutants were purified, the ACP-STAS interaction was analyzed using various biochemical and biophysical assays. The selection of an ideal method of protein-protein interaction studies depends on the properties of the proteins of interest and desirable results (quantitative or qualitative) among other factors. First, the ACP-STAS interaction was qualitatively determined using affinity pull-downs. As shown in Figure 12, GST-tagged STAS domain co-purified with His-tagged ACP validating the crystal structure.

To quantitatively determine the binding constant of the ACP-STAS interaction, surface plasmon resonance was initially identified as a method of choice. In this technique, a protein of interest (probe) is immobilized on a metal film (usually 50nm gold) that is placed on a glass prism and exposed to a dielectric medium (water phase) that contains the binding partner (analyte). Polarized light travelling via the prism is directed towards and reflected from the metal-probe surface. The SPR angle, which is sensitive to change in refractive index, depends on the amount of protein near the surface of the immobilized protein. Upon probe-analyte binding, a change in SPR angle is recorded in real-time and later converted to binding kinetic parameters. Using SPR, binding constants can be determined with low quantity of proteins.

To determine the dissociation constant of the ACP-STAS interaction, an attempt was made to couple apo-ACP and holo-ACP onto a carboxymethyl dextran sensor chip. However, the negatively charged carboxyl groups of the CM dextran sensor chip suppressed binding of the highly anionic protein ACP (pI ~4) to the surface. This underlines one of the challenges of binding studies using SPR – the effect of the sensor surface on the immobilized protein. Proper binding analysis using SPR requires one protein to be stably immobilized with its binding
epitopes presented in their native state. This makes the choice of immobilization chemistry and the sensor surface critical. For the STAS-ACP interaction amine coupling was used to crosslink either ACP or STAS onto the carboxymethyl surface. Immobilization chemistry was likely not an issue since the crystal structure of ACP shows four surface-exposed lysine residues with three of them found opposite the STAS binding interface. When attempts to immobilize ACP failed, perhaps due to repulsion of ACP by carboxyl groups of the binding surface, the STAS domain was used as a probe. After immobilization of the STAS domain onto the chip, however, increasing concentration of ACP did not result in change in response units. This is likely due to either the aggregation of the STAS domain on the chip or mass transport limitation. Aggregation of the immobilized protein, in this case the STAS domain, blocks potential binding sites. Efficient mass transfer between the flowing buffer and the surface is also vital for accurate SPR studies\textsuperscript{114}. Otherwise, the analyte concentration near the immobilized protein can be retained or depleted resulting in different local concentration of the analyte compared to the bulk solution. Finally the affinity of free ACP for immobilized STAS domain may be too weak to be detected by SPR.

Bio-layer interferometry, a technique that does not rely on micro-fluidic dynamics, was used to overcome the issue of mass transport limitation. BLI is also an optical technique that measures change in phase as generated from interference of visible light reflected from a reference optical layer and a bio-layer containing proteins of interest. In this method, light travels via an optic fiber sensor probe containing bio-layer (Streptavidin, Ni-NTA etc) and a reference layer. The light reflected back from this layer has an interference pattern that is dependent on the thickness of the biolayer. The change in interference pattern (phase) that results upon protein complex formation at the bio-layer is used to calculate binding parameters. Just like SPR, BLI is a very sensitive technique that enables label-free, real-time monitoring of molecular interactions.
Moreover, it requires low concentration of protein for analysis. Unlike SPR, the Octet® BLI instrument does not require the flow of bulk solution on the immobilized protein of interest. Instead, the protein is coupled to a sensor probe and immersed into a solution containing the binding partner. This difference in techniques makes BLI less susceptible to inefficient mass transfer. To analyze the ACP-STAS interaction, biotinylated STAS domain was immobilized on streptavidin sensors and the binding response was studied in the presence of ACP. As shown in Figure 11, higher concentrations of ACP resulted in increased response units. Binding analysis using a 1:1 binding model predicted a $K_d$ in micro-molar range. However, the raw binding data fit poorly to the model and definitive binding constant could not be calculated. Even though BLI is not influenced by mass transport limitations, protein aggregation can still affect the results.

