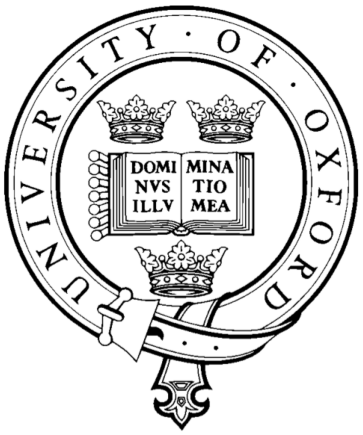
Structural Studies of Signal Transduction in Bacterial Chemotaxis and A Metabolism-Related Cancer Target



Madhvi Jayalakshmi Venkatesh

Oriel College

University of Oxford

Michaelmas Term 2015

A thesis submitted for the degree of

Doctor of Philosophy

Abstract

Structural Studies of Signal Transduction in Bacterial Chemotaxis and Metabolism-Related Cancer Targets

Madhvi Jayalakshmi Venkatesh

Oriel College, University of Oxford
Submitted for the degree of Doctor of Philosophy, Michaelmas Term

Transmission electron microscopy (TEM) has had a growing impact on the life sciences as both technological and methodological developments have allowed for cells and macromolecules to be resolved at increasingly higher resolutions. Variations of TEM can be used to visualize cellular ultrastructure as well as to determine macromolecular structure in a near-native state.

In this work, a number of TEM-based methods were utilized to structurally characterize proteins from different biological systems. The areas of study were the chemotaxis systems of *Escherichia coli* and *Rhodobacter sphaeroides* as well as human cancer metabolism. In each system, a key protein of interest was selected for investigation.

Bacteria use their chemotaxis systems to dynamically sense and respond to their environment by biasing their movements towards favorable directions. The *E. coli* serine chemotaxis receptor (Tsr) was studied using cryo-electron tomography of arrays formed by overexpressed Tsr in a thin *E. coli* cell strain and in membrane isolates. Subsequent sub-volume averaging yielded structures of the full-length receptor which showed that serine binding is accompanied by expansion of the signaling domain.

The protein selected for study in the *R. sphaeroides* chemotaxis systems, PpfA, is an ATPase responsible for segregating the cluster of cytoplasmic chemosensory proteins. Immuno-gold labeling of cryo-sections showed that both wild-type and mutant PpfA demonstrated some degree of clustering into foci.

The enzyme pyruvate kinase M2 (PKM2) plays an important role in various cancers by regulating metabolism to promote cell survival and propagation. Single particle cryo-electron microscopy was used to perform a structural comparison of wild-type PKM2 with a patient-derived mutant form of the protein. At the resolution of the structures solved, the molecular architecture of the patient and mutant forms of the protein were the same, except for the mutated residue. The structures revealed the potential presence of a stabilizing stacking interaction in the mutant that may explain differences in its biochemical behavior.

These studies illustrate the range of biological questions and systems to which TEM-based techniques can be applied with the goal of gaining structure-based mechanistic insights. For macromolecular targets of clinical relevance such as PKM2, such information can be used for rational drug design.

Declaration

The work described in this thesis was carried out in the Department of Biochemistry at the University of Oxford and the U.S. National Institutes of Health between October 2011 and January 2016. Except where specifically stated, this dissertation is the outcome of my own work and not the result of any collaboration. No part of this thesis has been submitted for a qualification at any other university.

Madhvi Jayalakshmi Venkatesh

Acknowledgements

I would like to thank Prof. Judy Armitage and Dr. Sriram Subramaniam for their support, mentorship and guidance throughout the DPhil process. They found the nuanced balance between guiding me towards meaningful projects and giving me the freedom to follow my own interests and intuitions.

I have had the privilege of working with wonderful colleagues in both the Armitage and Subramaniam labs. They were instrumental in helping to make my DPhil experience productive, enjoyable and intellectually stimulating. I would especially like to thank Dr. Soojay Banerjee, Alan Merk, Prashant Rao, Dr. Chris Jones, Lisa Hartnell, Dr. Doreen Mattheis, Dr. Lesley Earl, Dr. Joel Meyerson, Dr. Kathryn Scott, Dr. Gabriel Frank, and Dr. Kedar Narayan who were all very generous in sharing their time, thoughts, and expertise with me. I would also like to Dr. Alberto Bartesaghi, Dr. Xiongwu Wu, and Dr. Mario Borgnia for developing software that enabled my work and introducing me to cryo-EM image processing and its underlying concepts.

I am very grateful to my parents, Vidya and Warren Venkatesh, and my brother, Anirudh Venkatesh, who have been constant sources of support and encouragement throughout my various pursuits over the years. I am very fortunate that I have been surrounded by caring and compassionate friends who have helped keep me going throughout my journey in lab and in life. These include Jason Murray, Dr. Thao Do, Kasi Aysola, Dr. Yamini Saripalli, Jenna Shapiro, Jamie Solis, Sudha Radhakrishnan, and Shailaja Maru Jauhari. Most importantly, I am indebted to my spiritual gurus, Sri Shirdi Sai Baba and Sri Vishwanathan Srinivasan for keeping the successes and challenges of life in perspective.

I also thank the Oxford Biochemistry Department and the intramural program at the National Institutes of Health for generously funding my work.

Contents

Abstract i

Acknowledgements iii

List of Figures vii

List of Tables ix

Abbreviations x

I General Introduction 1

1 Introduction to Principles and Methodology of Structural Biology 2
1.1 Protein Structure . . . 2
1.2 Functional Implications of Protein Structure . . . 5
1.3 Structural and Imaging Methods in Biology . . . 8
1.4 Transmission Electron Microscopy . . . 10
1.4.1 Cryo-Electron Microscopy and Single Particle Analysis . . . 13
1.4.2 Cryo-Electron Tomography and Sub-volume Averaging . . . 19
1.5 Choice of Methods for the Questions Addressed in this Thesis . . . 22
1.5.1 Bacterial Chemotaxis . . . 23
1.5.2 Human Cancer Metabolism . . . 23

II Bacterial Chemotaxis 25

2 Introduction to Bacterial Chemotaxis 26
2.1 Bacterial Motility and Taxis . . . 26
2.2 Bacterial Chemotaxis . . . 28
2.3 Chemoreceptor Structure and Signal Transduction . . . 36
2.4 Chemoreceptor Arrays . . . 41
2.4.1 Architecture . . . 41
2.4.2 Positioning and Partitioning . . . 44
2.5 Aims of this Section of the Thesis . . . 46

3 Methods **49**
3.1 Bacterial Strains and Growth Conditions . . . 49
3.1.1 Strains and Plasmids . . . 49
3.1.2 *E. coli* Growth Conditions . . . 51
3.1.3 *R. sphaeroides* Growth Conditions . . . 51
3.1.4 Antibiotics . . . 51
3.1.5 Storage . . . 52
3.2 Genetic Manipulations . . . 52
3.2.1 Cloning . . . 52
3.2.2 DNA Transfer . . . 56
3.3 Protein Expression and Purification . . . 58
3.3.1 Tsr Overexpression in *E. coli* . . . 58
3.3.2 Tagged PpfA Expression in *R. sphaeroides* . . . 59
3.3.3 Membrane Isolation and Protein Purification . . . 60
3.3.4 Confirming Protein Expression . . . 63
3.4 Microscopy and Image Analysis . . . 65
3.4.1 Fluorescence Microscopy . . . 65
3.4.2 Cryo-electron Tomography . . . 66
3.4.3 Cryo-sectioning and Immuno-gold Labeling . . . 68

4 Optimization of *E. coli* Serine Chemoreceptor (Tsr) Expression and Purification **70**
4.1 Overview of Protocol . . . 70
4.2 Induction of Tsr Expression . . . 72
4.3 Membrane Isolation . . . 74
4.3.1 Sucrose Gradient . . . 74
4.3.2 Buffer Composition . . . 75
4.4 Solubilization and Purification of Untagged Tsr . . . 77
4.5 Solubilization and Purification of *Strep*-Tagged Tsr . . . 79
4.5.1 Creation of *Strep*-Tagged Tsr Construct . . . 79
4.5.2 Solubilization and Purification . . . 80
4.6 Discussion . . . 83

5 Structure of *E. coli* Tsr **86**
5.1 Different Sample Preparation Strategies for Cryo-ET . . . 87
5.1.1 Cell Lysis on the Grid . . . 88
5.1.2 Thin *E. coli* cells . . . 89
5.1.3 Isolated Membrane Assemblies . . . 91
5.2 Tsr Structure in the Presence and Absence of Serine . . . 92
5.3 Discussion . . . 96

6 Organization of the *R. sphaeroides* Cytoplasmic Chemosensory Partitioning Protein, PpfA **100**
6.1 Creation of Plasmids and Strains . . . 100
6.2 Fluorescence Microscopy for Phenotypic Validation . . . 101
6.3 Comparison of Organization of Wild-Type and ATP-locked PpfA by Immuno-gold Labeling . . . 103

6.4 Discussion . . . 104

7 Discussion **109**

III Human Cancer Metabolism 111

8 Introduction to PKM2 and its Role in Cancer Metabolism **112**
8.1 Cancer-related Modifications to Cellular Metabolism . . . 112
8.2 The Role of PKM2 in Cancer . . . 114
8.3 PKM2 Structure and Function . . . 116
8.3.1 Allosteric Regulation . . . 117
8.3.2 Post-translational Modifications and Mutations . . . 120
8.4 Aims of this Section of the Thesis . . . 121

9 Methods **123**
9.1 Strains, Plasmids, and Growth Conditions . . . 123
9.2 Protein Expression and Purification . . . 124
9.2.1 Overexpression in *E. coli* . . . 124
9.2.2 Protein Purification . . . 124
9.2.3 Validating Protein Expression and Quality . . . 125
9.3 Microscopy and Image Analysis . . . 126
9.3.1 Negative Staining . . . 126
9.3.2 Preparation of Vitreous Specimens . . . 126
9.3.3 Cryo-electron Microscopy . . . 127
9.3.4 Image Processing . . . 127
9.3.5 Structure Refinement and Analysis . . . 128

10 Biochemical and Structural Characterization of Pyruvate Kinase M2 (PKM2) **130**
10.1 PKM2 Purification and Biochemical Characterization of Oligomeric State . . . 130
10.2 Determination of Structure of Wild-Type and H391Y Mutant Forms of PKM2 . . . 133
10.3 Flexibility of the Domains with Monomers . . . 138
10.4 Discussion . . . 141

IV Concluding Remarks 146

11 Summary of Imaging Methods Utilized in This Study **147**

12 Appendices **149**
A Media and Reagents . . . 149
B Primers . . . 153

Bibliography **155**

List of Figures

1.1 Conformation and hydrogen bonding in α helices and β sheets 4
1.2 Relative scale of structural and imaging techniques used in biology . . 8
1.3 Schematic diagram of a transmission electron microscope 11
1.4 The workflow for cryo-electron microscopy and single particle reconstruction 15
1.5 Method for determining 3-D reconstruction from 2-D projection images 17
1.6 The workflow for cryo-electron tomography 20

2.1 Comparison of random walk and biased random walk 28
2.2 The chemotaxis pathway of *E. coli* 31
2.3 The chemotaxis operons and loci of *R. sphaeroides* 33
2.4 The chemotaxis pathway of *R. sphaeroides* 35
2.5 Structure of chemoreceptor trimer-of-dimers 37
2.6 Proposed mechanism of chemoreceptor signal transduction 39
2.7 Classification of MCPs by cytoplasmic domain heptad repeats 40
2.8 Architecture of membrane-associated and cytoplasmic chemoreceptor arrays 43
2.9 Schematic representations of membrane-associated and cytoplasmic chemoreceptor arrays 44

3.1 Overlap extension PCR 54

4.1 SDS-PAGE analysis of whole cell lysates after Tsr induction 74
4.2 SDS-PAGE analysis of sucrose gradient fractions 75
4.3 SDS-PAGE analysis of membranes in buffers of different ionic strengths 76
4.4 Purification of untagged Tsr 78
4.5 Purification of *Strep*-tagged Tsr in DDM 81
4.6 Detergent screen for solubilizing Tsr 82
4.7 Purification of *Strep*-tagged Tsr in DM 82

5.1 Cell lysis on grids induced by freeze-thaw 88
5.2 SDS-PAGE analysis of Tsr expression in B/r H266 90
5.3 Cryo-ET of Tsr assemblies in B/r H266 cells 91
5.4 Cryo-ET of isolated Tsr membrane assemblies 93
5.5 Tsr structure in the presence and absence of serine 94
5.6 Comparison of the structures from this study with previous study . . 95

6.1 Extent of cluster segregation by tagged variants of PpfA 102

6.2 Immuno-gold labeling of HA-tagged PpfA induced with 0.01mM IPTG 104
6.3 Immuno-gold labeling of wild type and mutant HA-tagged PpfA induced with 0.5mM IPTG 105

8.1 Schematic representation of the Warburg effect 115
8.2 Crystal structure of PKM2 with bound ligands 117
8.3 Isoform-specific region of PKM2 118

10.1 SEC analysis of the effect of FBP and TCEP on PKM2 oligomeric state 132
10.2 Cryo-EM of vitrified specimens of $PKM2^{H391Y}$ 134
10.3 Symmetrized cryo-EM maps of $PKM2^{WT}$ and $PKM2^{H391Y}$ 135
10.4 Comparison of FBP density in wild-type and mutant PKM2. 136
10.5 Comparison of electron density at residue 391 in wild-type and mutant PKM2 137
10.6 Potential stacking interaction formed in the mutant 137
10.7 Local resolution estimation for symmetrized cryo-EM maps of $PKM2^{WT}$ and $PKM2^{H391Y}$ 138
10.8 Resolution estimation for asymmetric cryo-EM maps of $PKM2^{WT}$ and $PKM2^{H391Y}$ 139
10.9 Local resolution estimation for asymmetric cryo-EM maps of $PKM2^{WT}$ and $PKM2^{H391Y}$ 140
10.10 Domain flexibility in asymmetric structures of $PKM2^{WT}$ and $PKM2^{H391Y}$ 141

List of Tables

3.1 Bacterial strains used in this study 49
3.2 Plasmids used in this study 50
3.3 Working concentrations of antibiotics 52
3.4 Composition of PCR reaction mixture 53
3.5 PCR reaction program 53
3.6 Reaction mixture composition for restriction enzyme digests 55
3.7 Reaction mixture composition for DNA ligation 55
3.8 Detergent Concentrations for Tsr Solubilization 61

4.1 Procedures optimized for the production of Tsr samples 71
4.2 Induction conditions tested for HCB721/pHSe5.tsr_{QEQE} 73
4.3 Succession of steps used for construction of pHSe5.*tsr-Strep* 80

9.1 Components in PKM2 cryo-EM samples 126
9.2 Data Collection and Image Processing Statistics 128

Abbreviations

2-D	Two dimensional
3-D	Three dimensional
Å	Ångstrom
Ala	Alanine
Amp	Ampicillin
APEX	Engineered ascorbate peroxidase tag
APXII	Higher activity APEX tag
Arg	Arginine
ATP	Adenosine triphosphate
B2H	Bacterial two-hybrid
bp	Base pair
Carb	Carbenicillin
CCD	Charge-coupled device
CCW	Counter-clockwise
CD	cytoplasmic domain
CFP	Cyan fluorescent protein
CHS	Cholesterol hemisuccinate
CMC	Cristical micelle concentration
Co^{2+}	Cobalt(2+)
Cryo-EM	Cryo-electron microscopy
Cryo-ET	Cryo-electron tomography
CTF	Contrast transfer function
CW	Clockwise
DAPI	4',6-diamidino-2-phenylindole
DDD	Direct electron detection devices
DDM	n-Dodecyl-β-D-maltopyranoside
DIC	Differential interference contrast
DM	n-Decyl-β-D-maltopyranoside
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	deoxynucleoside-5'-triphosphate
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
FBP	Fructose 1,6-bisphosphate
FFT	Fast Fourier transform
FIB	Focused ion beam
GFP	Green fluorescent protein
H	Histidine
HA	Human influenza hemagglutinin
ID	Identification
IPTG	Isopropyl β-D-thiogalactoside
Kan	Kanamycin
kb	Kilo base

kDa	Kilo Dalton
L	Liter
LB	Lysogeny broth
LDS	Lithium dodecyl sulfate
M	Molar
MCP	Methyl-accepting chemotaxis protein
mL	Milliliter
mRNA	Messenger ribonucleic acid
MW	Molecular weight
N/A	Not applicable
Nal	Nalidixic acid
NaCl	Sodium chloride
NCI	U.S. National Cancer Institutes
NIH	U.S. National Institutes of Health
NMR	Nuclear magnetic resonance
OD_{600}	Optical density at 600nm
OD_{700}	Optical density at 700nm
OG	Octyl glucoside
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
PEP	Phosphoenolpyruvate
PFA	Paraformaldehyde
PKM1	Pyruvate kinase M1 isoform
PKM2	Pyruvate kinase M2 isoform
Phe	Phenylalanine
PpfA	*Rhodobacter sphaeroides* chemosensory segregating protein
PPP	Pentose phosphate pathway
ROS	Reactive oxygen species
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
Ser	Serine
TCEP	Tris(2-carboxyethyl)phosphine
TEM	Transmission electron microscope
Tlp	Transducer-like protein
TlpT	Essential *R. sphaeroides* cytoplasmic chemoreceptor
TM	Transmembrane
Tsr	*Escherichia coli* serine chemotaxis receptor
Tyr	Tyrosine
WT	Wild type
Y	Tyrosine
YFP	Yellow fluorescent protein

Part I

General Introduction

Chapter 1

Introduction to Principles and Methodology of Structural Biology

1.1 Protein Structure

Proteins are macromolecules which are responsible for carrying out the majority of processes required for cells to live. Their diverse functions include catalyzing biochemical reactions for the formation of new molecules, sensing and transmitting signals, regulating the exchange of small molecules between a cell and its environment, and providing intracellular structural support [1, 2].

In order for individual proteins to be able to serve their functions, they need to have the correct surface chemistry and three-dimensional structure, or conformation. These attributes are determined by the sequence of amino acids that make up a given protein. Amino acids are linked together in unbranched chains by covalent peptide bonds with their order determined by the sequence of nucleotides in regions of an organism's DNA. Each of the twenty genetically-encoded amino acids has distinct chemical properties based on the composition of a variable region, known as a side chain, that is not involved in forming peptide bonds. These side chains can be classified as nonpolar or polar with some polar residues bearing positive or negative charges at physiological pH [1].

The amino acid sequence, or primary structure of a protein, dictates the folding of a polypeptide chain based on a number of noncovalent interactions that form between different parts of the chain. These include hydrogen bonds, electrostatic attractions, and van der Waals interactions that can involve atoms in the polypeptide backbone, side chains, or both [3, 4]. In addition, hydrophobic clustering forces push amino acids with nonpolar side chains together to minimize their disruptions of the hydrogen bond network of surrounding water molecules [5]. Though these noncovalent interactions are weak, they are collectively strong enough to hold a protein in the correct conformation under physiological conditions [6]. Some proteins, particularly those that are secreted or attached to the outside of a cell's plasma membrane, are also stabilized by covalent cross-linkages that reinforce the favored conformation. The most common type of cross-linkages are covalent sulfur-sulfur bonds, or disulfide bonds, that form between the sulfhydryl groups on the side chains of cysteine amino acids [7].

Though each protein has a distinct conformation, there are some folding patterns that are observed in numerous proteins. The two most common, α helices and β sheets, are a result of hydrogen bonding between the N-H and C=O groups of the polypeptide backbone **(Figure 1.1)**. An α helix occurs when a polypeptide chain twists on itself to form a regularly spaced helix with a hydrogen bond occurring between every fourth peptide bond [8]. α helices composed of nonpolar amino acids are common in the transmembrane regions of membrane proteins since the hydrophilic polypeptide backbone hydrogen bonds with itself and can be shielded from the nonpolar lipid environment by the hydrophobic side chains extending out from the helix [9]. In some proteins, between two and four α helices can twist around each other in such a way that their nonpolar amino acid side chains can face inward and thus, are shielded from the aqueous environment. This very stable structure, known as a coiled-coil, is found in many elongated proteins including the *E. coli* serine chemotaxis receptor, which is studied in this work [10, 11, 12]. A β sheet is formed when hydrogen bonds form between neighboring segments of the polypeptide backbone running either in the same orientation (parallel) or in the opposite orientation (an-

tiparallel). Both types of β sheets form rigid structures which are often found in the cores of many proteins [13].

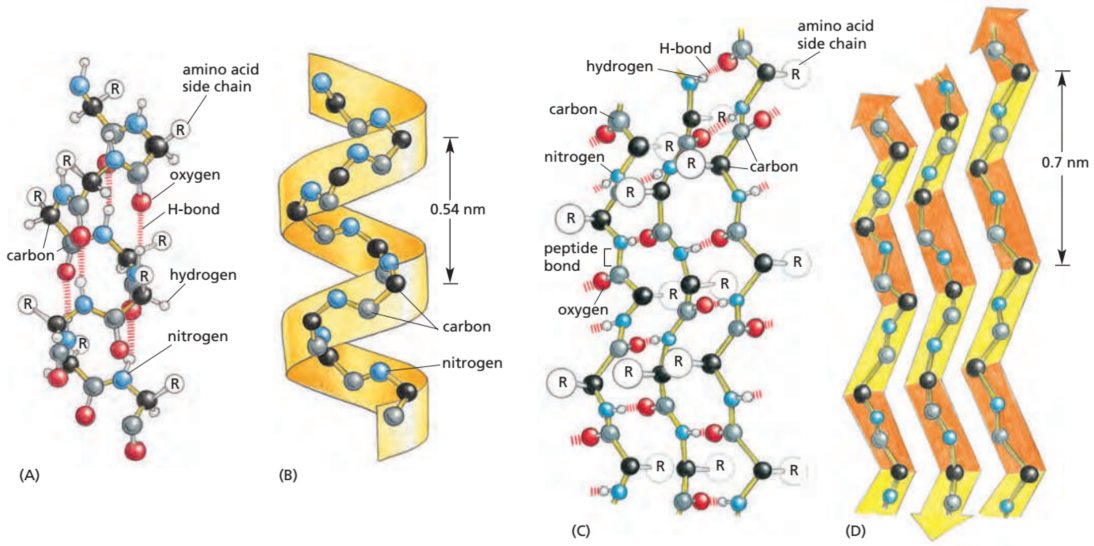


Figure 1.1: Conformation and hydrogen bonding in α helices and β sheets. (A) and (B) show an α helix where each N-H bond of the backbone forms a hydrogen bond with the C=O located four peptide bonds away in the same chain. In this diagram, the N-H groups all point up and the C=O groups all point down (towards the C-terminus), giving the helix a polarity. (C) and (D) show a β sheet that is held together by hydrogen bonds between different adjacent chains. The amino acid side groups in a given chain alternate between extending out above and below the β sheet. This particular example depicts adjacent peptide chains running in opposite (antiparallel) directions. (A) and (C) show all of the atoms in the peptide backbone with the side chain groups denoted by R. (B) and (D) show only the carbon and nitrogen atoms in the peptide backbone. Figure taken from [1].