The ACP-STAS interaction was then assessed using a tryptophan fluorescence assay. This method allows the detection of protein interaction using an intrinsic fluorophore - tryptophan. Tryptophan’s emission peak ranges from 330-360 nm depending on the hydrophobicity (polarity) of the environment it resides. Exposure of tryptophan to aqueous environment results in red shift and vice versa. Unlike SPR and BLI, immobilization of one of the binding partners is not required for the tryptophan fluorescence assay, which means the negative affects of tethering a substrate to a surface and potentially blocking binding sites are avoided. The STAS domain has a single tryptophan residue (Trp485) while ACP does not possess any tryptophan residues, making the two proteins ideal candidates for tryptophan fluorescence assay. If there is an interaction between ACP and STAS, it may result in a change in tryptophan emission spectrum of STAS as a result of quenching or a shift in wavelength. In order to determine the dissociation constant ($K_d$) of the ACP-STAS interaction, 5μM of STAS was incubated with various concentrations of ACP, and intrinsic tryptophan fluorescence of the STAS domain was measured by excitation at 295nm. As shown in Figure 13, the tryptophan fluorescence intensity of the
STAS domain did not exhibit a concentration-dependent decrease with increasing concentration of ACP. This result implies: 1) Trp485, which is at the core of the STAS domain, might not be affected by any conformational change the ACP-STAS interaction induces or 2) there is no interaction between ACP-STAS. The former is a more likely explanation since previous BLI and affinity pull-down experiments have shown an interaction. Hence, the ACP-STAS interaction was analyzed using Isothermal Titration Calorimetry (ITC).

ITC operates by measuring the change in enthalpy during a reaction. Power is supplied to thermocoupled sample and reference cells to maintain the same temperature $^{115}$. When a protein of interest is titrated into the sample cell, the power required to sustain the temperature of the sample cell changes relative to that of the reference cell. This change in power (delta power, DP) is the result of the difference in enthalpy that occurs upon protein binding. In a single titration experiment, the enthalpy, entropy and the Gibbs free energy of a binding reaction can be determined $^{115}$. These thermodynamic parameters are used to calculate the stoichiometry and association constants of a binding reaction. The caveat of using ITC is that a significantly larger quantity of protein is required compared to methods mentioned above.

In order to analyze the ACP-STAS interaction by ITC, the STAS domain was titrated into apo-ACP (ACP$^{S36A}$) or holo-ACP. The dissociation constant for the holo-ACP-STAS interaction was determined to be $0.7 \pm 0.1 \, \mu\text{M} (\text{Figure 14a})$. This result is in agreement with previous bio-layer interferometry results that predicted the $K_d$ to be in the micro-molar range. Based on the ITC raw data (Figure 14) titrating ACP$^{S36A}$ into wild-type STAS domain did not result in change in heat energy indicating the two proteins do not interact. The crystal structure of the ACP-STAS complex depicts the prosthetic group of ACP in a hydrophobic groove created by two helices of the STAS domain. In the absence of this prosthetic group, as in the case of ACPS36A, there was
no change in heat energy detected. This result indicates the 4’-phosphapantethiene prosthetic group of holo-ACP is crucial for the ACP-STAS interaction. Further analysis of the ITC result indicates a stoichiometry of N=0.3. This result deviates from the 1:1 stoichiometry shown by the crystal structure. Perhaps, the dynamic structural conformation of holo-ACP accounts for this discrepancy. Holo-ACP has been shown to exist in dynamic conformational equilibrium where the prosthetic group is either tucked in between two helices of ACP or protrudes away from the protein towards the solvent. Thus, in the presence of two structurally heterogeneous holo-ACP species, the STAS domain would interact with the species that has its prosthetic group exposed and accessible for binding. The holo-ACP that has its prosthetic group hidden possess a structure that is similar to apo-ACP. ACP’s conformational flexibility is thought to be imperative for substrate delivery to different enzymes. Once an acyl chain is conjugated to the prosthetic group, the hydrocarbon is hidden between helices II and III of ACP. The only exception is malonyl-ACP where the negatively charged carboxyl group of the malonyl is always solvent exposed and accessible for binding. Thus, in malonyl-ACP the prosthetic group is locked in a conformation that favors protein-protein interaction. Thus it is predicted that this acylated derivative of ACP would have a higher affinity for STAS, however making large enough quantities of malonyl-ACP \textit{in vitro} has been difficult. In contrast, the addition of longer acyl chains (longer than 3 carbons) would be expected to hinder ACP-STAS interaction. The discrepancy in stoichiometry might also be due to uncertainties in ACP concentration estimation since ACP does not contain Trp residues.