The regions containing α helices and β sheets make up the secondary structure of a protein. The overall fold of a single polypeptide chain that results from additional noncovalent interactions and in some cases, covalent cross-linkages, is known as a protein's tertiary structure. For many proteins, multiple folded polypeptide chains come together to form the final functional protein. This final level of organization containing multiple chains is referred to as quaternary structure. In addition to the their correctly folded polypeptide chains, some proteins also require the association of a small molecule or the covalent addition of sugars or oligosaccharides in order to function properly [1, 14].

Protein structural and functional studies have the importance of protein domains as organizational units. A domain is a substructure produced by a contiguous portions of a polypeptide chain which can form a stable structure independent of the rest of the protein. Individual domains are often linked to specific functions and form modular units from which larger proteins are constructed. Proteins can consist of a single domain or several domains linked by unstructured, flexible regions of a polypeptide chain. Some domains are found in several different proteins where they perform highly similar functions due to their structural similarities [14, 15].

The final conformation of a polypeptide chain is generally the fold that produces the lowest energy state of the chain [3]. Although polypeptide chains generally can fold into the correct conformation independently, they are often guided along the most energetically favorable folding pathway by molecular chaperones present in living cells [16]. Most proteins fold into a single conformation, which changes slightly during interactions with other molecules. These changes in shape are often essential for the protein to correctly perform its cellular functions [17].

1.2 Functional Implications of Protein Structure

All proteins perform their biological functions by binding to other molecules, known as ligands, with great specificity. Proteins can bind many different types of ligands including ions, small molecules, or macromolecules such as other proteins. The specificity of protein binding results from matches between the surface contours of the protein and its ligand. The binding site is formed by a very specific arrangement of amino acids that are in close proximity in the folded protein [18, 19]. The specific orientations of the amino acids in the binding site can result in interactions that make ligand binding energetically favorable [20]. Even small changes in amino acids even in regions of the protein far from the binding site can alter the protein conformation enough that they disrupt the binding site [21].

For some proteins, such as enzymes, ligand binding is only the first step in performing their functions. Enzymes convert their ligands, or substrates, into one or more products. Enzymes catalyze these reactions by forming, breaking, or changing bonds in a wide variety of substrates by stabilizing high-energy transition states that occur in the course of the reaction. This reduces the energy barrier required for a reaction to proceed, thereby increasing the rate of reaction by thousands or even millions of times [22].

Cells have multiple ways of regulating the activity of the many enzymes that are present in the cytosol. Cells can adjust the expression level, subcellular localization, or protein degradation rate of an enzyme [23]. However, the most common type of regulation involves the reversible binding of a regulatory molecule, which is often another product in the pathway in which a given enzyme is involved. Regulatory molecules can either compete with a substrate for binding in an enzyme's active site or can bind at a completely different regulatory site. Regulators that bind a site distinct from the active site are known as allosteric regulators and often have a very different shape from the enzyme's substrate. The binding of an allosteric regulator causes a small conformational change in the protein such that changes in the positions and orientations of the amino acids at the active site alter its substrate binding capability or catalytic activity [24, 25].

Most proteins, not just enzymes, are thought to be allosteric since their activities can be altered by small conformational changes that accompany the binding of a ligand. Such ligand-induced conformational changes and the resulting modifications to surface contours occur in proteins as diverse as receptors, structural proteins, and motor proteins [26, 27, 28]. A ligand will stabilize the conformation to which it binds most strongly. Thus, in cases where multiple ligands prefer the same conformation one ligand binding can increase the propensity of the of other ligands to bind to the same protein. A specific example of this is cooperativity in proteins that have multiple copies of the same subunit. In this case, ligand binding to a single subunit

causes an allosteric change that facilitates binding of the same ligand to the other subunit. This results in a much sharper cellular response to the presence of the ligand because binding to a single subunit can cause a conformational change that nearly completely actives or inactivates the protein [29, 30].

Another method that is used extensively to regulate protein activity is the covalent addition or removal of small modifier groups, most commonly phosphate groups. For instance, the enzyme-mediated addition of a phosphate group, which carries two negative charges, can cause a conformational change by attracting amino acids with positively charged side chains. This change can be propagated to other regions of the protein's surface in order to alter its activity [31]. The effect of phosphorylation can be reversed by another enzyme-mediated reaction that removes the phosphate group [32]. A phosphate group can also regulate protein activity by either forming a part of the structure that is recognized by another protein's binding site or by masking a binding site that might otherwise hold proteins together. Thus, phosphorylation can also influence the assembly or disassembly of protein complexes, thereby potentially influencing cellular localization [31]. The addition and removal of phosphates often occurs in response to signals that indicate a change in the cell's state. In many cases, a cascade of protein phosphorylation events can relay a signal from the cell's surface in order to bring about a change in cellular behavior, such as gene expression [33, 34].

Protein structure is critical for ensuring proper activity, localization and regulation of proteins, whether it is in the context of ligand binding, allostery, or phosphorylation. Thus, a thorough understanding of the function of a given protein requires a precise knowledge of its structure in all of its different physiological states as well as understanding of its localization in the cellular context. There are a variety of techniques that have been developed in order to study the structure and localization of proteins. They differ in the type and level of information that they can provide about these two aspects of proteins. A summary of the various techniques and their uses is provided in the following section.

1.3 Structural and Imaging Methods in Biology

Biological processes occur over various length scales ranging from molecular interactions on the sub-nanometer level to interactions of organisms and ecosystems over tens of meters. The diverse range of techniques used to study these biological phenomena is equally vast in its scope and methodology. Even the subset of techniques used for visualizing proteins, cells, and their respective structures is varied in both approach and practical resolution range **(Figure 1.2)**.

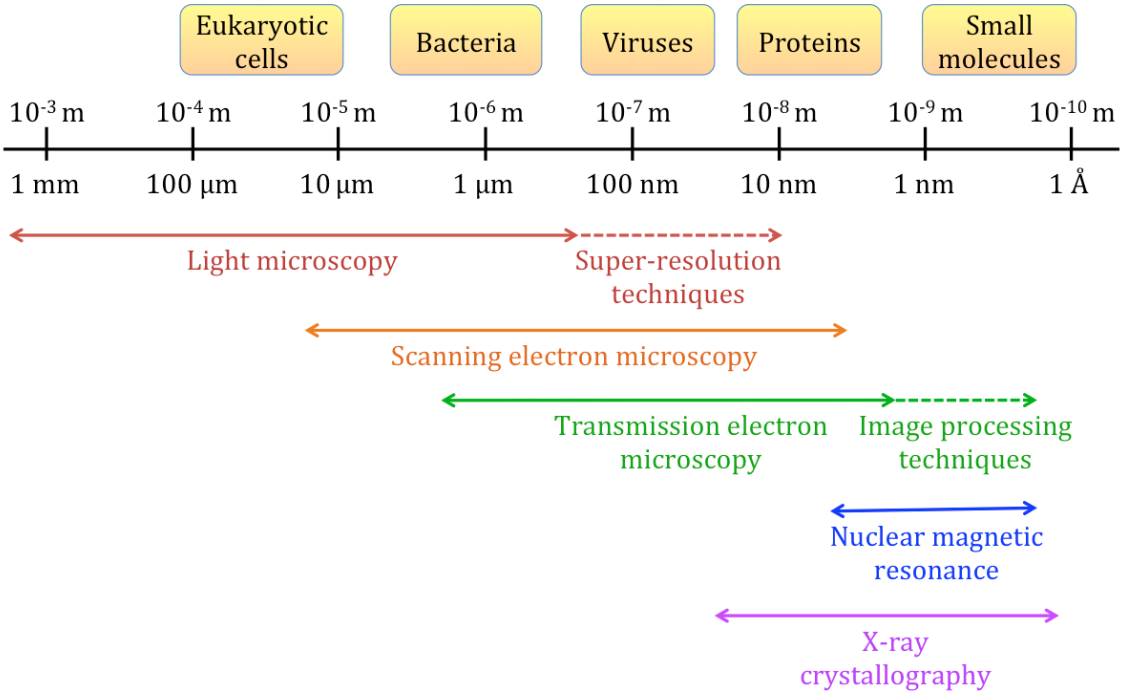


Figure 1.2: Relative scale of structural and imaging techniques used in biology.
Some of the commonly used methods for visualizing cell and protein structure are mapped by their practical limits of scale and resolution. In some cases, such as in the case of transmission electron microscopy, direct visualization of protein features in individual images is limited. However, subsequent image processing such as alignment and averaging to boost signal to noise ratios of the data can yield much higher resolution information about protein structure.

Traditional optical light microscopy can be used to view live mammalian and bacterial cells, but it is limited to a resolution of approximately 250 nm because of

the diffraction limit of light [35]. More detailed information on localization of cellular targets can be obtained using fluorescence microscopy, with targets of interest labelled using genetically encoded tags, photoswitchable tags or quantom dots [36]. Historically, fluorescence microscopy was limited to the same resolutions as traditional light microscopy, but the advent of inventive fluorescence microscopy techniques in recent years has allowed for biological imaging with spatial resolutions in the range of 10-100 nm. Examples of these techniques include photo-activated localized microscopy (PALM), stochastic optical reconstruction microscopy (STORM), structured illumination microscopy (SIM), and stimulated emission depletion (STED) [35, 36].

There is also a suite of electron microscopy techniques that have proven useful for biological applications. Most of them involve transmission electron microscopy, but one notable exception is focused ion beam scanning electron microscopy which now has been used in multiple instances to gain new insights into subcellular architecture and host-pathogen interactions [37, 38, 39]. While this technique has the ability to provide 3-D information on large volumes, including tissues, it is limited by a resolution of approximately 5nm [40]. Transmission electron microscopy, which offers higher resolution, can be performed on various types of samples including purified macromolecular complexes, viruses, and thin sections of cells. In order to reduce radiation damage and withstand the high vacuum in an electron microscope, biological samples must either be studied in a frozen hydrated state (cryo-electron microscopy or cryo-EM) or stained with a heavy metal solution and dried before imaging [41]. While both types of samples can be directly visualized at a resolution of 20-40 Å, frozen hydrated samples offer the opportunity to yield 3-D structural averages with resolutions up to 2 Å after subsequent alignment and averaging [42]. The achievable resolutions vary based on the nature of sample being studied and the type of imaging performed. Many types of samples can be imaged in a frozen hydrated state including bacterial cells, viruses, cell fragments, and purified proteins. In general, sample thickness is inversely correlated with achievable resolution as thicker samples decrease the recorded signal by increasing scattering. Data collection on

frozen hydrated samples can take the form of acquiring single projection images at each area of interest, as is typical in standard cryo-EM, or a series of tilted image at each region of interest, as is done in cryo-electron tomography (cryo-ET). At present, macromolecular structure determination at near-atomic resolution is limited to cryo-EM of purified macromolecular complexes, though cryo-ET has yielded sub-nanometer resolution structures of *in situ* macromolecular complexes [41, 43, 44, 45].

Other techniques of studying macromolecular structure include solution nuclear magnetic resonance (NMR) and X-ray crystallography. Solution NMR allows for protein structure determination up to a resolution of approximately 2 Å without crystallization or freezing [46, 47]. NMR also provides information on protein conformational flexibility through the measurement of inter-atomic distances and dihedral angles in solution. X-ray crystallography is currently the most effective method for determining structures of proteins and nucleic acids at atomic resolution (1.5-3.0 Å). While it is presently the highest resolution approach available in biology, its use is limited to cases where the macromolecular complex is able to form sufficiently ordered crystals and where the sample is present in sufficient quantities to pursue crystallization trials [41].

1.4 Transmission Electron Microscopy

The first transmission electron microscope (TEM) was developed by Ernst Ruska and Max Knoll in the 1930s [48]. Since then, there have been numerous developments to improve resolution when studying biological samples, but the overall architecture remains largely the same. In a TEM, illumination of the sample is achieved using condenser lenses to define the illuminated area and angle of convergence of an electron beam extracted from a cathode kept at high negative voltage. The objective lens is the first image-forming lens encountered by the electrons as they travel downward through the microscope column. As a result of the small focal length of the objective lens, the specimen needs to be immersed in the bore

of the lens. Beams that are transmitted through and scattered by the specimen are focused in the back focal plane of the objective lens. The objective aperture, which is located in this plane, is used to improve contrast of the final image by removing widely scattered electrons. The diameter of the objective aperture determines what portion of the scattered electrons participate in image formation. A series of projector lenses is located between the objective lens and the detection devices. These lenses are responsible for magnifying the image and projecting it onto the screen or camera for viewing or photographic recording, respectively **(Figure 1.3)** [49].

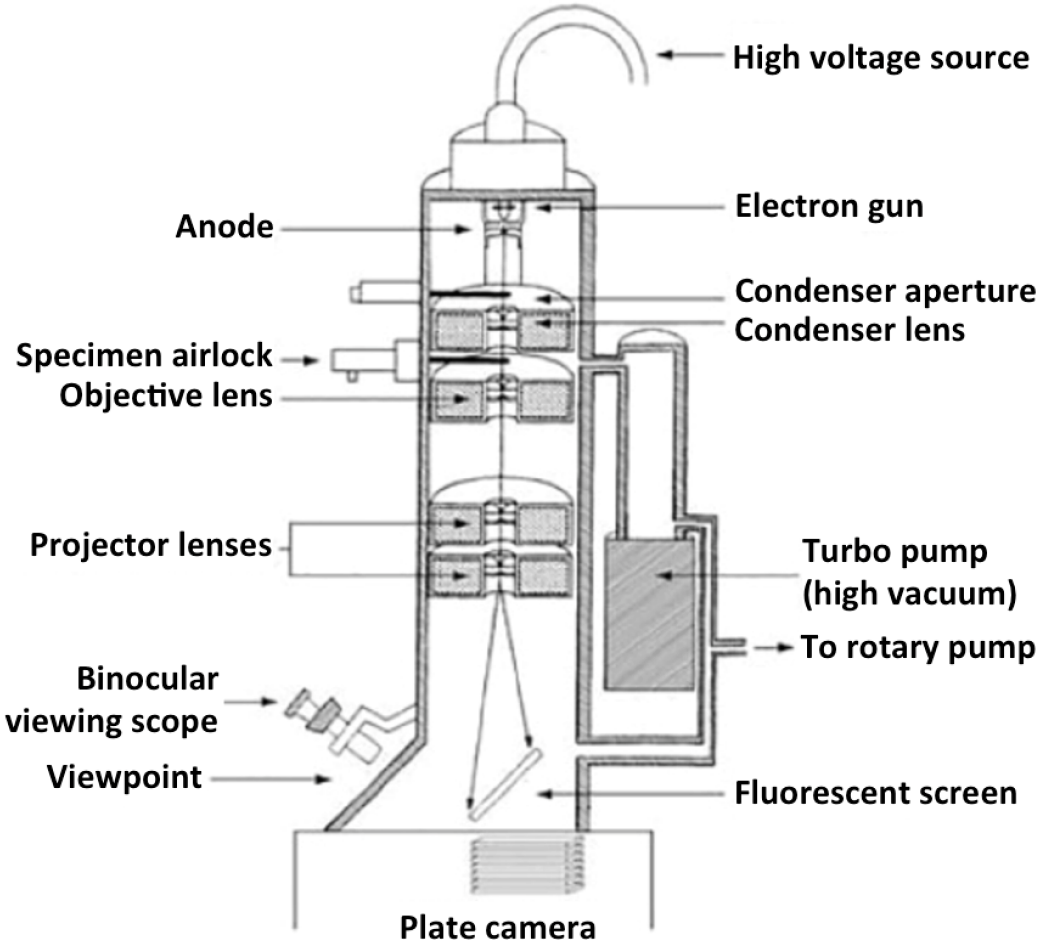


Figure 1.3: Schematic diagram of a transmission electron microscope. Electrons travel down the microscope column from the gun through a series consisting of condensor lenses, the objective lens, the specimen and projectors lenses in order to form a magnified image on the viewing screen or camera. Figure taken from [49].

Amplitude contrast in a TEM is generated by removing widely scattered electrons with the objective lens aperture. Higher mass or thicker areas of the sample will scatter more electrons than their lower mass or thinner counterparts. In order to enhance amplitude contrast in biological samples, they can be stained with a low-concentration solution of heavy metals (usually uranium salts) to generate areas of significantly higher molecular mass [49]. Such staining can be done on thin sections of cells or on purified macromolecular complexes. There are currently two techniques to specifically label subcellular protein targets that largely rely on amplitude contrast from high molecular weight atoms. The first, immuno-gold labeling, involves using gold-conjugated antibodies to detect the proteins of interest, usually in sections of cells that are chemically fixed. The second involves genetically encodable tags that produce singlet oxygen that reacts with diaminobenzidine (DAB) to generate an osmophillic precipitate around the protein of interest. Subsequent staining with osmium tetraoxide generates regions of dark staining that correspond to the localization of the desired target [36].

Image formation in a TEM is complex and involves contributions from phase contrast as well as amplitude contrast. Phase contrast results from the interference between unscattered electrons and elastically scattered electrons. In unstained, frozen hydrated biological samples, phase contrast dominates as there are minimal differences in the molecular mass of the atoms in the specimen [49]. In order to amplify phase contrast, images of cryo-EM samples need to be acquired away from focus, typically in the range of 1 µm to 5 µm under focus. Defocussing enhances phase contrast by increasing phase shifts in elastically scattered electrons relative to their unscattered counterparts, but also introduces artifacts in the image as a result of the imperfect lens system. These artifacts are modeled by the contrast transfer function (CTF) and computational corrections of the CTF are essential for high-resolution structure determination [50]. Another alternative for generating phase contrast that allows for the acquisition of images in focus is the Volta phase plate. This utilizes a beam-induced Volta potential on the surface of a continuous thin film placed in the back focal plane of the objective lens in order to generate a phase shift in the elastically

scattered electrons with respect to their unscattered counterparts [51]. The Volta phase plate has been reported to be comparable to or slightly better than the conventional defocus approach and has yielded near-atomic resolution structures. It is proposed that the phase plate could facilitate cryo-EM structure determination of difficult targets including those that are small, heterogeneous or flexible [52].

The two major data collection schemes on a TEM involve taking single 2-D projection images or taking a series of tilted images over regular increments for reconstruction into a 3-D volume (tomography). While both can be performed on either stained, dried or unstained, frozen hydrated samples, the granularity of the stain and the loss of fine biological features during sample preparation limits the resolution to 15-20 Å resolution for stained samples [41]. Higher resolution studies utilize cryo-EM or the related cryo-electron tomography (cryo-ET) combined with iterative alignment, classification, and averaging. A distinguishing feature of both of these techniques is that the sample is preserved in a near-native state as a result of embedding in amorphous ice [53]. While cryo-EM generally involves studying macromolecular complexes as purified entities, cryo-ET allows them to be studied in their native environments as long as the sample is sufficiently thin [54].

1.4.1 Cryo-Electron Microscopy and Single Particle Analysis

Cryo-EM is rapidly gaining popularity as a method for using TEM to determine the structure of macromolecules preserved in an unstained, frozen hydrated state. The appeal of cryo-EM lies largely in its ability to provide structural information without the need for crystallizing the target macromolecule. Furthermore, it requires only moderate amounts of the sample at concentrations of 1-2 mg/mL. These attributes make cryo-EM structure determination a tractable method for studying large or flexible targets that resist crystallization or macromolecular entities that are difficult to purify in large quantities [55]. For flexible proteins or complexes, cryo-EM offers the advantage that it can simultaneously determine multiple structural states from a single specimen, thereby allowing it to serve as a tool to study protein dynamics [41].

During the last few years, the cryo-EM field has been transformed by the introduction of direct electron detection devices (DDDs). Their high signal-to-noise ratios with respect to charge-coupled device (CCD) cameras was beneficial for unstained, frozen hydrated samples that have low contrast. Another key feature of these detectors was their ability to split the total electron dose used to acquire an image (typically 20-40 e/Å^2 over several seconds) into short frames (typically with a dose of 1 e/Å^2 each). Since biological specimens undergo beam-induced motion as a result of cumulative radiation damage, individual lower-dose frames have minimal beam-induced movement relative to the total exposure. The resulting series of frames can be computationally aligned after data collection to reduce the blurring of features due to beam-induced movement between the frames [56, 57]. This produces particle images with higher signal-to-noise ratios which can be aligned more accurately. Consequently, fewer particles are required in the average because they are less noisy and particles contribute to higher resolutions because of more accurate alignments. As the process iterates, these factors have a synergistic effect that has led to near-atomic resolution structures in several instances [42, 58, 59].

Cryo-EM structure determination is most commonly performed using the single particle reconstruction technique **(Figure 1.4)**. This approach entails imaging a purified protein or macromolecular complex in a thin layer of amorphous ice supported by a holey carbon film on an EM grid. The sample should contain the target in multiple orientations, ideally with a uniform angular distribution, within the layer of vitreous ice. While it is desirable to have a high particle density to reduce the number of images that need to be collected, ideally the particles should not touch each other [50].

Single particle data collection involves taking a single 2-D projection image at each area of interest. For high resolution structure determination, these images are collected using DDDs with many intermediate frames as discussed previously. Images must be taken in the electron microscope's bright-field mode while applying under-

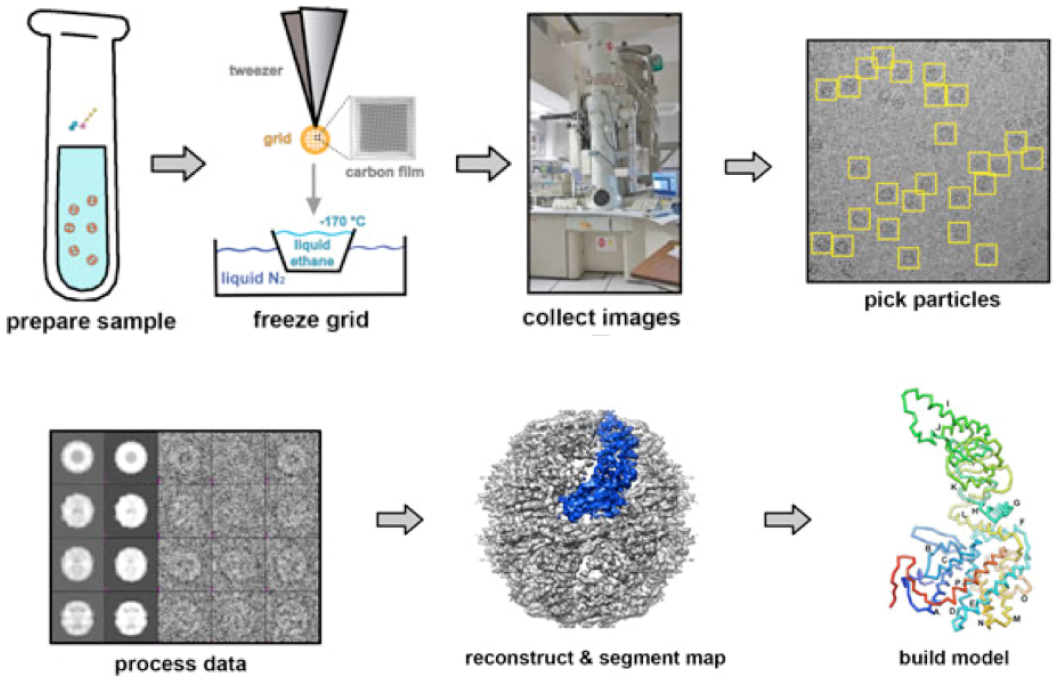


Figure 1.4: The workflow for cryo-electron microscopy and single particle reconstruction.
A solution containing the purified macromolecular complex is placed on an EM grid coated by a thin carbon film. Excess liquid is blotted away and the grid with a thin layer of sample solution is rapidly frozen in liquid nitrogen cooled ethane. While being maintained at liquid nitrogen temperatures, the sample is loaded into a cryo-transmission electron microscope where many projection images are collected. Particles are extracted from the images and iteratively aligned, classified, and averaged. Orientations are assigned to the 2D projections to obtain a 3D reconstruction of the macromolecular complex, which can be segmented into its constitutive parts if desired. If the resolution of the electron density map is sufficiently high, it can be use to build an atomic model. Figure taken from [60].

focus in order to provide sufficient contrast [49]. The images are modulated by the contrast transfer function (CTF), a quasi-periodic sine function in reciprocal space. The high resolution details in an image are weakened by the envelope function of the CTF, with a quicker fall off associated with higher defocus values. Thus, images are generally taken at the lowest defocus values where low-frequency contrast makes the particles visible. Since the CTF has multiple zero crossings (frequencies of destructive electron interference) that correspond to lost information from an image, it is also important that the images are taken at different underfocus settings (resulting in images with different periodicity of the CTF) in order to sample the

entire reciprocal space [61, 62]. During image processing, the CTF for each image is estimated based on the amplitude oscillations observed in the power spectrum of the image [63]. For full CTF correction, particles are multiplied in reciprocal-space by the estimated CTF of image from which they were extracted [61].

While CTF estimation is performed on entire images, subsequent processing steps are done using windowed particles extracted from the images. Particle coordinates can be specified manually or automatically by searching for regions that emerge above background level noise or that match specified templates [50]. The windowed particles suffer from a low signal to noise ratios, which can be improved by alignment and averaging. Alignment is performed by using cross-correlation to maximize the similarity between particle images. During 2-D analysis, this is generally also accompanied by classification in order to remove invalid particles and verify that there is an appropriate distribution of particle views. Classification can be done using a method such as principal component analysis to reduce dimensionality and a clustering method such as K-means to separate subsets of similar images [49]. In each step, the data is presented with K seed templates and the images are aligned and compared to all of the templates and then assigned to the template that provides the closest match. This is done iteratively, where the templates for one round of classification are the averages of the classes from the previous round of grouping. With each successive iteration, the averages improve because images are more precisely aligned to the references, which in turn improves the accuracy of the next round of classification and alignment [64, 65].

The collected images and the resulting windowed particles are 2-D projections of the 3-D macromolecules in the specimen. The ability to combine 2-D projections from different directions into a 3-D reconstruction rests on the central slice theory, which states that the Fourier transform of a 2-D projection is a central slice through the 3-D Fourier transform of the object **(Figure 1.5A)**. Therefore, if the directions of multiple 2-D projections is known, they can be positioned within the 3-D Fourier transform and then used to calculate the original 3-D structure by computing the

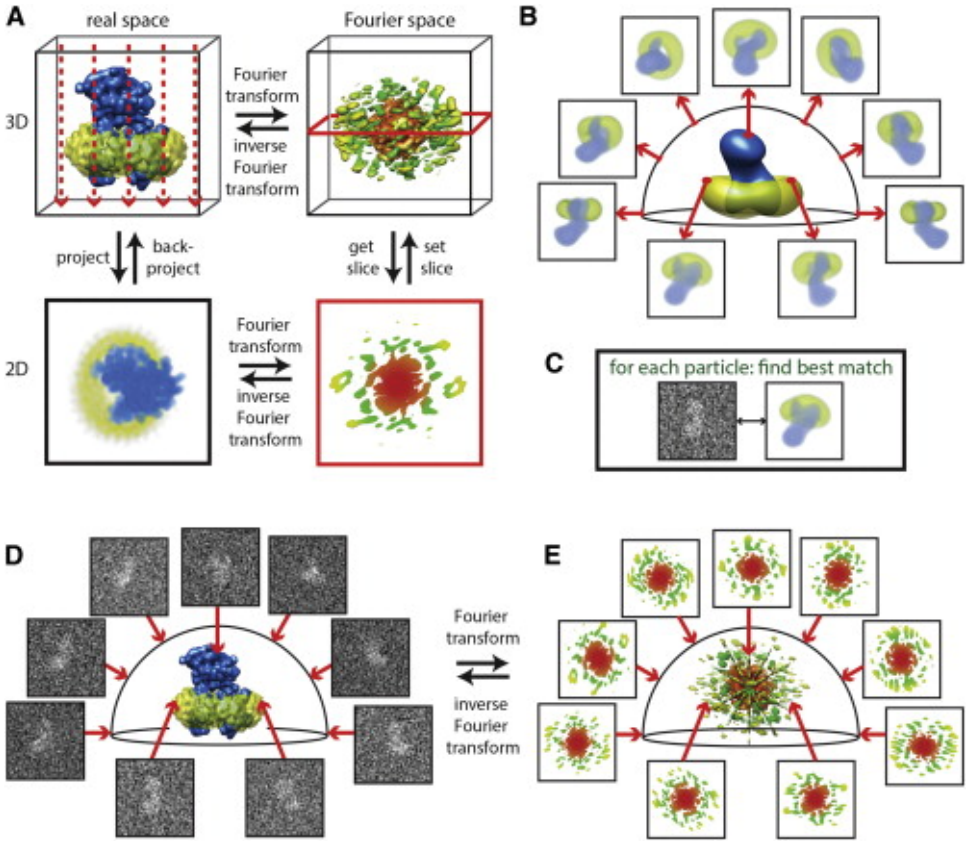


Figure 1.5: Method for determining 3-D reconstruction from 2-D projection images.
A) The central slice theory states that the Fourier transform of a real space 2-D projection of a 3-D object (left) is equivalent to a central slice in the 3-D Fourier transform of the object (right). The direction of projection of the 3-D object in real space is perpendicular to the central slice of the object in Fourier space. B-E) Projection matching is used to combine 2-D projections into a 3-D reconstruction. One compares each experimental image (C) to a set of projections of an initial 3-D reference (B). This orients the projections in real space relative to the particle (D) and allows them to be positioned in the Fourier space representation of the particle (E). An improved model of the particle can be obtained from the inverse Fourier transform and used to repeat steps B-E iteratively until the structure converges. Figure taken from [41].

inverse Fourier transform [41]. In order to determine the relative orientations of the particle in the images, they are generally compared to computationally generated projections of a 3-D reference structure **(Figure 1.5B,C)** [66]. With a good angular distribution of projection views, the reconstruction obtained from the particle images using the assigned orientations will be better than the initial reference. Therefore, applying this method iteratively should converge to the structure of the

particle of interest [41]. Most 3-D refinements are done using this sort of projection matching methodology where over-fitting (artificial creation of features from aligned noise) is prevented by independently refining two arbitrarily divided halves of the data and combining them at the end [50].

Initial 3-D reference structures can either come from similar existing structures, such as low resolution representations of homologs, or be determined *ab initio* from EM images. A variety of approaches exist to calculate 3-D reconstructions from tilting the specimen in a well-defined manner with respect to the electron beam [67, 68, 69]. Alternatively, one can utilize the fact that each pair of 2-D projections have a line of intersection in 3-D Fourier space. The resulting common line in the 3-D Fourier transform of a pair of projections can be used to determine relative orientations by maximizing the similarity between these common lines [70].

Cryo-EM 3-D reconstruction can often be complicated by the fact that many macromolecular samples are heterogeneous as a result of adopting different conformational or compositional states. Maximum-likelihood approaches have been implemented in order to separate different structural states in 3-D without preexisting knowledge of the variations [71, 72, 73]. These methods have made it possible to simultaneously determine multiple conformationally distinct structures from a given sample, thereby providing a unique opportunity to gain insights into the dynamics of macromolecules. In addition to separating conformationally homogeneous subsets of particles, 3-D classification can assist in reaching high-resolution by removing particle images that negatively contribute to the reconstruction process [41].

The interpretation of a cryo-EM structure depends primarily on the resolution of the map, which determines the visible features. There are often local variations in resolution, where lower resolution areas correspond to more flexible regions. For high-resolution maps ($<$4 Å), secondary structure and certain residues are clearly visible allowing for *de novo* backbone tracing. For maps in the range of 4-10 Å, secondary structure features and relative arrangement of subunits can be resolved

which permits unique fitting of segments solved by X-ray crystallography and the detection of conformational changes. Low resolution maps (>10 Å) reveal the overall shape and may be used to determine large quaternary structure changes. They can also be used as 3-D initial models for subsequent higher resolution studies [50].

1.4.2 Cryo-Electron Tomography and Sub-volume Averaging

Cryo-electron tomography (cryo-ET) is a technique closely related to cryo-EM that can be used for *in situ* macromolecular structural studies with a resolution range of about 8-30 Å **(Figure 1.6)** [43, 44, 45]. As is the case in cryo-EM, samples for cryo-ET are studied in an unstained, frozen hydrated state. However, cryo-ET differs in that each area of interest is imaged multiple times while the specimen is tilted through a range of angles [54, 74]. The resulting "tilt series" contains projection images of the specimen, which are limited to the angular range of approximately $\pm 70^{\circ}$ as a result of geometric constraints of the specimen in the microscope [75]. The image sampling is further constrained by the sensitivity of biological specimens to radiation damage from the electron beam as the total dose on a given field of view ($<100 e/Å^2$ and possibly lower to preserve high resolution features) must be split across all of the images in a tilt series [74]. Current technological limitations in the speed of tilting the sample and acquiring images do not allow images to be acquired continuously while remaining below the limit of the total dose that the sample can withstand. Thus, images are collected at discrete angles about the tilt axis, usually in 2-3° intervals. In order to improve the alignment of sequential images in cryo-ET tilt series, 5 nm or 10 nm gold fiducial markers are often used to provide high contrast objects for tracking [76]. The minimum number of tilted views required to determine the structure of an object with diameter D at a resolution of d is given by the Crowther criterion. Based on this criterion, the minimum number of tilted images required is equivalent to π multiplied by D/d, indicating that both smaller objects and higher resolution necessitate a greater number of views [77, 78]. The resolution of structures determined by cryo-ET can also be improved by acquiring each tilted image with intermediate frames in much the same way as the single

particle cryo-EM images described in **Section 1.4.1**. This allows for correction of sample drift that can be more severe at higher tilt angles and in portions of the tilt series with higher cumulative radiation damage [78].

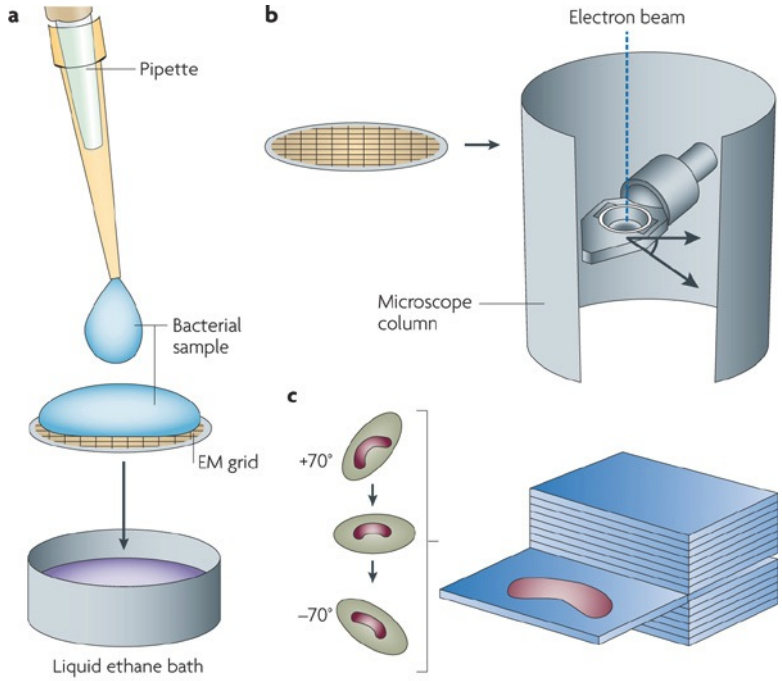


Figure 1.6: The workflow for cryo-electron tomography.
a) A solution containing the cells or assemblies of interest is placed on a holey carbon grid, which is blotted and plunge frozen in liquid nitrogen cooled ethane to create a thin layer of vitreous ice. b) The frozen specimen is loaded into the column of an electron microscope where images are collected over a range of tilt angles. c) The series of tilted images is reconstructed into a 3-D volume, known as a tomogram, which reveals the density profile of the imaged cells or assemblies. Figure taken from[79].

For experiments where the goal is to reach resolutions beyond the first zero crossing of the CTF curve, CTF needs to be estimated and corrected in each image of the tilt series. Though the CTF includes contributions from phases and amplitudes, the effects of CTF are generally only corrected for the phases in the case of cryo-ET. This is largely due to the facts that the amplitude contribution is relatively small and that the low level of signal in the image (as a result of the total dose being split across the tilt series) make it difficult to very accurately estimate the CTF. Tilting the sample results in differences in height across the specimen. As a result, the

defocus also varies and CTF estimation and correction need to be done on adjacent strips of the image parallel to the tilt axis. A common implementation involves utilizing overlapping strips and using interpolation of CTF values in adjacent strips to eliminate discontinuities [80].

Once CTF has been estimated, the tilt series is reconstructed into a 3-D volume known as a tomogram. The ability to perform this reconstruction is based on the central slice theory, which describes the relationship between a 2-D projection of a 3-D object in both Fourier and real space **(Figure 1.5A)** [76]. The main difference between the implementations of 3-D reconstruction in cryo-EM single particle analysis and cryo-ET is that the relative orientations of the 2-D projection images are known in cryo-ET, which removes the need for iterative refinement of the angular assignment relative to a 3-D reference. Cryo-ET reconstruction also involves using the entire field of view of projection images to construct a volume of the entire area of interest on which a tilt series was collected. This is in contrast to cryo-EM reconstructions which are done on particle images extracted from the full field of view as individual particles are assumed to adopt orientations independently of one another [76, 50].

The 3-D tomogram of the field of view imaged contains several sub-regions, or sub-tomograms, containing the macromolecule of interest. These sub-tomograms can be extracted using either manually or automatically defined coordinates. As is the case in single particle analysis, these individual sub-tomograms suffer from a low ratio of signal to noise, which can be overcome by iterative alignment and averaging [81]. Classification is generally also done during this process to separate different conformational states or remove particles with low information content. The methods used for iterative alignment and classification (maximizing similarity by cross-correlation, reducing dimensionality by principal component analysis, and separating image subsets by a clustering method like K-means) are generally the same as those used in cryo-EM single particle analysis, except the alignment must be done in 3-dimensional space [74, 82, 83].

Since cryo-ET tilt series have an angular range limited by geometric constraints of the specimen, they have a wedge of 3-D information that is missing in reciprocal space. This ultimately introduces distortions along the electron beam axis into all real space tomogram voxels of a reconstructed tomogram [75]. In order to account for the missing wedge and prevent its artifacts from biasing sub-tomogram alignment, the similarity measurements should be limited to regions in reciprocal space where there is signal for both objects being compared. It is most straightforward to apply such restrictions in Fourier space where the missing information is naturally separated [82].

Once the alignments of the sub-volumes have converged, a final map is created and interpreted. In cases where the particles were conformationally or compositionally heterogeneous and classification was employed, a map will be created for each class of sub-tomograms. Though cryo-ET has not yet yielded structures with resolutions higher than 8 Å, it is one of the only structural techniques that allows proteins to be studied *in situ*. This is especially relevant for membrane proteins and other complexes that become unstable outside of their native environments or for certain biological questions that require the cellular or subcellular context [74, 75]. At best, cryo-ET structures can yield information on secondary structure elements and conformational changes. However, lower resolution structures can still be valuable if they provide information on quaternary organization or the connectivity of fragments or domains that have been solved by X-ray crystallography in isolation.

1.5 Choice of Methods for the Questions Addressed in this Thesis

This thesis applies a range of imaging techniques to structurally characterize key functional aspects of three proteins involved in cell signaling and metabolic regulation. The areas of study were the chemotaxis systems of *Escherichia coli* and

Rhodobacter sphaeroides as well as the human cancer metabolic pathway. In each system, a key protein was selected for investigation using one or more imaging methods based on the rationale explained below.

1.5.1 Bacterial Chemotaxis

The investigation of bacterial chemotaxis is specifically focused on the *E. coli* serine chemotaxis receptor, Tsr, and the *R. sphaeroides* cytoplasmic chemosensory cluster segregating protein, PpfA. The Tsr studies involved analysis of the conformational states of the full length receptor with and without the ligand, serine. Biochemical efforts were made to detergent-solubilize the receptor for cryo-EM and single particle analysis. However, the detergent-solubilized receptor proved to be unable to withstand the blotting and freezing required for cryo-EM sample preparation. Efforts were made to reduce sample thickness in order to study Tsr in its native membrane environment using cryo-ET. Ultimately, the structure of the receptor was studied in isolated membrane assemblies using cryo-ET and subsequent sub-volume averaging. The question investigated for PpfA involved differences in cellular organization between the wild-type protein and a mutant locked in the ATP-bound state. PpfA was tagged for electron microcopy and light microscopy was used to validate that the tagged protein continued to serve its cellular function of segregating the cytoplasmic chemosensory cluster. In order to localize PpfA within the cellular context, stained cell sections with immuno-gold labeling were visualized with TEM. This provided a platform to selectively detect the tag on PpfA while providing contrast for other cellular features.

1.5.2 Human Cancer Metabolism

The metabolic studies examine the structure of the human M2 isoform of pyruvate kinase (PKM2) that is expressed in cancer cells and is responsible for promoting tumorigenesis, growth, and proliferation [84]. In order to determine the structural

implications of a cancer-correlated patient-derived mutation, we compared this protein variant to the wild-type. Since PKM2 is soluble in aqueous buffers and high-resolution structural differences needed to be resolved at near-atomic resolution, cryo-EM and single particle analysis were used.

Part II

Bacterial Chemotaxis

Chapter 2

Introduction to Bacterial Chemotaxis

2.1 Bacterial Motility and Taxis

Outside of controlled environments (such as those used in the laboratory), bacteria, like many other cells, must sense and respond to their environment and its dynamic changes in order to survive. Bacterial responses can include changes in gene expression or cellular behaviors in order to optimize the use of the available resources. One example of adaptive behavior is taxis, or moving towards favorable environments and away from unfavorable ones.

In order to perform taxis, bacteria have evolved a wide variety of mechanisms for moving across solid surfaces and swimming in liquid media [85]. Examples of the diverse ways in which bacteria move across solid media include: attachment and flow of outer membrane proteins in *Flavobacterium johnsoniae* [86]; extension and retraction of Type IV pili in *Myxococcus xanthus* [87]; attachment of focal adhesion complexes to a surface in order propel the *Myxococcus xanthus* cell body forward [88]; and coordinated rotation of semi-helical filaments extending from the cell body (flagella) [89].

Flagella-driven motility is used by bacteria to swim in liquid environments [90]. There are variations in the architecture, positioning and numbers of flagella across different bacterial species. Monotrichous bacteria, such as *Caulobacter crescentus*

and *Rhodobacter sphaeroides*, possess a single flagellum but vary in the way that they position it [91, 92]. *C. crescentus* has its flagellum located at a specific cell pole, while *R. sphaeroides* randomly positions its single flagellum. Bacteria that produce multiple flagella, such as *Escherichia coli* and *Salmonella enteric* ser Typhimurium, are known as peritrichous bacteria. Other bacterial species that also produce multiple flagella but position them with greater specificity are known as lophotricous. One example is *Vibrio fischeri* which positions the multiple flagella that it produces at one of the cell poles [93]. An interesting variation on the use of flagella for locomotion is demonstrated by spirochetes, such as *Treponema pallidum*, whose flagella are contained within the periplasmic space. These cells are propelled through viscous environments by rotation of the flagella, which results in rotation of the flexible spiral shaped cell body [94].

The semi-helical flagellar filament is attached to the flagellar motor by a flexible hook. In most species, the flagellar motor is able to rotate in both directions in order to alternate periods of forward motion (runs) with periods where the cell body reorients (tumbles). In peritrichous bacteria, counterclockwise motion of the flagellar motor causes the flagella to come together and form a bundle that drives the cell body forward. When the motor switches to clockwise motion, the flagella separate causing the cell to actively tumble. Lophotrichous and monotrichous bacteria usually move forward by counterclockwise motion and reorient by Brownian motion when the motor briefly reverses directions [95]. *R. sphaeroides* is unusual in that its flagellar motor is a unidirectional stop-start motor where forward motion of the cell body occurs during counterclockwise rotation and the cell is reoriented by Brownian motion when the motor stops [96, 97].

Motile flagellate bacteria are too small to sense a spatial gradient. Therefore, they lack the ability to identify environmental differences while they are stationary and must employ temporal sensing mechanisms as they move. Bacteria utilize a 'random walk' of runs in random directions for arbitrary lengths of time in order to move in space **(Figure 2.1A)**. However, cells can bias the frequency of their tumbles,

reverses, or stops in order to embark on a 'biased random walk' where their net movement is towards a favorable environment **(Figure 2.1B)** [98]. This biasing of movement, or taxis, depends on temporal sensing of changes in a range of stimuli such as magnetic fields (magnetotaxis) [99], oxygen (aerotaxis) [100], light (phototaxis) [101], temperature (thermotaxis) [102], nutrient concentrations (chemotaxis) [103], and internal energetic state of the cell (energy taxis) [104].