Once the $K_d$ for the ACP-STAS interaction was determined using ITC, the roles of two STAS domain residues were also analyzed using the same technique: R523 and R527. In most of the ACP-enzyme complexes that have been characterized (ACP and AcpS, FabD, FabH, FabG and FabI) arginine residues have been identified to be crucial for protein-protein interactions.
For example, mutations at R129 and R192 of the fatty acid biosynthetic enzyme, FabG, inhibit the enzyme’s interaction with ACP. In the crystal structure of the ACP-STAS(YchM) complex, R523 in the STAS domain forms hydrogen bonds with two residues of ACP, E48 and D57. The side-chain amide group of R527 of STAS also makes a salt-bridge with E48 of ACP. The raw ITC data (Figure 15) indicates that R523 STAS domain mutant (STAS_{R523E}) does not bind to holo-ACP. Similarly, no interaction was observed between holo-ACP and STAS_{R527A}. Guanidine denaturation assay showed that the arginine mutation did not completely disrupt the conformation of the STAS domain. A reciprocal mutation in ACP (ACP_{D47R}) also deterred interaction confirming the result was due to inhibition of the ACP-STAS interaction (Appendices). Taken together, these observations lead to the conclusion that R523 and R527 of the STAS domain are important for binding ACP. Hence, the STAS domain mimics the characteristic surface pattern displayed by ACP partner enzymes, which consists of electropositive/hydrophobic patches. In the STAS domain, the hydrophobic patch is the groove that interacts with the 4’phosphopantheine as mentioned above. The electropositive arginine residues bind the highly conserved electronegative residues of ACP.

The functional characterization of the YchM-ACP interaction requires development of an assay that qualitatively or quantitatively measures YchM’s activity. This functional assay could then be used to assess the role of R523 and R527 in YchM’s transport or other activities. Accordingly, two growth conditions were targeted: growth in high pH where bicarbonate is present and in media with fatty acids as the sole carbon source. Previously, Hfr C cells lacking YchM have been shown to exhibit growth defects in alkaline conditions. Accordingly, ychm\textDelta Hfr C cells grown in alkaline pH (8.3) exhibited a 13-fold reduction in growth compared to wild-type cells. To further investigate the role of YchM for growth in different pH’s, WT and ychm\textDelta BW25311, ychm\textDelta MG1655 E. coli cells were grown in various pH between 7 and 9. There was
no growth defect in any of the strains, suggesting that YchM does not confer selective advantage for growth under alkaline condition. HfrC E.coli cells are unstable and undergo frequent genetic recombination. Mutations or recombination events that involve essential genes during the growth assay can result false positive phenotypes. This might explain the previously reported growth defect in ychmΔ Hfr C cells grown in alkaline pH.

To explore the link between YchM and fatty acid degradation, cells devoid of YchM were grown on fatty acid as their sole carbon source. Long chain fatty acids can be imported into the cytosol of bacteria where they enter the β-oxidation pathway. As shown in Figure 17, ychmΔ BW25113 cells utilized oleic acid as efficiently as the wild type cells. Hence, YchM is not necessary for fatty acid degradation. Perhaps, YchM becomes essential in the absence of one or more of the default fatty acid metabolic enzymes, even when the cells depend on fatty acid as their carbon source. Further studies are required to explore the role of YchM in strain devoid of a fatty acid degradation component.
Chapter 5 Conclusion and Future Directions
The co-purification and co-crystallization of endogenous ACP with the STAS domain implicated the involvement of YchM in fatty acid metabolism. In order to investigate the specificity of the interaction, two approaches were used: 1) identification of residues of YchM important for the ACP-YchM interaction, and 2) the development an *in vivo* assay used to study the role of residues rendered to be binding determinants. Analysis of the ACP-STAS interaction with different biophysical assays has identified isothermal titration calorimetry as the ideal method of study. Using ITC, the dissociation constant for STAS-holoACP interaction was calculated to be $\sim 1\mu\text{M}$. Moreover, the 4’-PPa group of holo-ACP and the hydrogen bond donors, R523 and R527, of the STAS domain were found to be essential for the ACP-STAS interaction in agreement with the predictions of the crystal structure. *In vivo* pH and fatty acid growth assays indicate that YchM does is not essential for growth in alkaline pH or utilization of fatty acids as carbon source.