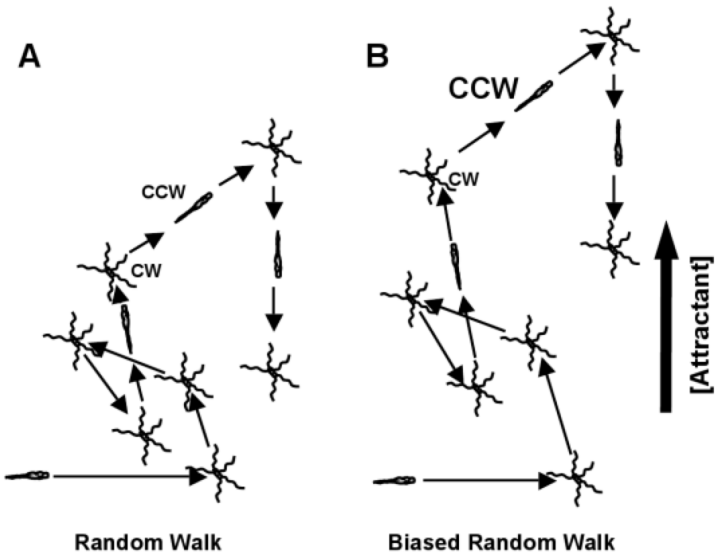


Figure 2.1: Comparison of random walk and biased random walk.
A) In homogeneous environments, swimming bacteria randomly change directions at stochastically distributed intervals, leading to no net movement. B) In non-homogeneous environments, the chemotaxis pathway suppresses direction switching when the bacterium is swimming towards an attractant, thereby causing a net movement in that direction. Figure taken from reference [26].

2.2 Bacterial Chemotaxis

Chemotaxis, or the directed movement in response to environmental chemical stimuli, is essential for a number of bacterial behaviors such a biofilm formation [105, 106], interspecies symbiosis [107, 108], and pathogenicity [109, 110, 111]. In most cases, motility alone without the accompanying chemical sensing and signaling is insufficient for these phenomena to occur. Thus, an understanding of the chemo-

taxis pathway and how it can be controlled thus has implications for processes in a variety of sectors ranging from agriculture to healthcare.

Bacteria use two-component signaling to respond to changes in nutrient concentration as they move through their environment. Unlike one-component signaling where the same protein both senses the stimulus change and effects a cellular response, two-component signaling utilizes different proteins for sensing and responding. Sensing of the stimulus is performed by a histidine kinase enzyme, which in turn activates its cognate response regulator either by phosphorylation or dephosphorylation [112]. Though bacteria use two-component systems to respond to a variety of stimuli including chemical gradients, oxygen levels, pH,osmolarity and light intensity, two-component pathways operate without significant crosstalk despite having largely homologous components [113].

Bacteria use a modified two-component signaling pathway in order to control their flagella in response to chemical stimuli. The best-studied chemotactic pathway is that of *E. coli* [114]. Like other two-component pathways, the *E. coli* chemotaxis pathway uses a histidine kinase (CheA) and a response regulator (CheY). However, CheA differs from the canonical histidine kinases in that it doesn't directly sense stimuli. Instead, it interacts with transmembrane chemoreceptors known as methyl accepting chemotaxis proteins (MCPs) [113]. The adaptor protein, CheW, assists in modulating the interaction between the MCPs and CheA.

MCPs form homodimers, which then organize as trimers of dimers in order to control CheA activity [115]. *E. coli* as four MCPs (Tsr, Tar, Trg, and Tap) as well as a redox sensor (Aer) that sense a variety of chemical stimuli, including some sugars, amino acids, dipeptides and redox potential [116, 117]. Certain MCPs can also sense and respond to more than one stimulus. For instance, the MCP Tar employs different mechanisms to sense both aspartate and maltose. It detects aspartate through direct binding while it senses maltose through interactions with the maltose-binding-protein [118].

Ligand binding to the periplasmic domain of a chemoreceptor causes a shift and rotation in one of the transmembrane helices that is propagated to the distal tip of the MCP where there is a change in interaction with CheA [119]. This results in a change in the rate of CheA autophosphorylation such that attractant binding inhibits CheA autophosphorylation while repellant binding stimulates it [120]. Phosphorylated CheA (CheA-P) can transfer the phosphoryl group to the response regulator, CheY, which in turn can bind to the flagellar motor protein, FliM. Binding of sufficient levels of phosphorylated CheY (CheY-P) cause the cell to tumble by making the motor switch from CCW to CW rotation [121, 122].

Most ligands that bind to the MCPs are chemoattractants and therefore inhibit CheA autophosphorylation. This in turn reduces the levels of CheY-P and causes the cell to continue swimming in the given direction. Therefore, increasing levels of attractant prolong runs, while decreasing levels cause more frequent tumbles. The net effect is that movement is biased towards regions of higher attractant concentration **(Figure 2.1B)**.

Rapid suppression of tumbles is facilitated by CheZ which dephosphorylates CheY-P to terminate the signal [123]. CheZ reduces the half time of CheY-P from ∼2 seconds to ∼200 milliseconds, which allows signal termination to be fast enough for the bacteria to sense and respond to changes while moving [113]. Signal termination in mutants lacking CheZ is so slow that they tumble excessively as a result of persistent signal.

A key feature of the chemotaxis system is a process known as adaptation, which resets the system to an intermediate activity level thereby allowing the cell to respond either negatively or positively to future chemoeffector concentration changes. The adaptation is mediated by methylation and demethylation of specific glutamate residues on the MCPs by the methyltransferase CheR and the methylesterase CheB, respectively [124, 125]. Methylation by the constitutively active CheR increases the

activity of CheA, thereby countering the effects of ligand binding. Negative feedback is also provided by CheB which is only active as a methylesterase when it is phosphorylated by CheA-P. MCPs demethylated by phosphorylated CheB have a reduced ability to induce CheA autophosphorylation, even in the presence of low attractant concentrations, so the rate of CheA activity and thus the rate of direction-switching returns to pre-stimulus levels. In addition, adaptation is facilitated by the substrate specificity of the adaptation enzymes in that CheR preferentially methylates inactive receptors and CheB preferentially demethylates active receptors [126, 127, 120].

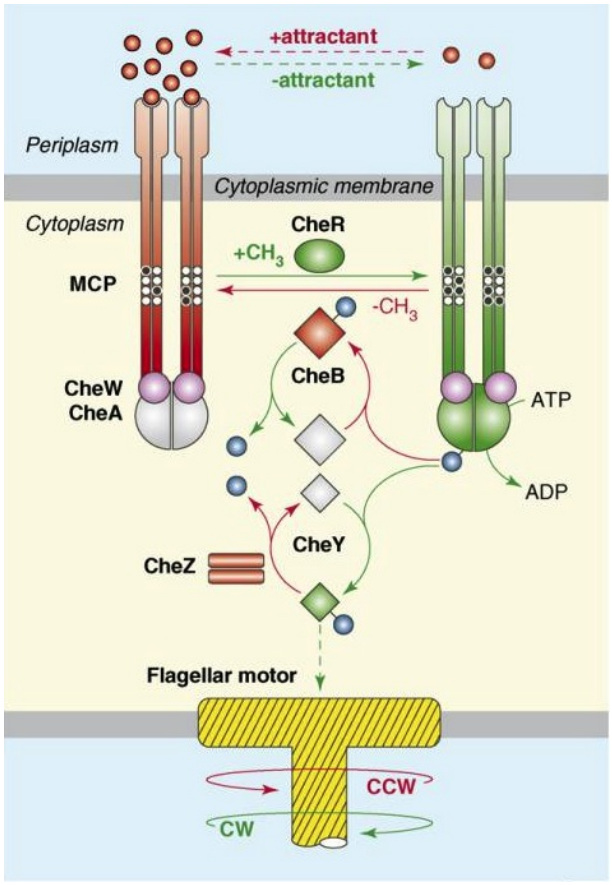


Figure 2.2: The chemotaxis pathway of *E. coli*.
Components shown in red lead to counterclockwise (CCW) movement of the flagellar motor while those in green lead to clockwise (CW) rotation. Gray components represent inactive forms. Phosphoryl groups are shown as blue circles. Receptor modification sites are shown as black (methylated) or white (unmethylated) circles. Figure taken from reference [128].

As a result of its simplicity, the *E. coli* chemotaxis system became an attractive

paradigm for bacterial chemosensory pathways. However, more than 50% of motile bacterial species contain multiple homologues of core chemotaxis proteins (CheA, CheW, CheY, CheB, and CheR) [129]. Therefore, the *E. coli* chemotaxis pathway may not reflect the complexity of chemotactic signaling in other bacterial species. An example of another species that contains a more complex chemosensory pathway with multiple homologues of core proteins is *R. sphaeroides*. This species was selected as an experimental system because it has two distinct chemosensory clusters that localize to different parts of the cell [113]. One cluster is membrane associated and usually found at a cell pole, while the other is located at approximately mid-cell as is believed to be associated with the nucleoid [130, 131]. The presence of two spatially distinct clusters not only begs the question of how signals from the two clusters are integrated into a single output, chemotaxis, but also how the different homologues that are present in each cluster are reliably trafficked, positioned and partitioned to maintain the two clusters.

R. sphaeroides has multiple sets of homologous chemotaxis genes encoded in three operons (*che*Op_1, *che*Op_2, *che*Op_3) and several loci (*che*BRA and *che*Y_4-*mcp*G) on its two chromosomes **(Figure 2.3)** [132]. Of these, *che*Op_1 and *che*BRA are not expressed or required for chemotaxis under normal laboratory conditions [92]. The *R. sphaeroides* chemotaxis system includes thirteen chemoreceptors, four CheA homologues, four CheW homologues, six CheY homologues, two CheB homologues, and three CheR homologues. It does not include a CheZ homologue. Interestingly, the *R. sphaeroides* genome also encodes two different sets of flagellar motor genes *fla*1 and *fla*2, of which only *fla*1 (which produced a single flagellum) is expressed under laboratory conditions. However, selecting for motile δ*fla*1 mutant cells has allowed the isolation of cells expressing *fla*2 genes [133]. Further analysis of these cells revealed that the expression of *fla*2 produced a polar tuft of flagella controlled by the proteins encoded in *che*Op_1 [134].

The proteins encoded by *che*Op_2 and *che*Op_3 follow a modified version of the *E. coli* signaling pattern in their control of the *fla*1 flagellum **(Figure 2.4)**. The use

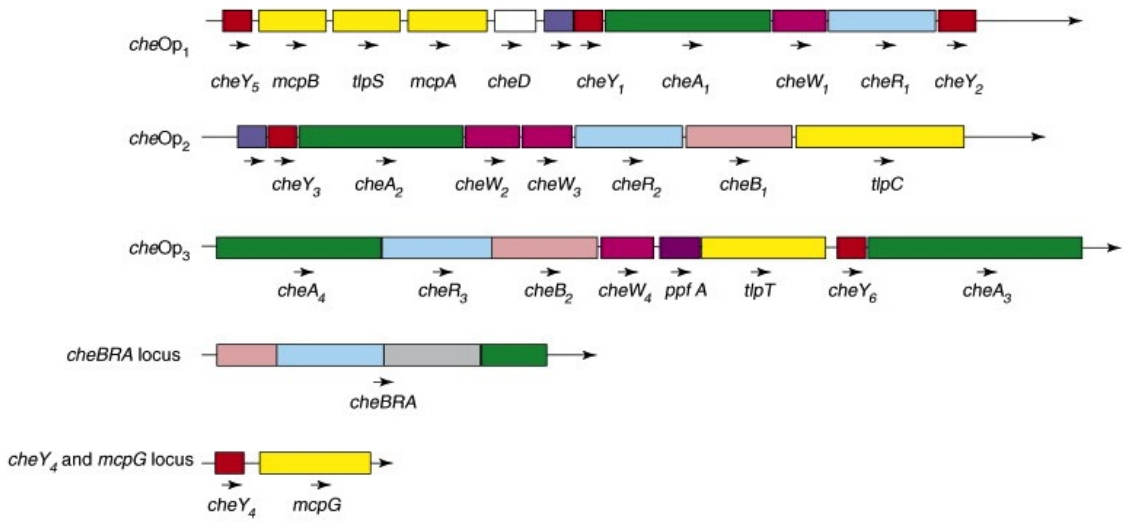


Figure 2.3: The chemotaxis operons and loci of *R. sphaeroides*.
cheOp_1, *che*Op_2, *che*Op_3 and *che*BRA are encoded on chromosome I. The *che*Y_4-*mcp*G locus is encoded on chromosome II. Figure taken from reference [92].

of immuno-gold electron microscopy and fluorescent protein fusions has shown that *R. sphaeroides* chemotaxis proteins localize to two distinct regions of the cell [130]. The proteins encoded by *che*Op_2 (with the exception of TlpC) localize to the cell pole along with the transmembrane chemoreceptors (MCPs) [135]. *che*Op_3 chemotaxis proteins localize to a cluster in the cytoplasm of the cell. The chemoreceptors which localize to the cytoplasmic cluster are known as transducer-like proteins (Tlps) because they lack a transmembrane domain but are otherwise homologous to MCPs. Both TlpT and TlpC have been shown to localize to the cytoplasmic cluster [130, 136].

The polar chemotaxis cluster contains one CheA homologue ($CheA_2$) and two CheW homologues ($CheW_2$ and $CheW_3$) [131]. Due to the similarity in composition to the *E. coli* chemotaxis cluster, it is believed that signaling in this cluster closely mimics that of *E. coli*. However, it remains unclear why there is a requirement for two CheW proteins. Like the *E. coli* CheA, the polar $CheA_2$ autophosphorylates and can phosphotransfer to CheB and CheY response regulators. *In vitro* studies have shown that $CheA_2$-P can phosphorylate $CheB_1$, $CheB_2$, $CheY_3$, $CheY_4$, and $CheY_6$,

but *in vivo* localization suggests that $CheB_2$ and $CheY_6$ are primarily associated with the cytoplasmic cluster rather than the polar one [92].

As compared to the *R. sphaeroides* polar cluster, the *R. sphaeroides* cytoplasmic cluster differs more significantly from the *E. coli* cluster, especially in relation to its CheA proteins. The two CheA homologues associated with the cytoplasmic cluster, $CheA_3$ and $CheA_4$ are both atypical CheA proteins. Unlike canonical CheA proteins, which contain five domains (P1-P5), $CheA_3$ and $CheA_4$ contain different subsets of these domains. Since $CheA_3$ contains only P1 and P5 domains and $CheA_4$ contains P3, P4, and P5 domains, neither can autophosphorylate as a homodimer. Instead, the P4 domain of $CheA_4$ phosphorylates the P1 domain on $CheA_3$, which them phosphotransfers to the response regulators $CheB_2$ and $CheY_4$ [92]. Interestingly, the P5 domains of each CheA are very different and bacterial two-hybrid (B2H) studies show that while TlpT will interact with the P5 domains of both $CheA_3$ and $CheA_4$, TlpC only interacts with $CheA_3$, suggesting it may only control the phosphatase activity.

R. sphaeroides does not have a CheZ homologue and therefore has adapted one of its CheA protein to serve as a phosphatase in order to achieve rapid signal termination. The 794 amino acid linker region that connects P1 and P5 of $CheA_3$ has been shown to act as a phosphatase for $CheY_6$-P. This activity is required for chemotactic ability, demonstrating the central importance of a phosphatase for signal termination in the chemotaxis pathway [137].

Both the polar and cytoplasmic chemosensory clusters are essential for *R. sphaeroides* chemotaxis and deletion of any components of either of the two clusters results in a loss of taxis. It is believed that the polar cluster senses changes in the external environment, while the cytoplasmic cluster is used to sense the internal metabolic state of the cell. Most *R. sphaeroides* chemotaxis components can complement deletions of their homologues in *E. coli*, but will not complement deletions of their homologues within the other *R. sphaeroides* chemosensory cluster. Exchanging the domains re-

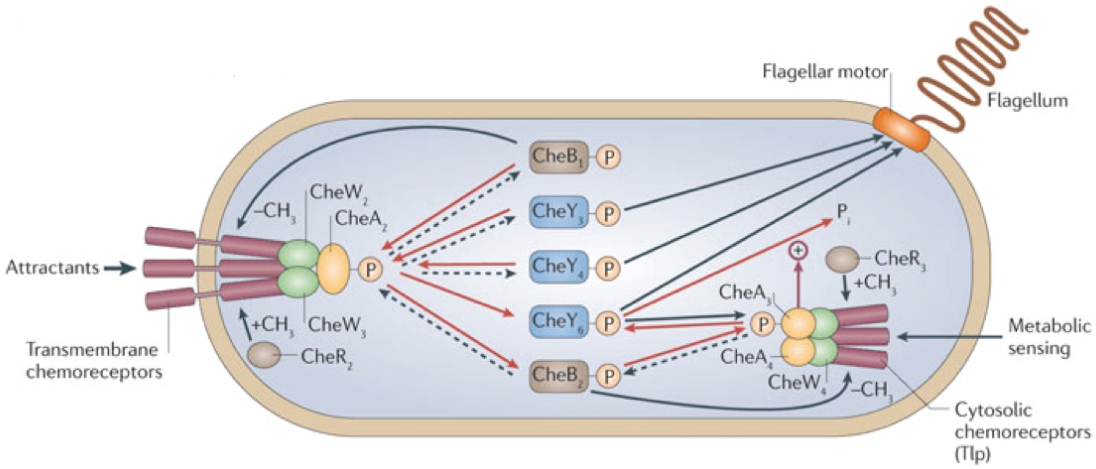


Figure 2.4: The chemotaxis pathway of *R. sphaeroides*.
The chemotaxis proteins controlling the *fla*1 flagellum from two distinct sensory clusters, one containing cytoplasmic chemoreceptors (Tlps) and another containing transmembrane chemoreceptors (MCPs). $CheB_2$-mediated phosphorelay (represented with dashed arrows) connects the polar and cytoplasmic clusters such that after being phosphorylated by $CheA_3$-P, $CheB_2$ can phosphorylate $CheA_2$ which can then phosphorylate any of its cognate response regulators. The reactions that allow $CheY_6$ to act as a phosphate sink are shown with red arrows. Figure taken from reference [138].

sponsible for CheA localization between polar and cytoplasmic homologues results in a reversal of localization, but a loss of chemotaxis [139]. This is indicative that the function of the homologous proteins is specific to the cluster to which they are natively localized.

Both operons encode CheB and CheR homologues that are believed to localize along with the other components of the operon in which they are encoded. Each pair is thought to be responsible for the adaptation of the chemoreceptors at the cluster to which it is localized. It is possible that there is cross-talk between the two clusters since $CheA_2$ can phosphotransfer to both homologues of CheB. Intriguingly, deletion of $CheB_1$ of $CheR_3$ lead to smooth swimming while deletion of $CheB_2$ or $CheR_2$ lead to tumbly swimming, suggesting the pattern of adaptation is reversed between the 2 clusters, with the polar cluster phosphorylating $CheY_3$/$CheY_4$ on addition of attractant.

R. sphaeroides has six CheY homologues, of which three ($CheY_3$, $CheY_4$, and $CheY_6$) are expressed under laboratory conditions. $CheY_3$ and $CheY_4$ interact exclusively with the polar cluster, while $CheY_6$ mainly interacts with the cytoplasmic cluster [140]. $CheY_6$ is essential for chemotaxis and is the only CheY homologue that has been shown to be able to stop the flagellar motor. However, either $CheY_3$ or $CheY_4$ is required in addition to $CheY_6$ in order to maintain a chemotactic phenotype [140]. The roles of $CheY_3$ and $CheY_4$ remain unclear though they seem to have some functional redundancies.

2.3 Chemoreceptor Structure and Signal Transduction

MCPs are largely α-helical proteins that span the bacterial cytoplasmic membrane. The periplasmic ligand binding domain exists in between two transmembrane (TM) helices. A HAMP signaling domain (found in histidine kinases, adenylyl cyclases, methyl-accepting chemotaxis proteins, and phosphatases) connects the second TM region to the long cytoplasmic domain **(Figure 2.5b)**. The cytoplasmic domain is further divided into a methylation, flexible bundle and signaling subdomains. The C-terminus of the *E. coli* MCP contains a conserved CheR binding motif known as the pentapeptide tail.

MCPs form stable homodimers in the cytoplasmic membrane and exist *in vivo* as trimers-of-dimers as demonstrated by cysteine cross-linking studies **(Figure 2.5a)** [141]. Nanodisc segregation of homodimers has shown that they are able to bind ligands and transduce signals, but the minimal functional unit for CheA kinase activation is a trimer-of-dimers [115]. Trimers-of-dimers containing a mixed composition of chemoreceptor homodimers can exist, but it is believed that only receptors with the same cytoplasmic domain length can form such interactions [142]. A deviation from the trimers-of-dimers arrangement was observed when the cytoplasmic domain

of *Thermatoga maritima* MCPI143C crystallized as linear 'hedgerows' [143]. It remains unclear whether this arrangement is a result of the crystallization conditions or whether this results from an inherent difference between MCPI143C and other receptors such as Tsr, whose cytoplasmic domain crystallized as a trimer-of-dimers [12].

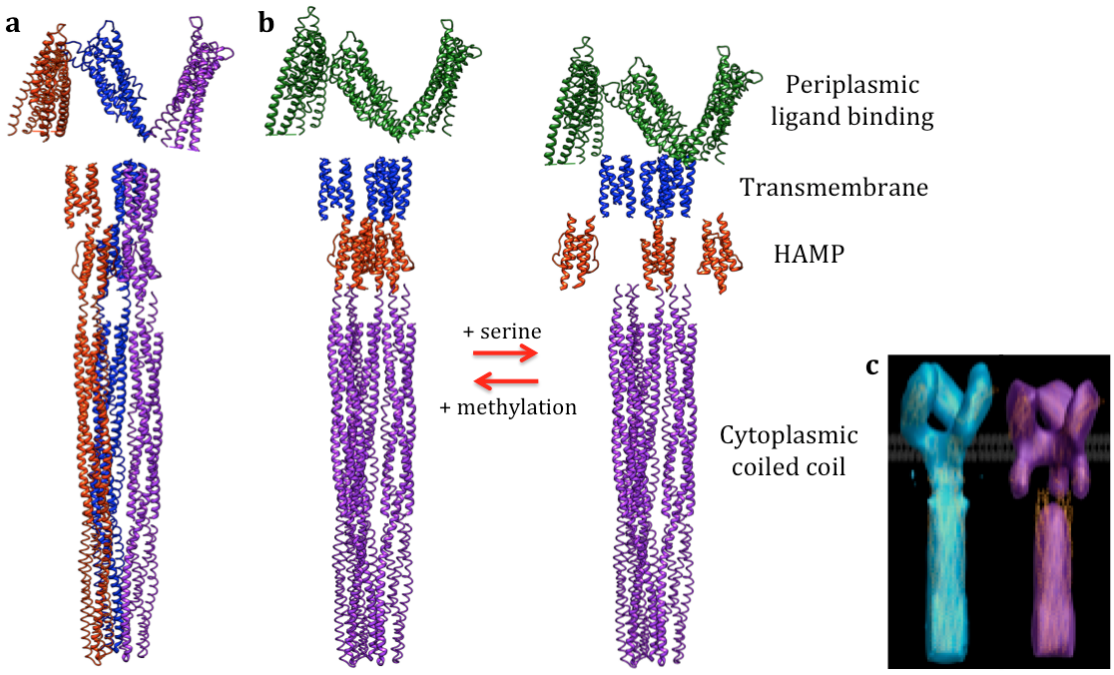


Figure 2.5: Structure of chemoreceptor trimer-of-dimers.
a) Model for the Tsr trimer-of-dimers denoting each of the three homodimers in a different color. b) Two conformations of the Tsr trimer of dimers determined based on electron microscopy show a large expansion of the HAMP domain that accompanies binding of the ligand, serine. The protein domains are colored as follows: the periplasmic ligand-binding domain in green, the transmembrane domain in blue, the HAMP domain in orange, and the cytoplasmic domain in purple. c) The averaged density maps from from cryo-electron tomography that were used to create the models in (a) and (b) from the atomic models of the Tsr periplasmic domain (PDB ID code 1VLS), the HAMP domain from an archaeal protein (PDB ID code 2ASW), the Tsr cytoplasmic domain (PDB ID code 1QU7) and a modeled transmembrane domain [144]. Coordinates in (a) and (b) taken from [145] and colored here for clarity. Figure in (c) taken from [145].

The periplasmic ligand binding domains can vary greatly between different MCPs are they each recognize and bind to different ligands. In the case of *E. coli* where the four receptors (Tsr, Tar, Trg, and Tap) likely arose from gene duplication, they

contain very similar sequences for their periplasmic domains. However, comparison of the crystal structures of the periplasmic domains of both Tar and Tsr show that though many of the interacting residues involved in ligand-binding are conserved, some of the residues are arranged differently in order to create specificity in ligand recognition and binding [146].

E. coli MCPs contain two transmembrane α-helical regions, TM1, which is the on the N-terminus of the ligand binding domain, and TM2, which connects the periplasmic and HAMP domains. TM2 transmits signals from changes in ligand binding rates to the cell via the HAMP domain. Attractant binding to *E. coli* Tar results in a 1-3Å piston-like displacement of one TM2 helix in a homodimer **(Figure 2.6)** [147]. Upward or downward shifts of TM2 within the membrane by changing its length alter signaling through TM2 and how readily the receptor can be modified by the adaptation enzymes [148, 149]. Moreover, cysteine cross-linking experiments have shown that the upward and downward motions of one transmembrane helix with respect to another can alter the signaling state of the receptor [141]. It is proposed that the piston-like displacements of TM2 are transmitted to the HAMP signaling domain by a five-residue control cable segment, whose helicity and interactions with a single trigger residue are responsible for modulation of signal transmission [150]. In addition to the displacement of TM2, there may also be a rotational movement of the transmembrane helices that extends from the rotations observed in the periplasmic helices upon ligand binding [151].

The proposition of helix rotation as a signaling mechanism is further supported by the solution NMR structure of an archaeal HAMP domain. The structure suggests that the HAMP domain, which consists of two amphipathic helices from each member of the homodimer, alternates between a canonical coiled coil and a homodimeric, four-helical parallel coiled coil [152]. It has been suggested that the HAMP domain may exist in one or more conformations that greatly differ from the archaeal NMR structure because the HAMP domain is extremely dynamic [26]. Another proposed movement of the HAMP domain in response to ligand binding involves a change in

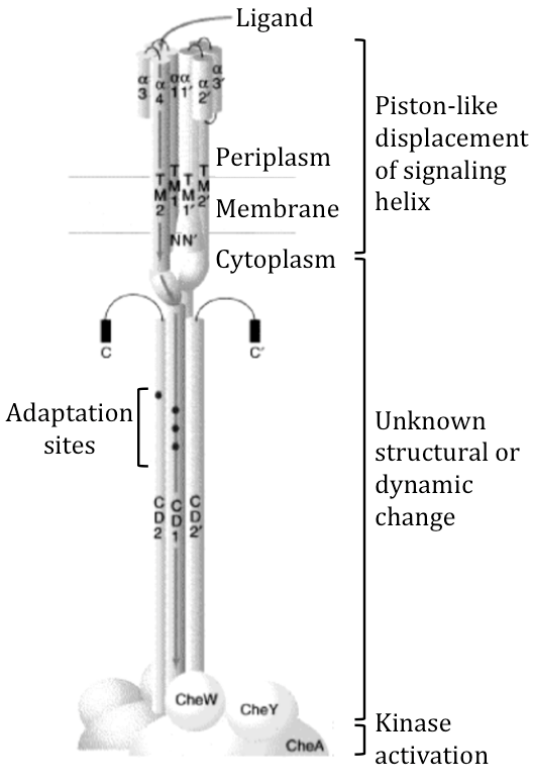


Figure 2.6: Proposed mechanism of chemoreceptor signal transduction. A schematic representation of an *E. coli* receptor homodimer showing the piston-like displacement of TM2 that is associated with ligand binding. The experimentally-determined methylation sites are denoted at their approximate positions. Figure taken from [26].

the tilt angle between the helices that causes them to open outward [153, 154]. The crystal structure of three concatenated HAMP domains from *Pseudomonas aeruginosa* provides evidence for yet another structural state with a very different parallel four-helix bundle [155]. Cryo-electron tomography of overexpressed Tsr in *E. coli* has shown that the quarternary conformation of the HAMP domain of a trimer-of-dimers expands to nearly double the diameter as a consequence of ligand binding **(Figure 2.5b,c)** [145]. Further studies will be required to determine whether these large movements of the HAMP domains occur in physiological assemblies of Tsr or whether they are by-products of the zippered receptor assemblies formed from overexpression. Thus, the precise conformational changes associated with signaling in the HAMP domain, both at the level of tertiary and quaternary levels, require further work in order to be conclusively determined.

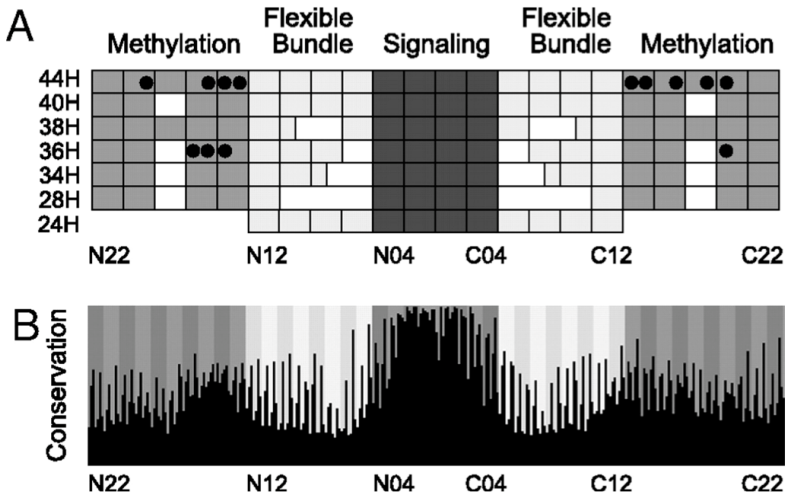


Figure 2.7: Classification of MCPs by cytoplasmic domain heptad repeats.
A) Heptad repeats in the seven major MCP classes (24H, 28H, 34H, 36H, 38H, 40H, 44H) are shown as rectangles where the three subdomains are indicated with different shades of grey. Each rectangle represents a group of two heptads. Heptads are numbered symmetrically from the N and C termini beginning at number 22 (N22 and C22) and reducing to N01 and C01 at the central hairpin located at the cytoplasmic tip of each monomer. B) Amino acid sequence conservation across all MCPs studied shows that the highest levels of conservation are in the signaling subdomain and the first and fourth residues of each heptad are strongly conserved. Figure taken from reference [156].

The cytoplasmic domain consists of two antiparallel helices (CD1 and CD2) connected by a hairpin turn at the membrane-distal tip of the receptor. In receptor homodimers, the helices of two monomers come together to form a long, supercoiled four-helical bundle, which in turn interact at a highly conserved region near the hairpin turn to form trimers-of-dimers [12]. Large-scale comparative genome analysis has revealed that the cytoplasmic domain is formed by symmetric pairs of multiple seven-residue (heptad) repeats, each of which corresponds to exactly two α-helical turns. Based on the presence of symmetric heptad repeats, seven major MCP classes have been identified **(Figure 2.7)**. The same analysis has revealed the presence of a highly conserved signaling subdomain closest to the hairpin turn as well as a flexible coiled-coil bundle and a methylation region required for adaptation

[156].

The C-terminal region of the MCP consists of an unstructured, flexible linker, which has been shown to end in the pentapeptide NWETF in *E. coli* and *Salmonella typhimurium* Tar and Tsr. In *E. coli*, this pentapeptide motif is the sole site of CheR binding and loss of this motif reduces methylation, and thus chemotaxis, in these organisms. Structural studies and multiples sequence alignments were used to suggest a general consensus sequence (X-[HFWY]-X-X[HFWY]) for the pentapeptide tether [156, 157, 158]. However, this consensus sequence appears only in about 10% of the MCPs analyzed, implying that this tether may not be a general mode of interaction between MCPs and adaptation enzymes [156].

2.4 Chemoreceptor Arrays

MCPs form large protein clusters, containing tens of thousands of copies of the receptors and associated chemotaxis proteins [159]. These clusters vary in size, are generally circular or ellipsoid in shape, and are often, but not always localized to the cell poles [160, 161, 162]. The higher order organization of these receptors is believed to be responsible for a number of attributes of the chemotaxis system including high sensitivity, wide dynamic range, extensive cooperativity, and precise adaptation [128].

2.4.1 Architecture

MCPs are not however usually found as single trimers of dimers, but form extended arrays with the CheA kinase and CheW coupling protein. Cryo-electron tomography has revealed that MCPs are organized perpendicular to the membrane with CheA and CheW binding to the cytoplasmic tips of the receptors and forming a "base plate" parallel to the membrane [163]. Further studies have shown that MCP trimers-of-dimers pack in ordered hexagonal arrays that have been evolutionarily

conserved across species from various bacterial phyla [164]. One proposed arrangement for these hexagonal arrays involves a ring of three alternating CheA P5 domains and three CheW proteins surrounded by a lattice of six receptor trimers-of-dimers [165]. This array arrangement prevents the dimeric CheA from being present in every receptor hexagon, necessitating that one empty hexagon is surrounded by six hexagons containing a CheA/CheW ring. An alternate arrangement proposes that the empty rings are occupied by six CheW proteins [166]. Experimentally measured stoichiometries of CheA:CheW:MCPs vary depending on the sample and preparation, indicating that not all CheA or CheW positions may need to be filled in order to form a stable array and that CheW may substitute for CheA at certain places in the lattice [165].

In the absence of CheA and CheW, MCPs at physiological concentrations do not cluster and are instead diffuse in the membrane, with the degree of polar localization varying by species [168, 135]. However, overexpression of MCPs in the absence of CheA and CheW has been shown to result in "zipper-like" assemblies with receptor-receptor interactions forming at the cytoplasmic tips of two opposing layers of membrane-associated receptors [163]. These zipper-like assemblies also demonstrate hexagonal packing at the overlapping cytoplasmic ends of the receptors, though the center-to-center hexagonal spacing is 9 nm in the zipper-like assemblies rather than 12 nm in native arrays [169]. This demonstrates that CheA and CheW play a structural role in maintaining appropriate clustering and spacing of MCPs in the arrays in addition to their role in signal transduction.

The described arrangements relate to membrane-associated arrays involving MCPs. However, transducer-like proteins (Tlps) that lack a transmembrane domain have been shown to localize to a cytoplasmic chemotaxis cluster [130, 136]. Recent cryo-electron tomography studies of *R. sphaeroides* and *Vibrio cholerae* have shown that like MCPs, Tlps adopt a trimer-of-dimers arrangement with 12 nm hexagonal spacing. However, the arrangement of cytoplasmic chemoreceptor arrays differs from membrane-associated arrays in that two CheA/CheW base plates sandwich two sets

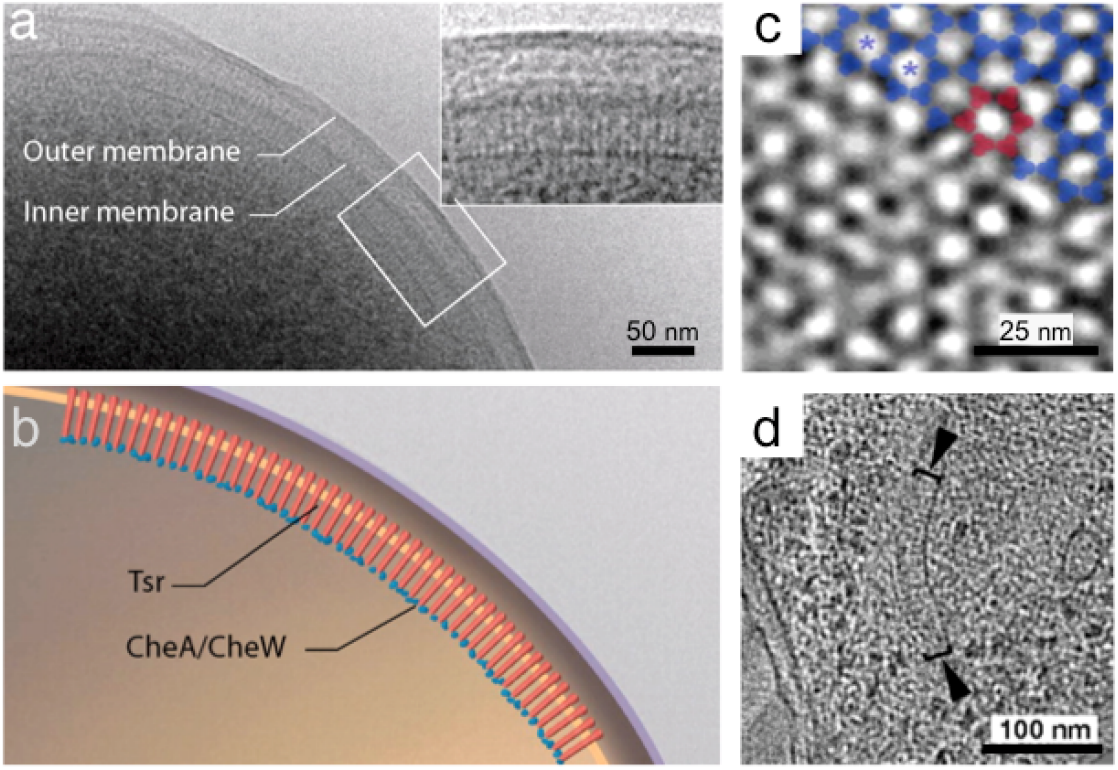


Figure 2.8: Architecture of membrane-associated and cytoplasmic chemoreceptor arrays.
a) Cryo-electron micrograph of a side view of a membrane-associated chemoreceptor array in *E. coli*. b) Schematic representation of a side view of a Tsr array showing the receptors and the CheA/CheW base plate. c) Receptor trimers-of-dimers denoted on the tomographic slice of hexagonal lattice of a chemoreceptor array from *Vibrio cholerae*. d) Cryo-section of *R. sphaeroides* showing the cytoplasmic chemoreceptor array. Figure taken from references [163, 164, 167].

of receptors oriented in opposite directions [167]. The results suggest that the 12 nm hexagonal spacing is an essential feature for chemoreceptor arrays and that sandwiching two sets of receptors may be a mechanism for stabilization, substituting for the membrane.

The existence of two differently structured chemoreceptor arrays at distinct parts of the cell begs the question of how the arrays are correctly positioned and partitioned.

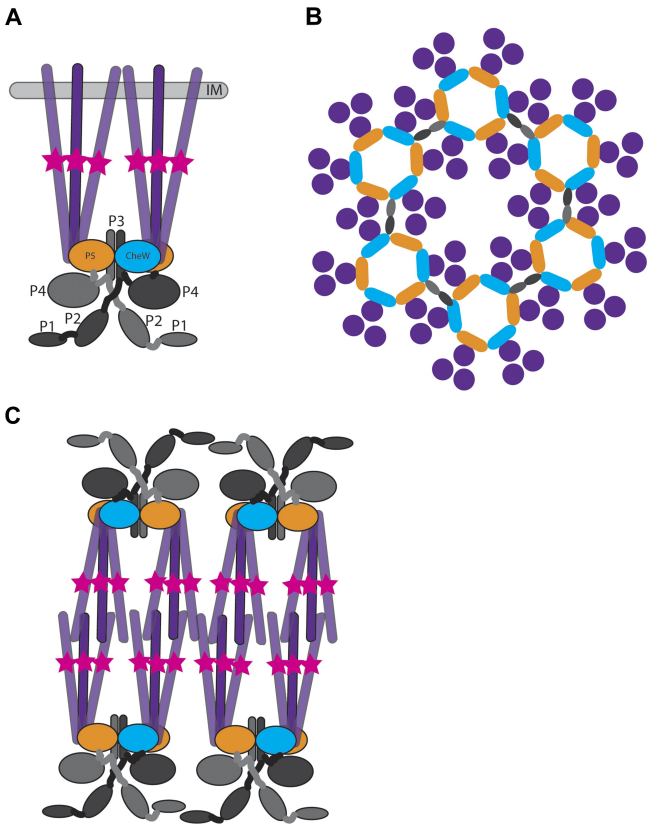


Figure 2.9: Schematic representations of membrane-associated and cytoplasmic chemoreceptor arrays.
A&B) Side view (A) and top view (B) of the arrangements in the membrane-assocaited arrays showing the interactions between the receptors, CheA and CheW. C) Side view of the cytoplasmic chemoreceptor array showing the sandwich of receptors interacting at their presumably ligand-binding ends. Figure taken from reference [167].

2.4.2 Positioning and Partitioning

Although membrane-associated chemoreceptor arrays are often localized to the cell poles, it is common to find that cells with multiple clusters will have some of them localized laterally [170]. Studies of chemoreceptor cluster positioning in *E. coli* suggest that receptors inserted in the membrane are positioned independently of any cellular markers of anchors and can either join an existing cluster or nucleate a new one [171]. A continuous distribution of cluster sizes is indicative that small clusters form and are either absorbed into larger clusters or grow to form a new large cluster. This is supported by the observation that the greatest density of small clusters is furthest from the largest cluster, as nearby clusters would be absorbed into the large

cluster [172]. This stochastic self-assembly model provides an elegant explanation for *E. coli* cluster formation, positioning, and partitioning during cell division.

Other species, such as *R. spheroides*, show deviations from the *E. coli* paradigms for cluster positioning. As in *E. coli*, membrane-associated chemoreceptor clusters in *R. sphaeroides* can be positioned both at the cell poles and laterally. However, *R. sphaeroides* differs from *E. coli* in that its position away from the FtsZ-ring that marks the septation site. Studies have also suggested the formation of mobile clusters of ∼800 receptors that move to the poles where they become trapped, suggesting a preference for the membrane curvature of the cell pole [173]. These results suggest that the positioning and partitioning of membrane-associated chemotaxis clusters in *R. sphaeroides* is a combination of stochastic self-assembly, diffusion-capture by the polar regions, and exclusion from regions around the FtsZ ring.

The cytoplasmic cluster of *R. sphaeroides* localizes to mid-cell with duplicated clusters found at $\frac{1}{4}$ and $\frac{3}{4}$ positions prior to cell division [174]. The positioning and partitioning of the cytoplasmic chemosensory cluster of *R. sphaeroides* has been shown to depend on the ParA homologue, PpfA [174]. Like ParA, PpfA is a Walker box ATPase where ATP binding to the Walker box promotes dimerization, DNA binding and polymerization [175, 176]. Interaction with a partner protein that is part of a complex to be segregated (ParB in the case of ParA [177] and the N-terminal region of TlpT in the case of PpfA [178]) stimulates ATPase activity, resulting in the depolymerization of the ParA filament [179]. This series of events in the ParA/B system is a common mechanism for plasmid and chromosome segregation although the mechanism by which polymerization and depolymerization result in segregation is unclear.

PpfA is encoded immediately upstream of the cytoplasmic chemoreceptor TlpT in *che*Op$_3$ **(Figure 2.3)**. Studies have shown that genomic replacements of WT PpfA with mutants defective in dimerization, ATP binding, ATP hydrolysis, and DNA binding all resulted in a loss of cytoplasmic chemosensory cluster segregation in *R.*

sphaeroides [178]. The expression of a PpfA-CFP fusion in *R. sphaeroides* showed that PpfA was distributed across the nucleoid with bright foci corresponding to the cytoplasmic cluster. Mutants that could not dimerize, bind ATP or bind DNA did not display such foci. However, the mutant that was locked in the ATP-bound state and could not hydrolyze ATP (D39A mutant) produced brighter foci than WT [178]. Expression of CFP tagged variants of PpfA in *E. coli* showed that DNA binding as dependent of ATP binding and dimerization. However, the D39A mutant of PpfA that cannot hydrolyze ATP remained co-localized with the nucleoid [178].

These results show that segregation of the cytoplasmic cluster by PpfA is accomplished by dimerization, ATP hydrolysis, and DNA binding. Although these steps are similar to those used by ParA for plasmid and chromosome segregation, PpfA does not appear to oscillate over the chromosome as is typical of many ParA proteins [180]. Therefore, the exact mechanism by which PpfA segregates the cluster remains unclear.

Studies to date suggest that the overall organization of all chemotaxis systems is similar, with receptors and kinases similarly organized into arrays. It follows that there should also be mechanisms in place to ensure that each daughter cell inherits a functioning cluster on division and is therefore able to compete in a dynamic environment.

2.5 Aims of this Section of the Thesis

Given the apparent conservation of the architecture of chemosensory arrays across species, this section focuses on the use of imaging techniques to gain insights on the structure and organization of two proteins involved in bacterial chemotaxis, the *E. coli* serine chemoreceptor (Tsr) and the *R. sphaeroides* cytoplasmic chemosensory segregation protein (PpfA).