5.1 Future *in vitro* analysis of the ACP-STAS interaction

The ACP-STAS interaction can be further characterized to study the specificity of the STAS domain for different acylated derivatives of ACP. As discussed above, ACP exhibits conformational flexibility depending on the type of acyl-chain or prosthetic group conjugated to it. Based on the ITC results reported above, a solvent exposed prosthetic group of ACP is essential to bind the STAS domain. To test this hypothesis, the binding affinity of the STAS domain to malonyl-ACP and decanoyl-ACP could be assessed. It is expected that the STAS domain will have higher affinity to malonyl-ACP, a derivative with a protruding prosthetic group. Decanoyl-ACP, which has its decanoyl-4’-PPa group hidden between helices II and III of ACP is not expected to interact with the STAS domain of YchM. The crystal structure of the
ACP-STAS domain shows Gln531 of STAS forming two hydrogen bonds with the carboxyl group of malonyl-ACP. The role of this residue for the ACP-STAS interaction can be analyzed by mutating this residue to Glu, Asn or Ala. In addition, the ACP-STAS crystal structure illustrates STAS domain residues Phe 539, Phe 540, Pro 532, Pro 541 and Ile 530 make direct contact with the 4’-PPa group. These residues are found within the hydrophobic pocket of the STAS domain that houses the prosthetic group of ACP during the proteins’ interaction. The affinity of malonyl-ACP to STAS domain variants harboring mutations at any of those residues could be analyzed using ITC.

5.2  *In vivo* functional assay of YchM

5.2.1  *YchM’s substrate selectivity*

The physiological role of YchM and the biological importance of the ACP-STAS interaction remain unclear. eSGA results indicated that cells lacking YchM and one component of the fatty acid degradation pathway (FadL, FadD, FadE, etc) are non-viable. This phenotype can be exploited to elucidate the role the STAS domain as well other residues of interest. A rescue approach can be undertaken to complement double knockouts harboring mutations in genes encoding for one of the fatty acid components (eg. FadD) and *ychm*. It is expected that a plasmid expressing wild-type YchM or FadD would rescue the knockouts. The residues that are important for YchM’s proper function can be identified through an attempt to rescue the double knockouts with various YchM mutants such as a STAS deletion. Lethality of the double knockouts in the presence of YchM mutant will demonstrate the importance of the mutated residue.
YchM mutations can either hinder the folding or insertion of the protein into the membrane or directly inhibit the transporter’s function. To identify mutants that affect proper folding or membrane insertion, the expression level of YchM mutants at the inner membrane can be assessed. Inner membrane fraction can be separated using sucrose-gradients and immuno-blotted for YchM. Mutants that affect folding or insertion will show reduced level of YchM in the membrane. If the level of a Ychm mutant in the membrane is similar to wild-type, the specific mutant can be deemed important for YchM’s function.

YchM is a putative bicarbonate transporter. However, the transporter’s potential in allowing the passage of other anions has not been studied. As shown in Table 1, most human SLC26 proteins, such as SLC26A3, transport more than one solute. Thus, a better understanding of YchM’s substrate affinities and specificities can shed light into YchM’s function. The uptake of radiolabelled substrates (e.g. $^{35}$S sulfate) into intact bacteria could be measured. If YchM-dependent uptake is demonstrated, competition studies with other anions can be performed. Electrophysiological studies can be performed on YchM expressed in E. coli spheroplasts or in the heterologous Xenopus oocytes expression system. Transport activities in the presence of different anions can be studied using 2-electrode voltage clamp. Once a transport assay has been developed, site-directed mutagenesis can be carried out to analyze the role of different residues in YchM’s transport activity. The transport efficiency of YchM lacking the STAS domain or with mutations in the domain can also be assessed.
References


(13) Melvin, J. E.; Park, K.; Richardson, L.; Schultheis, P. J.; Shull, G. E. Mouse down-regulated in adenoma (DRA) is an intestinal Cl(-)/HCO(3)(-) exchanger and is up-regulated in colon of mice lacking the NHE3 Na(+)/H(+) exchanger. *J. Biol. Chem.* **1999**, *274*, 22855-22861.


(89) Zhang, Y. M.; Marrakchi, H.; White, S. W.; Rock, C. O. The application of computational methods to explore the diversity and structure of bacterial fatty acid synthase. J. Lipid Res. 2003, 44, 1-10.