Tsr was studied in an attempt to determine the movements of individual helices in

the full length protein by determining a sub-nanometer resolution structure. As a transmembrane receptor with a flexible signaling domain, full-length Tsr has long been a challenging structural target which has not yet yielded to crystallization. Previous electron microscopy studies have shown that cryo-electron tomography and sub-volume averaging are viable techniques for probing Tsr structure in a membrane environment both in systems where Tsr is overexpressed or expressed at physiological levels.[145, 165]. However, the previous studies have been limited in resolution to approximately 30Å as a result of technological limitations and the thickness of whole cell samples. This study sought to take advantage of methodological developments in image detectors in order to obtain higher resolution structures. In an attempt to further improve the achievable resolution, efforts were taken to explore other types of full-length Tsr samples that would yield thinner EM specimens. These included attempts to reduce the thickness of whole *E. coli* cells, express Tsr in thin *E. coli* cells, and isolate membranes containing Tsr assemblies. Additionally, though it removes the receptor from its native arrays and membrane environment, detergent solubilization of Tsr was pursued in hopes that it may allow for single particle cryo-EM structure determination since that is currently the highest resolution electron microscopy technique available. Notably, near-atomic resolution structures have recently been reported for other membrane proteins solubilized using amphipathic polymers or detergents [58, 181].

The partitioning protein PpfA is of interest because its organization and segregation mechanism may provide a basis for how other Walker box ATPases use variations on the ParA/B mechanism to segregate protein clusters. This particular study looked at the organization of PpfA, which has previously been studied using photo-activated localized microscopy (PALM) in the Armitage laboratory. The poor z resolution in PALM makes it difficult to determine whether PpfA is occluded by the nucleoid since fluorescence spots above and below the plane of focus remain visible. Visualization of cryo-sections offers the advantage of having distinct z slices while the immuno-gold labeling provides a methods for specifically identifying the location of PpfA in these slices.

The specific aims of this section are:

1. Study the structure and organization of Tsr:

   (a) Optimize whole cell, membrane isolate, and detergent solubilized samples of Tsr for subsequent cryo-electron tomography or microscopy studies (including optimization of expression and purification of Tsr as well as reducing sample thickness for tomography studies).

   (b) Gain insights into the organization of zipper-like receptor assemblies that result from Tsr overexpression using cryo-electron tomography. Since these arrays are non-functional, comparison to the higher order structure of functional arrays may yield insights into the requirements for array functionality.

   (c) Determine the structure of full-length Tsr in the native membrane environment using cryo-electron sub-tomogram averaging. Structures of Tsr were sought in both the unliganded state and after incubation with the receptor's native ligand, serine.

2. Investigate the localization of PpfA:

   (a) Phenotypically validate the cluster segregation ability of tagged PpfA at different induction levels using fluorescence light microscopy.

   (b) Compare the differential localization of the wild-type and ATP-locked mutant forms of PpfA using immuno-gold labeling and electron microscopy.

Chapter 3

Methods

3.1 Bacterial Strains and Growth Conditions

3.1.1 Strains and Plasmids

Several *Escherichia coli* and *Rhodobacter sphaeroides* strains were used in this study. The specific strains used are listed in **Table 3.1** and plasmids used are listed in **Table 3.2**.

Table 3.1: Bacterial strains used in this study.

Strain	Species	Genotype/description	Source
XL1-Blue	*E. coli*	General cloning strain, used for plasmid creation and modification. Tetracycline resistant.	Stratagene
NEB 5-alpha	*E. coli*	General cloning strain, used for plasmid creation and modification. Resistant to nalidixic acid.	New England Biolabs
One Shot BL21(DE3)	*E. coli*	General expression strain, used for inducible protein expression.	Invitrogen
HCB721	*E. coli*	$\Delta tsr(7021)$ *trg*::Tn*10* $\Delta(cheA\text{-}cheY)$::*XhoI*(Tn*5*) (does not express the receptors Tar, Tsr, Trg, and Tap, the cytoplasmic signaling proteins CheA and CheW, and the adaptation proteins CheR and CheB). Kanamycin resistant.	[182]

B/r H266	*E. coli*	Thin B strain containing no motility or chemotaxis machinery.	[183]
S17-λ pir	*E. coli*	Plasmid mobilizing strain, used for introducing plasmids into *R. sphaeroides* strains. Streptomycin resistant.	[45]
WS8N	*R. sphaeroides*	Spontaneous nalidixic acid resistant strain of the wild-type strain WS8.	[184]
JPA1558	*R. sphaeroides*	*tlpT-yfp*	[131]
JPA1519	*R. sphaeroides*	*yfp-cheW4* Δ*tlpT*	[131]

Table 3.2: Plasmids used in this study.

Plasmid Name	Description	Source
pHSe5.tsr_{QEQE}	Wild-type form of Tsr under the control of a double pTac IPTG-inducible promoter, Ampicillin resistant	[185]
pET28b	Commercial T7 cloning and expression vector, Kanamycin resistant	Novagen
pET28b-*Strep*	pET28b containing a thrombin cleavage site and *Strep*-tag	This study
pET28b-*tsr*-*Strep*	pET28b containing *tsr* with a C-terminal thrombin cleavage site and *Strep*-tag	This study
pHSe5.*tsr*-*Strep*	pHSe5.tsr_{QEQE} containing a C-terminal thrombin cleavage site and *Strep*-tag	This study
pIND4	IPTG inducible plasmid for gene expression in *R. sphaeroides*, Kanamycin resistant	[186]
pIND4-*ppfA* -CFP	pIND4 containing a C-terminal fusion of CFP to *ppfA*	Chris Jones (Armitage laboratory)
pIND4-*ppfAD39A*-CFP	pIND4 containing a C-terminal fusion of CFP to the D39A mutant of *ppfA*	Chris Jones (Armitage laboratory)
pIND-*ppfA-APXII-HA*	pIND4 containing a C-terminal *ApexII-HA* fusion to *ppfA*	This study
pIND-*ppfAD39A-APXII-HA*	pIND4 containing a C-terminal *ApexII-HA* fusion to the D39A mutant of *ppfA*	This study

pIND-*HA-APXII-ppfA*	pIND4 containing a N-terminal *HA-ApexII* fusion to *ppfA*	This study
pIND-*HA-APXII-ppfAD39A*	pIND4 containing a N-terminal *HA-ApexII* fusion to the D39A mutant of *ppfA*	This study

3.1.2 *E. coli* Growth Conditions

Frozen stocks of *E. coli* or liquid cultures requiring antibiotic selection were spread on Lysogeny broth (LB) agar (1.5% w/v agar) plates and grown overnight at 37°C. *E. coli* liquid cultures were grown overnight at 37°C with 225 rpm shaking in LB broth. Note that all buffer and media conditions are listed in Appendix A.

3.1.3 *R. sphaeroides* Growth Conditions

R. sphaeroides frozen stocks or liquid cultures requiring antibiotic selection were spread on LB agar plates and grown for 2-3 days at 30°C. *R. sphaeroides* liquid cultures were grown in succinate medium at 30°C. Liquid cultures were either grown photoheterotrophically without shaking under white light or aerobically in the dark with shaking at 30°C. In both cases, growth for 2-3 days was required to obtain a stationary culture.

3.1.4 Antibiotics

Antibiotics were added to the liquid culture media or agar plates in order to select for successful plasmid uptake or the correct bacterial strain. Antibiotics were added at the working concentrations shown in **Table 3.3** and were made up at 1,000-fold the working concentration, filter sterilized, and stored at -20°C. For transformations prior to Tsr protein induction **(Section 3.3.1)**, carbenicillin was used as an alternative to ampicillin in order to reduce the presence of satellite colonies on LB agar

plates.

Table 3.3: Working concentrations of antibiotics used for selection.

Antibiotic	Concentration for *E. coli*	Concentration for *R. sphaeroides*
Kanamycin (Kan)	50 μg/mL	25 μg/mL
Nalidixic acid (Nal)	25 μg/mL	25 μg/mL
Ampicillin (Amp)	100 μg/mL	–
Carbenicillin (Carb)	100 μg/mL	–
Tetracycline (Tet)	25 μg/mL	–

3.1.5 Storage

Freezer stocks for long-term storage of bacterial cells at -80°C were made by mixing 600 μL stationary culture and 400 μL of 60% glycerol in a cryotube.

3.2 Genetic Manipulations

3.2.1 Cloning

Polymerase Chain Reaction (PCR)

PCRs were performed with Phusion High-Fidelity DNA Polymerase (New England Biolabs) using the reaction mixtures described in **Table 3.4**. For the production of small tags that were generated by overlap extension of the primers **(Figure 3.1A)**, no template DNA was added. In cases were DNA fragments were joined by overlap extension PCR, the two fragments to be attached were both used as templates with the outer primers **(Figure 3.1B)**.

All primers were synthesized by Sigma-Genosys or IDT based on the sequences listed in **Appendix B**. PCR reaction mixtures were placed in a thermocycler and run using the program outlined in **Table 3.5**. The annealing temperature for each reaction was determined based on the melting temperatures (T_m) of the primers.

Table 3.4: Composition of PCR reaction mixture.

Component	Quantity
Template DNA	0.5 μL
Primers	2.5 μL of each at 10 μM
dNTPs	1 μL at 10mM
5x Phusion GC Buffer (New England Biolabs)	10 μL
Phusion High-Fidelity DNA Polymerase (New England Biolabs)	0.5 μL
MilliQ water	Make up to 50 μL

Table 3.5: PCR reaction program.

Stage	Temperature	Time
Initial denaturation	98°C	2 min
Repeat the following cycle 25 times		
Denaturation	98°C	10 sec
Annealing	T_m specific (50-65°C)	20 sec
Extension	72°C	20 sec per kb
Final extension	72°C	10 min
Hold	10°C	Indefinite

Gel Electrophoresis

DNA fragments from PCR or plasmid digestion reactions were separated by gel electrophoresis. Gels were made with 1-2% agarose dissolved in 0.5x TBE buffer. The agarose concentration was selected to optimize separation of the DNA fragments of interest. DNA samples mixed with 5x loading dye were loaded into the wells of the gel alongside 1-kb plus DNA ladder (Invitrogen). The gel was run in 0.5x TBE buffer with a potential difference of 120V. The gel was subsequently stained in ethidium bromide for 15 minutes and DNA bands were visualized using a UV transilluminator with a wavelength of 360 nm. If DNA fragments were required for further cloning steps, their corresponding DNA bands were excised from the gel.

Gel Extraction

DNA fragments from PCR or digested plasmids were purified from agarose gels using the Gel Extraction Kit (QIAGEN). The procedure was performed using the

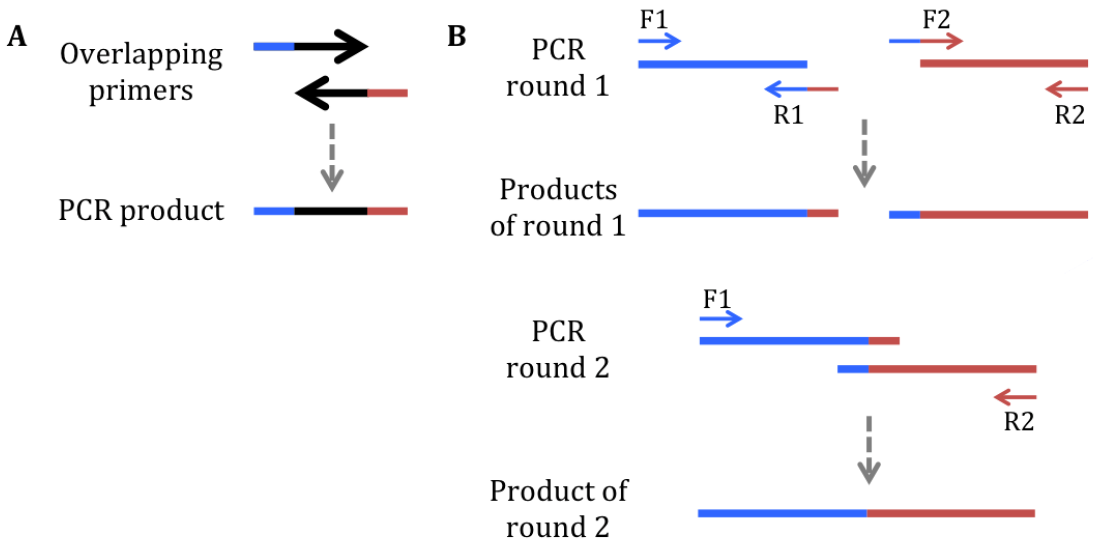


Figure 3.1: Overlap extension PCR.
Overlap extension PCR was used to extend primers containing a region of overlap to create small fragments (A) and to join larger fragments by creating a region of overlap (B). Attaching larger fragments together required two round of PCR, where the first amplified each of the two fragments separately. The two inner primers (R1 and F2) used in this step contained a region of overlap with the other fragment. The two fragments produced in the first round of PCR were both used as templates for the second round of PCR, which was performed with the outer primers (F1 and R2). Since the two fragments contain a region of overlap, they can anneal to each other and act as primers for elongation, thereby resulting in a fused sequence.

manufacturer's provided protocol and DNA was eluted in 50μL of MilliQ water.

Restriction Enzyme Digests

Rectriction enzyme digests were performed on purified PCR fragments or plasmids for subsequent ligation or for confirmation of plasmid identity. Restriction enzymes were obtained from New England Biolabs (NEB) and digests were performed using NEB's supplied buffers. Restriction digests were performed based on the mixture composition in **Table 3.6** and were incubated for at least 2 hours at 37°C.

Following restriction digestion, the enzymes were either inactivated by heating to 80°C for 20 min or removed using the standard gel extraction protocol without dis-

Table 3.6: Reaction mixture composition for restriction enzyme digests.

Component	Quantity
DNA	40μL for PCR products 10μL for plasmids
10 x buffer	5μL
Enzyme	1μL each
MilliQ water	make up to 50μL total volume

solving agarose. For plasmid digests, this was also followed by a 1 hour incubation at 37°C with 1 μL calf intestinal phosphatase (NEB) added to the 50 μL reaction or elution volume. This step was included to remove 3' phosphoryl groups, thereby preventing self-ligation of single-cute plasmids.

Ligations

Ligations of the digested products were set up as shown in **Table 3.7**. The volume of insert and vector were adjusted to give a molar ratio of approximately 3:1. Ligations were incubated for 1-2 hours at room temperature.

Table 3.7: Reaction mixture composition for DNA ligation.

Component	Quantity
Digested vector	1μL
Digested insert	3μL
5x T4 DNA ligase buffer (Invitrogen)	4μL
T4 DNA ligase (Invitrogen)	1μL
MilliQ water	11μL

Plasmid Purification

Sequencing quality plasmid DNA was extracted from 3-5 mL of overnight *E. coli* liquid cultures using the QIAGEN Plasmid Mini-Kit. The manufacturer's protocol was followed and the DNA was eluted in MilliQ water and stored at -20°C.

Sequencing

DNA sequencing was performed by Source BioScience Oxford or the NIH CCR Genomics Core using the BigDye termination method (PE Biosystems). Sequences were analyzed using ApE - A plasmid Editor version 2.0.47 (M. Wayne Davis).

3.2.2 DNA Transfer

Preparation of Competent Cells

Chemically competent *E. coli* cells were prepared or purchased from commercial vendors as described below.

XL1-Blue and S17-λ pir cells were made chemically competent using standard transformation buffer (TFB). An overnight *E. coli* liquid culture was diluted 50-fold in fresh LB broth and grown at 37°C to an OD_{600} of 0.4-0.6. Cells were transferred to 50 mL Falcon tubes and incubated on ice for 15-30 min and then centrifuged at 1000 x g for 10 min at 4°C. The pellets were resuspended in 16 mL ice cold TFB I buffer and incubated on ice for 15 min. Subsequently, cells were centrifuged at 750 x g and resuspended in 2 mL ice cold TFB II buffer. The cells were then split into 200 μL aliquots on ice, flash frozen in liquid nitrogen and stored at -80°C.

HCB721 and B/r H266 cells were made chemically competent using TSS and KCM. An overnight liquid culture was diluted 100-fold in LB broth and grown at 37°C to an OD_{600} of 0.5-0.8. Cells were placed in pre-chilled 50 mL falcon tubes and centrifuged at 3210 x g for 10 min at 4°C. The pellets were resuspended in 5 mL ice cold TSS and were split into 90 μL aliquots on ice and stored at -80°C. KCM was added to thawed cells immediately prior to transformation as described in the next section.

Chemically competent One Shot BL21(DE3) and NEB5-alpha cells were purchased from Invitrogen and New England Biolabs, respectively.

Transformation of *E. coli*

Each aliquot of XL1-Blue or S17-λ pir cells was thawed on ice. For each transformation, 100 μL of cells was combined with 10 μL of ligation mixture or 1 μL of purified plasmid DNA. Cells were incubated on ice for 30 min, heat shocked at 42°C for 1 min, and then immediately incubated on ice for 2 min. 800 μL of LB broth was added to cells and they were incubated at 37°C for 1 hour. Cells were then spread on LB agar plates containing the appropriate selective antibiotics.

For HCB721 and B/r H266 cells, 10 μL of 1x KCM was added to each 90 μL competent cell aliquot after the cells were thawed on ice. 1 μL of purified plasmid DNA was added to each aliquot and the resulting mixture was incubated on ice for 10-30 min. The cells were then heat shocked at 42°C for 90 sec and incubated immediately on ice for 2 min. 100 μL of LB broth was added to the cells and they were incubated at 37°C with 225 rpm shaking for 1 hour. Cells were then spread on LB agar plates containing the appropriate selective antibiotics.

BL21(DE3) and NEB5-alpha cells were transformed as per the manufacturer's instructions and were plated on LB agar containing selective antibiotics.

Conjugation into *R. sphaeroides*

R. sphaeroides cannot be made chemically competent, so plasmids with the pIND4 backbone were introduced into *R. sphaeroides* by conjugation with S17-λ pir *E. coli* cells. Plasmids were transformed into S17-λ pir cells by transformation **(Section 3.2.2)**. Overnight liquid cultures of S17-λ pir cells containing the desired plasmid were diluted 50-fold into LB broth and grown at 37°C to an OD_{600} of 0.4-0.6. 1 mL of this S17-λ pir culture and 1 mL of a two-day *R. sphaeroides* liquid culture grown photoheterotrophically were centrifuged separately at 6000 x *g*. Each pellet was washed in 1 mL of LB broth (without antibiotics) before being resuspended in 100 μL LB broth. During resuspension, special care was taken to avoid sheering the pili on the S17-λ pir cells. 10 μL of S17-λ pir cells was combined with 100 μL of

R. sphaeroides cells and the resulting mixture was placed on a nitrocellulose filter on an LB agar plate without antibiotics. After overnight incubation at 30°C, the filter paper was placed in an eppendorf tube with 800 μL of LB broth and vortexed to resuspend cells. The resulting cell suspension was spread on 2-4 LB agar plates containing nalidixic acid and the appropriate selection antibiotic for the introduced plasmid.

3.3 Protein Expression and Purification

3.3.1 Tsr Overexpression in *E. coli*

HCB721 or B/r H266 cells were transformed with the appropriate Tsr expression plasmid and plated on LB agar plates containing carbenicillin. pHSe5.*tsr-Strep* was used for affinity purification of detergent-solubilized Tsr and pHSe5.tsr_{QEQE} was used for all other experiments. After overnight growth at 37°C, a few milliliters of LB broth were added to the surface of the plate and colonies were scraped off of the plate and suspended in the liquid. This concentrated cell suspension was diluted in the desired media supplemented with ampicillin for growth and subsequent induction.

Induction in HCB721 cells

The concentrated cell suspension from the plate surface was diluted either in fresh LB broth or super broth (KD Medical) such that the initial OD_{600} was 0.01 for cell cultures that will be induced at an OD_{600} of ∼0.1 and 0.1 for cell cultures that will be induced at an OD_{600} of ∼0.8 or ∼1. Cultures were grown with 225 rpm shaking at the temperatures as specified in **Section 4.2**. In instances where the temperature was decreased at the time of induction, cultures were cooled prior to the addition of IPTG. When the desired OD_{600} for induction was reached, IPTG was added such that the final concentration was either 0.5mM or 1mM, depending on

the experiment **(Section 4.2)**. Cells used for subsequent membrane isolation and protein purification were harvested 10.5-14.5 hours after induction by centrifugation for 10 min at 6,000 x g at 4°C and cell pellets were stored at -80°C. Cells used for whole cell imaging were withdrawn from culture when the OD_{600} reached 0.5-0.8 and were immediately used to prepare vitreous specimens as described in **Section 3.4.2**.

Induction in B/r H266 cells

The cell suspension scraped from the plate was diluted in M9 minimal media supplemented with 0.04% glycerol such that the initial OD_{600} was 0.05. Cultures were grown at 37°C with 225 rpm until cells reached an OD_{600} of ∼0.1. Cultures were cooled to 25°C and IPTG was added to a final concentration of 0.5mM. The induction was allowed to proceed for 6 hours until the cells reached an OD_{600} of 0.68. Cells were withdrawn directly from the culture and used to prepare vitreous specimens as described in **Section 3.4.2**.

3.3.2 Tagged PpfA Expression in *R. sphaeroides*

Two-day photoheterotophic liquid cultures of *R. sphaeroides* with appropriate pIND4 plasmids were diluted 50-fold in 10 mL of succinate medium containing kanamycin. Cultures were grown aerobically overnight at 30°C and subsequently diluted 40-fold in 3 mL of succinate medium containing both kanamycin for selection and IPTG for PpfA induction (concentrations used were 0mM, 0.01mM, 0.1mM, 0.5mM, 1mM). These cultures were grown aerobically to an OD_{700} of 0.3-0.4. Cells were immediately placed on agarose pad slides for fluorescence microscopy **(Section 3.4.1)** or fixed for cryo-sectioning and immunogold labeling **(Section 3.4.3)**.

3.3.3 Membrane Isolation and Protein Purification

The Tsr membrane isolation and protein purification were implemented in collaboration with Dr. Soojay Banerjee (Subramaniam laboratory). Specifically, the sucrose gradient, detergent solubilization, and size exclusion chromatography were all performed by Dr. Banerjee while other protocols were independently implemented.

Membrane Isolation

The buffer used for the production of isolated membrane assemblies was 25mM Tris pH 8.0 with 1mM EDTA and either 100mM or 200mM NaCl **(Section 4.3.2)**. For preparations from which Tsr was to be subsequently solubilized, the buffer used was PBS pH 7.4 with 1mM EDTA since this appeared to increase purity, though it reduced yield. All buffers were supplemented with protease inhibitors (complete ULTRA Tablets, EDTA-free from Roche Applied Science).