### Appendices

#### Appendix 1 Primers

<table>
<thead>
<tr>
<th>Construct</th>
<th>Primer name and sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>GST-STAS\textsubscript{436}</td>
<td>5'ACGGGGATCCATGACTCGCCTGGCACCAGG 5'CCGGGTCACTTTATAAATCCGCCATCGCC</td>
</tr>
<tr>
<td>GST-STAS\textsubscript{447}</td>
<td>pGEX_447STAS F 5' TTC CAG GGG CCC CTG GGA TCC AUG CCA GAC GAT GTC CTG GTT CTG CGC 3' pGEX_447STAS R 5' GGCAGAACCAGGACATCGTCTGGCATGGATCCCAGGGGCCCCCTGGAA 3'</td>
</tr>
<tr>
<td>STAS\textsubscript{R523E}</td>
<td>R523E F 5' GAA TTC CAG CCA CTG GAG ACT ATG GCT CGC GCA 3' R527E R 5' TGC GCG AGC CAT AGT CTC CAG TGG CTG GAA TTC 3'</td>
</tr>
<tr>
<td>STAS\textsubscript{R527E}</td>
<td>R527E F 5' CTG CGC ACT ATG GCT GAG GCA GGC ATT CAA CCG 3' R527E R 5' CGG TTG AAT GCC TGC CTC AGC CAT AGT GCG CAG 3'</td>
</tr>
<tr>
<td>ACP\textsubscript{S36A}</td>
<td>S36A ACP F 5' GAAGACCTGGGCGCGATGGCTTG S36A ACP R 5' CAGCTACAAGGTGTCAAGCAGCATCC</td>
</tr>
<tr>
<td>ychm\textsubscript{Δ} Colony PCR</td>
<td>pKD3CAT F: 5' GTGTAAGGCTGGAGCTGCTTC 3' YchM-100 R: 5' CTATGACAAACGTGTGGCCG</td>
</tr>
<tr>
<td>ychm\textsubscript{Δ} Colony PCR</td>
<td>YchM+140 F 5' CATGTTCGAA CACTAATCGA AC 3' pKD3CAT R 5' ATGGTCCATATGAATATCCTCTCC</td>
</tr>
</tbody>
</table>

86
Appendix 2  Size exclusion elution profile of STAS_{R527E} (a) and STAS_{R523E} (b). Peak fractions from glutathione-affinity chromatography of GST-tagged STAS mutants were pooled and incubated with PreScission protease. The cleavage products were and loaded onto a Superdex75 (a) or Superdex200 (b) gel filtration columns.
Appendix 3 Thermal denaturation of the STAS domain.

Appendix 4 Purification of ACP by size exclusion chromatography. (a) Fractions collected from anion exchange chromatography were loaded onto Superdex75 column that was equilibrated with 10mM HEPES and 50mM NaCl, pH 8 buffer. Coomassie stained SDS-PAGE of peak fractions of gel filtration of ACP (b) ACP monomer and dimers were separated using Superdex75 column.
Appendix 5 Mass spectrometry results illustrate the predominant ACP species is holo-ACP. Apo-ACP was incubated with malonyl-CoA and Acp synthase. The molecular weight of the product was 11012.50 Da, higher than the mass of apo-ACP (10671.59 Da). 358 Da corresponds to the molecular weight of the 4'-phosphopantethiene group.

Appendix 6 YchM does not provide selective advantage for growth in alkaline pH. Growth curve of WT, YchM Δ, STAS Δ, and YchM Δ Hfr C E. coli cells complemented with pET-41b plasmids expressing YchM (YchM:KO~YchM) and TMD domain of YchM (YchM:KO~STAS:KO). Cells were grown in LB media of pH 7 and pH 8.3 containing 100mM HEPES.
Appendix 7 The role of YchM under alkaline conditions. WT (blue) and ychm Δ (red) MG1655 E coli cells were grown in LB media of pH 7, pH 8, pH 8.5, and pH 9 buffered with 40mM MOPS/Tricine. Overnight cultures of wild-type and ychm knock out cells were used to inoculate 100µL of LB media under the different pH conditions. Prior to inoculation, aliquots of overnight cultures were adjusted to an OD$_{600}$=0.1 and a serial 1:10 dilution was performed in 100-well titer plates. OD$_{600}$ measurements were made every 15 minutes for 24hrs. Growth differences observed at pH 8.5 and 9 are not significant and not represented in other dilution series.