Pellets of induced cells were thawed, resuspended in buffer, and lysed by high pressure homogenization. Cell debris were removed by centrifugation for 45 min at 10,000 x g at 4°C. Membranes were then pelleted by ultracentrifugation for 1 hour at 100,000 x g at 4°C. The supernatent was retained to check for loss of Tsr by polyacrylamide gel electrophoresis. The pelleted membranes were resuspended and washed in buffer at 9,000 x g before being used for polyacrylamide gel electrophoresis **(Section 3.3.4)**, cryo-electron tomography **(Section 3.4.2)**, or detergent solubilization **(Section 3.3.3)**. Unused membrane suspensions were separated into single-use aliquots, flash frozen in liquid nitrogen and stored at -80°C.

Sucrose Gradient

Solutions of sucrose (in Tris pH 8.0) ranging from 30% to 70% (w/v) at 10% increments were layered sequentially (500 μL each) in an open-top, thinwall ultracentrifuge tube (Beckman Coulter) from most dense to least dense. The membrane suspension was applied to the top of the density gradient and the entire setup was

subject to ultracentrifugation for 4 hours at 100,000 x g at 4°C. Fractions were eluted by piercing the bottom of the centrifuge tube.

Detergent Solubilization

The detergents that were attempted for Tsr solubilization along with the concentrations at which they were used are listed in **Table 3.8**. Of the detergents used, only n-dodecyl-β-D-maltopyranoside (DDM) and n-decyl-β-D-maltopyranoside (DM) showed an appreciable ability to solubilize Tsr **(Sections 4.4 and 4.5.2)**.

For Tsr solubilization, membrane suspensions were pelleted at 9,000 x g and resuspended in PBS pH 7.4 with detergent, 1mM EDTA and protease inhibitors (cOmplete ULTRA Tablets, EDTA-free from Roche Applied Science). Samples were incubated overnight at 4°C with gentle shaking to promote exchange of phopholipids and detergent surrounding the protein. Subsequently, any remaining membranes and insoluble protein aggregates were removed by ultracentrifugation for 30 min at 100,000 x g at 4°C. The resulting supernatent was further purified by affinity chromatography (for *Strep*-tagged protein only) and size exclusion chromatography **(Section 3.3.3)**. The solubilization efficiency was analyzed by polyacrylamide gel electrophoresis **(Section 3.3.4)**.

Table 3.8: Concentrations of detergents in PBS/EDTA buffer used for Tsr solubilization. Each separate line represents a separate detergent or combination that was attempted for solubilization.

Detergent	Critical Micelle Concentration (CMC)	Concentration
n-Dodecyl-β-D-maltopyranoside (DDM)	0.2mM	20mM
Brij 35	0.09mM	9mM
n-Decyl-β-D-maltopyranoside (DM)	1.8mM	180mM
Fos-choline-14 (FC14)	0.12mM	12mM
Lauryl maltose neopentyl glycol (LMNG)	0.01mM	1mM
DDM & cholesterol hemisuccinate (CHS)	–	1mM DDM & 0.2mM CHS

Affinity Chromatography

Detergent solubilized Tsr was added to 2 mL of *Strep*-Tactin Superflow Plus (QIAGEN), a resin that contains an engineered form of streptavidin. The solution was mixed gently by shaking at 4°C for 1 hour and then loaded into a column. The flow through and washes with 4 mL of solubilization buffer were collected for analysis of the loss of protein by failure to bind to the column. Three elutions were performed with 500 μL of D-desthiobiotin and collected for polyacrylamide gel electrophoresis and size exclusion chromatography. The highest concentration of eluted protein was usually found in the second elution. Protein concentration was measured using absorbance at 280nm as measured by a NanoDrop 2000 Spectrophotometer (Thermo Scientific).

Size Exclusion Chromatography

Remaining impurities after solubilization (for untagged protein) or affinity purification (for *Strep*-tagged protein) were removed by size exclusion chromatography (SEC). Protein samples were injected into a Superdex 200 10/300 GL column (GE Healthcare Life Sciences) attached to an ÄKTA system, using a flow rate of 0.25 mL/min. The SEC run was performed using PBS pH 7.4 with 1mM EDTA and either 2mM DDM or 18mM DM, in order to maintain the protein in the same detergent used for solubilization. 0.5mL fractions were collected and those at the peak corresponding to Tsr (fractions B5 - B7 corresponding to elution volumes of 8.5-10mL) were retained and analyzed by polyacrylamide gel electrophoresis **(Section 3.3.4)**. Selected fractions were used immediately for the preparation of vitreous specimens **(Section 3.4.2)** or were split into 10-20 μL aliquots, flash frozen in liquid nitrogen and stored at -80°C for later use.

3.3.4 Confirming Protein Expression

The presence of a protein in whole cell lysates or purified protein was indicated by size comparison after polyacrylamide gel electrophoresis (PAGE), either under denaturing or native conditions. For purified protein, denaturing gels were also used to reveal the presence of contaminating proteins and native gels were used to confirm that the protein existed in its functional oligomeric state. The identity of the protein of interest was more conclusively established by Western blot analysis.

Denaturing Polyacrylamide Gel Electrophoresis

For determining protein expression, 1 mL of cell culture was pelleted by centrifugation at 13,000 x g and resuspended in 60 μL MilliQ water. Equal volumes of resuspended cells and 2x reducing sample buffer were mixed and the resulting samples were heat treated at 94°C for 3 min and vortexed for 1 min. In order to normalize samples, the amount of each cell sample to be loaded was calculated by comparing the OD_{600} values from the cell cultures of the given sample and the least dense sample. Prior to loading, all cell lysate samples were brought up to the same volume using 1x reducing sample buffer. Purified protein samples were mixed with 4x LDS Buffer (Invitrogen) prior to loading.

Samples were loaded on NuPAGE 4-16% Bis-Tris Gel (Invitrogen) alongside Benchmark Protein Ladder (Invitrogen) or SeeBlue Plus2 Prestained Standard (Invitrogen). Gels were run at 200V for 45-50 min in MOPS running buffer (Invitrogen). Gels were subsequently removed from their cassettes, washed 3 times in MilliQ water, stained for at least 2 hours in SimplyBlue SafeStain, and destained for at least 1 hour in MilliQ water.

Native Polyacrylamide Gel Electrophoresis

Purified protein samples were mixed with 4x NativePAGE Sample Buffer (Invitrogen) and loaded on a precast NativePAGE Novex 4-16% Bis-Tris Gel (Invitrogen) alongside NativeMark Unstained Protein Standard (Invitrogen). The anode and dark blue cathode buffers were prepared from NativePAGE 20x Running Buffer (Invitrogen), NativePAGE 20x Cathode Buffer Additive (Invitrogen), and MilliQ water as per the manufacturer's instructions. Gels were run for 105-120 minutes at 150V in an XCell *SureLock* Mini-Cell with ∼200 mL dark blue cathode buffer in the upper buffer chamber and ∼600 mL anode buffer in the lower buffer chamber. Gels were then removed from their cassettes, destained in MilliQ water for 1-2 hours and stained overnight using SimplyBlue SafeStain (Invitrogen).

Confirming Protein Identity by Western Blot Analysis

An SDS-PAGE gel was run as previously described using SeeBlue Plus2 Prestained Standard (Invitrogen) as the marker. The blotting pads and filter paper were pre-soaked in transfer buffer. The PVDF membrane was activated in methanol, rinsed in MilliQ water, and soaked in transfer buffer. The gel was removed from the cassette, wet in transfer buffer and the pre-soaked membrane was placed on the gel. Filter paper and then blotting pads were placed on either side of the gel-membrane combination. Care was taken to avoid air bubbles between the components as they would reduce conductivity and interfere with the transfer. The entire assembly was placed in the XCell II Mini-Cell Blot Module with the gel closer to the cathode and the membrane closer to the anode. The blot module was filled with transfer buffer just until the gel-membrane sandwich was submerged. The other chamber was filled with ∼650 mL of MilliQ water and the transfer was performed by applying a potential difference of 30V for 1 hour.

TBS solutions were made from using 10x TBS pH 7.4 stock solution supplied by Quality Biologicals. After the transfer, the membrane was blocked with 30 mL of TBS with 5% skim milk for 30 min. The membrane was washed 3 times in TBST

(250 μL of Tween-20 in 500 mL TBS). Tsr was labeled using an α-Tsr polyclonal antibody gifted by Dr. Peter Ames (Parkinson laboratory) at a 1:1000 dilution in TBS with 1% skim milk. After incubation with the primary antibody for 1.5-2 hours, the membrane was washed 3 times with TBST and incubated for 30 min with an anti-rabbit antibody conjugated to alkaline phosphatase (Abcam). The membrane was subsequently washed 3 times in TBST and once in TBS. The immunoblot was visualized by placing the membrane in a 10 mL solution containing a SIGMA*FAST* BCIP/NBT tablet (Sigma-Aldrich) dissolved in MilliQ water. After sufficient color developed from the reaction of alkalinne phosphate with its substrate (usually 2-5 min), the membrane was rinsed in MilliQ water and air dried.

3.4 Microscopy and Image Analysis

3.4.1 Fluorescence Microscopy

Preparing Agarose Pad Slides

Slides were prepared by adding a thin layer of 0.8% agarose in succinate medium between two coverslips. The adjacent coverslips and excess agarose were subsequently removed to leave an agarose pad. 2 μL of cells as prepared in **(Section 3.3.2)** were placed on the agarose before addition of a cover slip.

Microscopy

A Nikon Eclipse Ti microscope was used to acquire fluorescence and phase contrast images of cells. Images were collected using the 100x oil objective (Nikon) with exposure times of 40 ms and 500 ms for the phase and YFP channels, respectively. The microscope was controlled using NIH Elements software (Nikon) and the images were captured using an Andor iXON CCD camera.

Image Analysis

Image analysis was performed using MicrobeTracker (Emonett and Jacobs-Wagner laboratories), a MATLAB-based software package that segments phase contrast images and analyzes the fluorescent signal within the segmented cells [187]. Automatic segmentation of cells is performed by thresholding and edge detection, combined with applying a contour which describes the model shape of an individual cell. The constraints for cell shape are based on user defined parameters such as the typical length, width, area, and curvature of cells of the bacterial species under study at the given magnification. Parameters for this study were based on parameters previously used in the lab for *R. sphaeroides* and were empirically optimized in the MicrobeTracker software. Detected objects that do not fit the parameters for cell shape and area are excluded from further analysis.

Manual inspection of the automatically segmented images was used to remove any incorrectly identified cells. Once cell contours were generated and manually inspected, an additional feature of MicrobeTracker, SpotFinderM, was used to manually identify fluorescence spots (indicative of clustered protein) within contours.

3.4.2 Cryo-electron Tomography

Preparation of Vitreous Specimens

Isolated membrane preparations were diluted in their buffer such that their total protein concentration was approximately 1.5 mg/mL. Whole cell samples were collected at mid-log phase (OD_{600} of 0.4-0.6) and used directly for study. Prior to sample preparation, 5 μL of protein-A gold conjugate fiducial markers (Utrecht University) were added to 20 μL of the sample to be studied. This volume was empirically determined in order to have 10-20 fiducials per imaging field at a nominal magnification of 34,000x.

A Mark IV Vitrobot (FEI Company) was used to prepare vitrified specimens for

cryo-electron tomography. The Vitrobot was operated with its chamber at 22°C and 90% humidity. Prior to the preparation of samples, the coolant bowl was filled with liquid nitrogen an allowed to equilibrate. Gaseous ethane was condensed into the central cup, which was thermally insulated by the surrounding liquid nitrogen. The coolant bowl was then loaded on the Vitrobot. Samples were prepared using 200 mesh Quantifoil MultiA holey carbon grids (Quantifoil Micro Tools). Prior to use, grids were cleaned using hydrogen-oxygen plasma for 7 seconds in a Solarus 950 unit (Gatan). Each grid was held with forceps and 2.5-3 μL of specimen solution was placed on the the cleaned surface of the grid. The forceps were loaded on the Vitrobot, which automatically blotted away excess liquid prior to plunging the grid into the nitrogen cooled liquid ethane which was maintained at approximately -180°C. Frozen grids were then transferred to cryo-EM storage boxes (Pacific Grid Tech) and stored in liquid nitrogen.

Microscopy

Grids were maintained at liquid nitrogen temperatures while they were placed in cartridges and loaded into the cryotransfer system of an FEI Polara G2 microscope. The field emission gun of the microscope was operated at 200 kV and images were collected on a 2K x 2K post-energy filter CCD (Gatan). A series of low-dose projection images was acquired over the angular range of $\pm$ 60° at 2° intervals to yield a total of 61 images per tilt series. Each image was acquired at a nominal magnification of 34,000x (4.1Å effective pixel size at the sample plane) with an average underfocus value of 2.5 µm and an electron dose of 1-2 $e^-/Å^2$.

Subtomogram Image Processing

Automated fiducial detection and image alignment were performed on collected tilt series using RAPTOR (Horowitz lab) as implemented in IMOD (Boulder Laboratory for 3-D Electron Microscopy) [188, 189]. The resulting aligned tilt series were reconstructed into a 3D volumes (tomograms) in IMOD using R-weighted back projection.

Volumes containing individual membrane assemblies were displayed in IMOD and were manually extracted from the reconstructed tomograms. From these regions, subtomograms were generated, aligned and averaged using algorithms developed and implemented by Dr. Xiongwu Wu (NIH Laboratory of Computational Biology). For each experimental condition, an initial model was generated as described in [145]. Briefly, subtomograms of a small repeating unit including a receptor and its six immediate neighbors were generated and subjected to several rounds of alignment. The aligned subvolumes were projected along the z axis to create 2D projection images that were subsequently iteratively classified using the EMAN software package [190]. The subvolumes corresponding to the projection images in the dominant class were aligned and averaged to create a single density map that served as an initial model. Individual subvolumes were then compared to this initial model and assigned a correlation coefficient as described in [145]. This was performed for each experimental condition, with individual subvolumes from a given sample being compared to the initial model generated for the corresponding experimental condition. The distribution of correlation coefficients was analyzed for each of the two experimental conditions and one major cluster was identified in each case. The subvolumes that corresponded to the correlation coefficients present in the major cluster were subjected to further alignment and averaging until the orientational assignments and the average density maps converged. A cylindrical mask with a diameter of 25nm and a height of 100nm was used during all steps of alignment and averaging.

3.4.3 Cryo-sectioning and Immuno-gold Labeling

Cells were grown for cryo-sectioning and immuno-gold labeling as described in **(Section 3.3.2)**. Once cells reached an OD_{700} of 0.3-0.4, they were fixed by the addition of an equal volume of 8% paraformaldehyde in 2x succinate media (pH adjusted to 7.4). After incubation with paraformaldehyde for 10-20 min, cells were pelleted by centrifugation at 3,210 x g for 10 min at 25°C and were stored at room temperature. Cryo-sections were prepared as described in [191] by Lisa Hartnell (Subramaniam laboratory). Briefly, the pellets were then placed in warm gelatin, cooled on ice, and cut into cubes. 15% polyvinyl pyrrolidone and sucrose at a final concentration of

1.7 M were used to cryoprotect these cubes. These cubes were subsequently cut into 70-100 nm sections at -130°C. Sections were picked up on a methyl cellulose/sucrose droplet and placed on a carbon coated copper grids.

Immuno-labeling and subsequent imaging was performed by Lisa Hartnell (Subramaniam laboratory). For labeling the HA epitope, grids were incubated with a rat anti-HA antibody (Roche) followed by a rabbit anti-rat antibody (Jackson Laboratories). For labeling the YFP epitope, grids were incubated with rabbit anti-GFP (AbCam). The antibody distribution was then visualized with 10-nm protein-A gold (Utrecht University), which recognizes the Fc portion of the rabbit antibody. After labeling, the sections were stained with 2% methylcellulose with uranyl acetate as described in [192] in order to enhance contrast. Projection images of grids with stained sections were collected using a 4K x 4K CCD camera mounted on a Tecnai 12 electron microscope (FEI Company). The microscope was equipped with a LaB6 filament and was operated with an accelerating voltage of 120 kV.

Chapter 4

Optimization of *E. coli* Serine Chemoreceptor (Tsr) Expression and Purification

For any imaging or structural study, a suitable sample must first be prepared. In the case of the *E. coli* serine chemoreceptor, Tsr, it was initially unclear whether the best method for structural study would be cryo-electron tomography of Tsr in its native membrane environment or cryo-electron microscopy of the detergent solubilized receptor. As a result, the protocol for producing Tsr was modularly optimized to provide the ability to study three types of samples: whole cells overexpressing Tsr, purified membranes containing Tsr, and detergent solubilized Tsr. This chapter summarizes the efforts to optimize the production of these samples for subsequent structural studies.

4.1 Overview of Protocol

The protocol involved three overarching procedures that were optimized: *a*) induction of Tsr expression; *b*) cell lysis and membrane fractionation to isolate Tsr containing membranes; and *c*) solubilization and purification of intact Tsr protein. While the endpoint of each procedure provided one of the types of samples for fur-

ther study, the earlier steps also served as intermediates for the other sample types. Namely, cells overexpressing Tsr were themselves potential subjects of study as well prerequisites for producing membranes containing Tsr. These, in turn, could be directly studied as well as used to produce detergent solubilized receptor. Thus, the protocol is a progression used to produce detergent solubilized Tsr with intermediate stopping points for the other types of samples. The specific steps and parameters

Table 4.1: Procedures optimized for the production of Tsr samples.
The three procedures are listed sequentially along with each of their specific steps and the parameters that were optimized.

Procedure	**Basic Steps**	**Parameters Varied and Tested**
Induction of Protein Expression	1. Growth of *E. coli* cells in media 2. Addition of IPTG to growing cells to activate promoter for protein expression 3. Harvest cells after several hours of growth with IPTG	• OD_{600} at time of induction • IPTG concentration • Growth temperature before and after induction • Length of time after induction
Cell Lysis and Membrane Fractionation	1. High pressure homogenization of resuspended pellet to lyse cells 2. Centrifugation to remove cell debris and isolate membranes	• Buffer composition • Inclusion of sucrose gradient
Solubilization and Purification	1. Incubation of isolated membranes with detergent to remove protein from membranes 2. Affinity purification (if protein is tagged) 3. Size exclusion chromatography to isolate monodisperse species	• Solubilizing detergent • Presence or absence of affinity tag

optimized for each of the three procedures within the protocol are summarized in **Table 4.1** and elaborated in the subsequent sections.

4.2 Induction of Tsr Expression

Tsr expression was performed in the *E.coli* host strain HCB721, which lacks the endogenous transmembrane chemoreceptors (Tar, Tsr, Trg, and Tap), cytoplasmic chemoreceptor array proteins (CheA and CheW), and chemotaxis adaptation proteins (CheB and CheR) [182]. Tsr was expressed in HCB721 from the construct pHSe5.tsr_{QEQE} which encodes for the wild-type form of Tsr under the control of an IPTG-inducible promoter [185]. The induction conditions previously used by Khursigara *et al.* for imaging studies of this strain and construct (growth at 27°C, induction with 1mM IPTG at an OD_{600} of 0.1 for 3-4 h) provided a starting point for the parameters tested for this study [145]. However, the conditions selected by Khusigara *et al.* were optimized for whole cell imaging at mid-log phase ($OD_{600} \approx 0.5$) and sought to have high Tsr expression at that growth phase. This differs from general strategies for protein overexpression in *E. coli*, where induction is generally performed at mid-log phase and allowed to proceed for several hours until the cells reach stationary phase [193]. Therefore, conditions based on both previous whole cell studies and established protein overexpression protocols were tested with the intention of increasing the amount of Tsr containing membranes that could be isolated. Temperature schemes were also varied such that both constant growth at a given temperature and reduction of temperature at the time of induction were examined **(Table 4.2)**.

For each set of conditions tested, whole cell lysates were evaluated at different time points using SDS-PAGE analysis to find the optimum length of incubation time after induction. For both sets of conditions where cells were induced at an OD_{600} of ∼0.1 (highlighted in cyan in **Table 4.2**), Tsr expression was highest 7-9 hours after induction (data not shown). Relative to those conditions, Tsr expression was

Table 4.2: Induction conditions tested for HCB721/pHSe5.tsr_{QEQE}.
The first two sets of conditions (in cyan) utilized early induction and constant temperature before and after induction based on the whole cell experiments in [145]. The second two conditions (in yellow) followed general strategies for protein overexpression in *E. coli* with induction later in the growth cycle and temperatures being decreased at the time of induction [193].

OD_{600} at time of induction	**IPTG concentration for induction**	**Growth temperature before induction**	**Growth temperature after induction**
~0.1	1mM	27°C	27°C
~0.1	1mM	32°C	32°C
~0.8	0.5mM	37°C	25°C
~0.8	0.5mM	37°C	30°C

significantly increased by inducing at an OD_{600} of ~0.8 and reducing the temperature at the time of induction **(Figure 4.1)**. Under these conditions (highlighted in yellow in **Table 4.2**), a larger temperature decrease (37°C to 25°C) at the time of induction slightly increased the level of Tsr relative to other proteins at a given cell density (data not shown). Thus, this set of conditions was tested in further detail to determine the optimum incubation time after induction. Different samplings of time points for two inductions showed that maximal Tsr expression is achieved 10.5 - 14.5 hours after induction **(Figure 4.1)**. Thus, the optimal conditions for the production of Tsr containing membranes and subsequent detergent solubilization were determined to be initial growth at 37°C with induction at OD_{600} of ~0.8 followed by incubation at 25°C for 10.5 - 14.5 hours.

The aforementioned conditions were optimized in lysogeny broth (LB). However, later inductions to produce Tsr containing membranes and solubilized protein were done in super broth in order to raise yield by increasing cell density for a given culture volume. When performing the inductions in super broth, the optimized parameters remained the same with the exception that induction was performed at an OD_{600} of ~1 since the richer media allowed for a longer logarithmic growth phase. For each induction, sufficient Tsr expression was confirmed using SDS-PAGE analysis of whole cell lysates prior to membrane isolation.

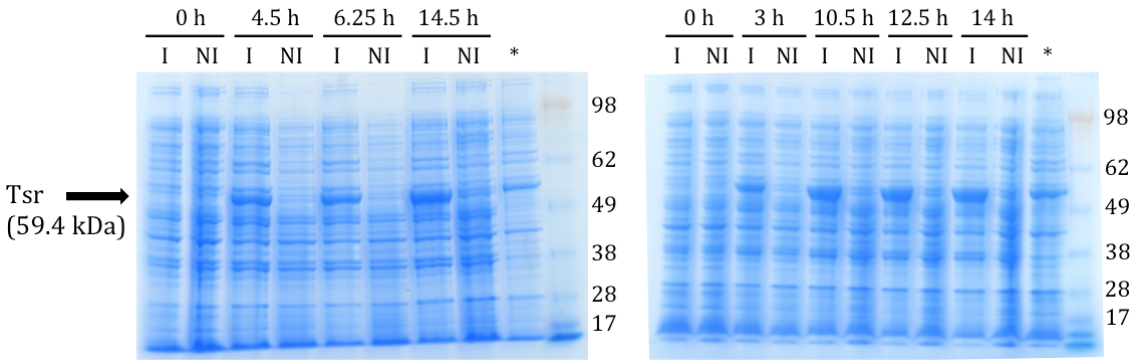


Figure 4.1: SDS-PAGE analysis of whole cell lysates after Tsr induction. Analysis from two separate induction experiments done to evaluate induction time. Inductions for these experiments were performed at an OD_{600} of ∼0.8 with 0.5mM IPTG with a post-induction growth at 25°C. At each time point, induced (I) and uninduced (NI) cells were compared. Time points were designated based on the length of incubation from the time of induction. The starred columns were cell lysates from the 8h time point of an induction at an OD_{600} of ∼0.1 with 1mM IPTG in cells that were grown at a constant temperature of 27°C. The theoretical molecular weight of Tsr (59.4 kDa) is indicated in the figure. All lysates were normalized by cell density (as determined by OD_{600}) prior to loading.

4.3 Membrane Isolation

Induced cells were harvested by centrifugation and cell pellets were frozen at -80°C. After thawing, pellets were resuspended and lysed by high pressure homogenization as described in **Section 3.3.3**. Cell debris were removed by centrifugation and membranes were pelleted by spinning at 100,000 x *g*. Membranes were then washed and resuspended in Tris buffer with EDTA and NaCl. Subsequently, different methods were attempted to increase the purity and yield of Tsr in the membrane suspension.

4.3.1 Sucrose Gradient

In order to remove any impurities resulting from outer membrane proteins, components of the membrane suspension were separated using a sucrose density gradient

(Section 3.3.3) [194]. The membrane suspension was applied to a gradient containing sucrose solutions ranging from 30% to 70% at 10% increments. SDS-PAGE analysis of the fractions eluted after ultracentrifugation revealed that the membranes containing Tsr primarily eluted in the fraction containing 60% sucrose **(Figure 4.2)**. Moreover, it appeared that the protein composition in the 60% sucrose fraction was approximately the same as the initial membrane suspension and no fractionation was achieved. Therefore, we extrapolated that there was no significant contamination from outer membrane proteins (though no outer membrane control was run). As a result of the negligible gains in purity of the Tsr containing membranes from the sucrose gradient, this procedure was omitted from subsequent implementations of the protocol.

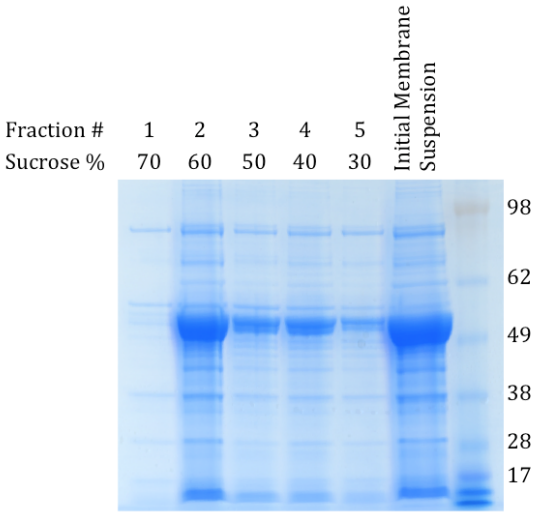


Figure 4.2: SDS-PAGE analysis of sucrose gradient fractions.
Tsr containing membranes eluted primarily in fraction 2 (60% sucrose), but seemed to have approximately the same composition as the initial membrane suspension.

4.3.2 Buffer Composition

The amount of sodium chloride (NaCl) in the buffer was varied in order to determine the effect of the buffer's ionic strength on the solubility of the membranes. Equal amounts of pelleted membranes were resuspended in parallel in buffer containing either 100mM or 200mM NaCl. These were subjected to additional centrifugation at

9,000 x g and the resulting pellets and supernatents were analyzed by SDS-PAGE. The size of the pellet was visibly smaller for the membranes resuspended in the buffer with higher salt concentration, indicating membrane solubility increased with higher ionic strength. However, the SDS-PAGE analysis revealed that the Tsr content of membranes in both buffers was so high that the increased solubility with 200mM NaCl was likely to be inconsequential **(Figure 4.3)**. Furthermore, higher salt concentration may reduce contrast in vitrified specimens by increasing density of the solution [195]. Even though such differences in contrast are likely minimal between the two concentrations in question, membranes suspended in the buffer with lower salt concentration were selected for use in further studies.

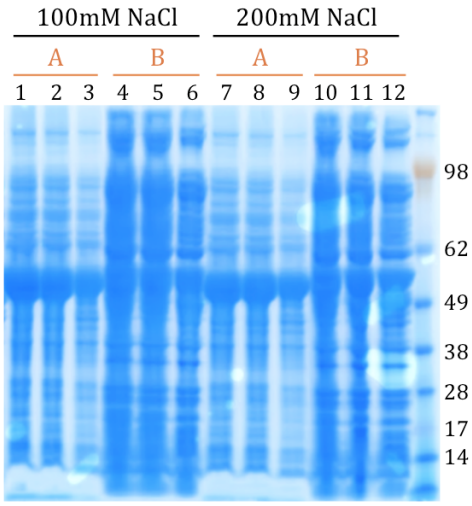


Figure 4.3: SDS-PAGE analysis of membranes in buffers of different ionic strengths.
Comparison of the resuspended pellet (A) and supernatent (B) after isolated membranes were washed in buffer (25mM Tris pH 8.0, 1mM EDTA) containing either 100mM or 200mM NaCl. Each set of 3 adjacent lanes contains decreasing amounts of the given sample. On the gel, there were minimal observable differences between the content of the pellet and supernatent at the two salt concentrations.

4.4 Solubilization and Purification of Untagged Tsr

The isolated membranes containing Tsr were solubilized overnight in n-dodecyl-β-D-maltopyranoside (DDM) to extract Tsr from the membranes. DDM was selected because of its low critical micellar concentration (CMC) relative to other non-ionic detergents as well as its prior use in solubilizing membrane proteins that retained their functionality [196]. After solubilization, ultracentrifugation at 100,000 x g was used to remove remaining membranes and insoluble proteins. Native-PAGE analysis confirmed the presence of Tsr trimers of dimers (356.4kDa) in the resulting supernatent **(Figure 4.4)**. Though some Tsr appeared to be lost in the pellet (as seen in the Native gel), most remained in the supernatent. Size exclusion chromatography (SEC) of the DDM-solubilized protein was subsequently used to isolate a more pure species. Peak fractions were analyzed to confirm the presence of Tsr and observe the relative proportions of contaminating proteins. Based on the SDS-PAGE analysis, the B7 fraction appeared to have the highest proportion of Tsr relative to total protein content **(Figure 4.4)**.

In this protocol, the only method for selection of the desired species after solubilization from the membranes was SEC. In spite of this, the resulting sample was predominantly Tsr. However, the existing heterogeneity would likely complicate later image processing steps by requiring more extensive classification in order to separate contaminating species [50]. In an effort to minimize the heterogeneity of the sample, selective purification of Tsr with an affinity tag was pursued.

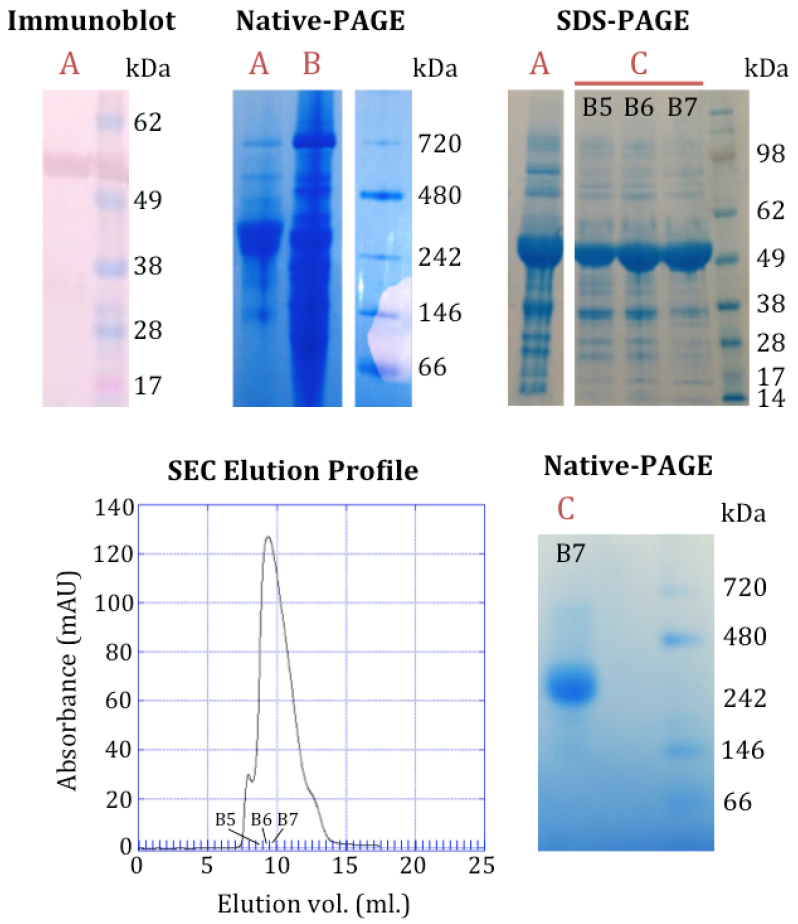


Figure 4.4: Purification of untagged Tsr.
After overnight solubilization in 20mM DDM and ultracentrifugation, the majority of Tsr was retained in the supernatent (A), though some was lost in the pellet (B). Immunoblotting with α-Tsr polyclonal serum confirmed the identity of Tsr and Native-PAGE revealed a species consistent with the molecular weight of a Tsr trimer of dimers (theoretical molecular weight of 356.4kDa). The elution profile from size exclusion chromatography showed large amounts of protein but a somewhat broad peak close to the void volume. Peak fractions (B5-B7 corresponding to elution volumes of 8.5-10mL in 0.5mL increments) after SEC (C) showed that the molecular weight of native Tsr is maintained and contaminants are reduced. For clarity, extraneous lanes have been removed and relevant lanes have been placed adjacent to the other lanes from the same gel.

4.5 Solubilization and Purification of *Strep*-Tagged Tsr

4.5.1 Creation of *Strep*-Tagged Tsr Construct

The primary considerations for choice of an affinity tag for Tsr were high specificity and mild elution conditions. The 8 amino acid *Strep*-tag satisfied these conditions and had many other desirable attributes such as being biochemically inert, having a balanced amino acid composition, and maintaining activity in the presence of mild detergents [197, 198]. In order to prevent disruption of membrane localization, the *Strep*-tag was placed on the C-terminal end of Tsr. This also ensured that it would be exposed since the C-terminal end of Tsr contains the exposed tether for the chemotaxis adaptation enzymes [199].

Though the *Step*-tag is unlikely to require cleavage due to its small size and inert nature, it was attached to Tsr with a thrombin cleavage site followed by a serine-alanine linker. The *Strep*-tag was inserted into the existing pHSe5.tsr_{QEQE} construct in order to minimize changes to the induction and expression. Since Tsr is not flanked by restriction sites in this vector, the Tsr-*Strep*-tag junction was first made in a commercial vector. To accomplish this, Tsr (with an added thrombin cleavage site) and the *Strep*-tag were sequentially inserted into the multiple cloning site of the vector pET28b using standard PCR, restriction digestion, and ligation **(Table 4.3)**. Then, common restriction sites in the middle of both Tsr and neighboring gene YjiZ were located in pHSe5.tsr_{QEQE} and used to excise the fragment in between. The excised fragment was replaced by ligation with a fragment generated by overlap extension PCR which contained the end of the Tsr gene, the *Strep*-tag, and the start of the YjiZ gene **(Table 4.3)**. The resulting plasmid, called pHSe5.*tsr-Strep* was sequenced to confirm the inclusion of the cleavage site and *Strep*-tag.

Table 4.3: Succession of steps used for construction of pHSe5.*tsr-Strep*. Each line of the table represents a single PCR reaction using a primer pair and, if applicable, insertion into a suitable vector by restriction digestion and ligation. Sequences for the listed primers can be found in **Appendix B**.

Forward primer	Reverse primer	Template	Restriction Enzymes	Vector	Product
mv001	mv002	none	BamHI/ HindIII	pET28b	pET28b-*Strep*
mv005	mv007	pHse5.tsr_{QEQE}	NcoI/ BamHI	pET28b-*Strep*	pET28b-*tsr-Strep*
mv008	mv009	pET28b-*tsr-Strep*	N/A	N/A	fragA
mv010	mv011	pHse5.tsr_{QEQE}	N/A	N/A	fragB
mv008	mv011	fragA + fragB	NdeI/ HindIII	pHse5.tsr_{QEQE}	pHSe5.*tsr-Strep*

4.5.2 Solubilization and Purification

Expression of Tsr from the pHSe5.*tsr-Step* construct was done in HCB721 using the optimized conditions discussed in **Section 4.2**. Membranes were isolated as was described in **Section 4.3** and incubated overnight at 4°C in 20mM DDM to solubilize Tsr. The resulting solution was bound to a *Strep*-Tactin column and eluted with D-desthiobiotin. SDS-PAGE analysis showed that two rounds of elution both produced very pure Tsr, though some Tsr is lost in the flow through **(Figure 4.5)**. SEC gave a tight peak and the peak fraction (B6) appeared as an intact trimer of dimers in Native-PAGE analysis **(Figure 4.5)**. Although Tsr appeared to be intact in solution based on these biochemical assays, it was unstable in the thin layer of vitreous ice on EM grids, perhaps due to interactions with the air-water interface during blotting and freezing. Solubilization and purification in other detergents was pursued in an effort to see if they assist in stabilizing the protein on cryo-EM grids.

Four other detergents were screened for their ability to solubilize Tsr from purified membranes at a concentration 100 times their critical micelle concentration (CMC). In addition, a combination of the detergents DDM and cholesterol hemisuccinate (CHS) was also tested. Only one of these five detergents, n-Decyl-β-D-Maltopyranoside (DM), was able to solubilize an appreciable amount of Tsr **(Figure**

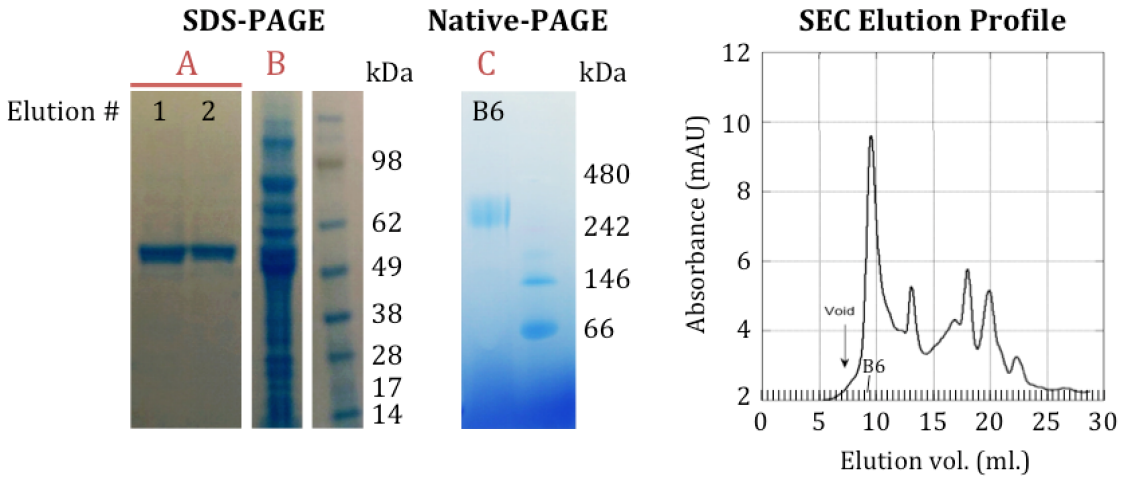


Figure 4.5: Purification of *Strep*-tagged Tsr in DDM.
After overnight solubilization in 20mM DDM, the resulting solution was affinity purified using a *Strep*-Tactin column. Two successive elutions with D-desthiobiotin (A) contained very pure Tsr, but some Tsr did not bind to the column and instead was present in the flow through (B) from the column. Native-PAGE showed that the peak fraction (B6, corresponding to an elution volume of 9-9.5mL) after SEC (C) maintains the molecular weight expected of a Tsr trimer of dimers (theoretical molecular weight of 356.4kDa). For clarity, extraneous lanes have been removed and relevant lanes have been placed adjacent to the other lanes from the same gel.

4.6). Therefore, purification of Tsr in solubilized in DM was performed using *Strep*-Tactin affinity purification and SEC. Though the Tsr purified in DM ran at the correct molecular weight for a trimer of dimers in Native-PAGE analysis **(Figure 4.7)**, it appeared to dissociate on cryo-EM grids.

Amphipathic polymers, known as amphipols, are an established alternative to detergents for maintaining membrane proteins in solution [200]. Efforts to exchange detergent to the most commonly used amphipol, A8-35, did not appear to stabilize the protein on cryo-EM grids. Addition of *E. coli* lipids not only failed to stabilize the protein, but also added an undesirable background to cryo-EM specimens. Attempts to use different types of EM grids including those that are continuous carbon-coated or functionalized with a self assembled monolayer were also unsuccessful [50, 201].

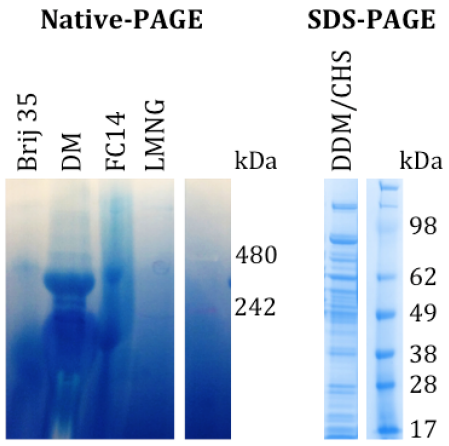


Figure 4.6: Detergent screen for solubilizing Tsr.
Tsr was solubilized overnight in 9mM Brij 35, 180mM n-decyl-β-D-maltopyranoside (DM), 12mM Fos-choline-14 (FC14), 1mM Lauryl maltose neopentyl glycol (LMNG), and 1mM DDM with 0.2mM cholesterol hemisuccinate (DDM/CHS). After insoluble components of the suspension were removed by centrifugation, gel electrophoresis showed that only DM extracted Tsr from membranes. For clarity, extraneous lanes have been removed and relevant lanes have been placed adjacent to the other lanes from the same gel.

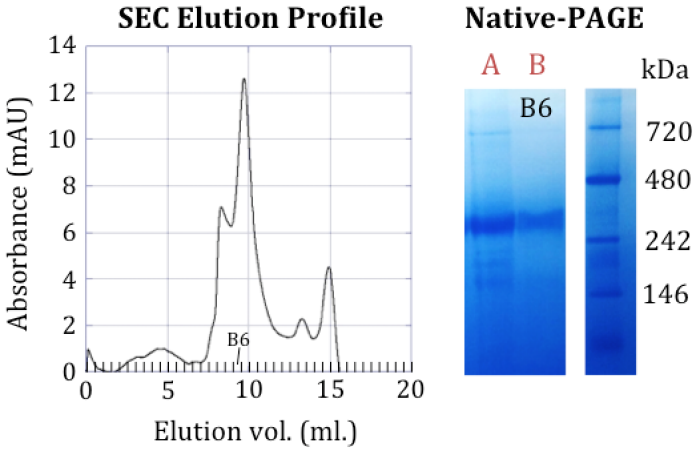


Figure 4.7: Purification of *Strep*-tagged Tsr in DM.
Tsr was solubilized in DM and analyzed by Native-PAGE after affinity purification (A) and SEC (B). In both the elution from the *Strep*-Tactin column and the peak fraction (B6, corresponding to an elution volume of 9-9.5mL) from SEC, Tsr appeared as an intact timer of dimers. For clarity, extraneous lanes have been removed and relevant lanes have been placed adjacent to the other lanes from the same gel.

4.6 Discussion

This chapter involved the optimization of three procedures, namely expression, membrane isolation, and protein purification to create samples for subsequent studies using cryo-electron microscopy or tomography. Specifically, Tsr expression was increased relative to the previously published protocol by inducing with IPTG later in the growth cycle, reducing temperature at the time of induction, and allowing a longer incubation after induction. The procedure established in this study was optimal for increasing yield in subsequent isolation of Tsr containing membranes and purified protein. However, for whole cell studies, where imaging is best done at mid-log phase, the previously published protocol continues to be the optimal choice [145].

In contrast to previous studies where Tsr containing membranes have been isolated and visualized, the procedure developed and optimized in this study does not require a sucrose gradient or detergent treatment of the membranes [185]. Instead, the membranes are kept in aqueous buffer throughout the process, thereby maintaining conditions relatively close to a physiological state. While the previously published protocol was sufficient for visualization of membrane assemblies, higher resolution structural studies are likely to have increased relevance if membrane isolation is done under conditions that are more closely aligned with native conditions.

Though untagged Tsr could be purified from isolated membranes using DDM solubilization and size exclusion chromatography, affinity purification with *Strep*-Tactin improved the purity of the resulting sample. In both DDM and DM, Tsr appeared to be an intact trimer-of-dimers in solution but was fragmented on cryo-EM grids. The origin of this phenomenon is uncertain, but it may be a result of exposure to the air-water interface during sample preparation. One potential approach for stabilizing Tsr for cryo-EM is the Gradient Fixation (GraFix) protocol during which macromolecular assemblies under mild chemical cross-linking during density gradient ultracentrifugation [202]. However, this method does involve fixation of the protein which limits the physiological relevance of the results.

Alternately, Tsr could be reconstituted into nanodiscs in order to be maintained in a lipid environment prior to cryo-EM sample preparation. However, previous electron microscopy of negatively stained samples of Tsr in nanodiscs revealed that Tsr was present as homodimers rather than trimers-of-dimers [203]. Nonetheless, It is possible that differences in the protein purification protocol between the previous nanodisc study and this study may yield a different result and produce nanodiscs containing functional receptors. Further investigation is necessary in order to determine whether the loss of trimerization of Tsr in nanodiscs is inherent to the nanodisc preparation protocol or is specific to the methodology employed by Li *et al* in their study [203]. Regardless, unpublished data from the Subramaniam lab suggests that there may be technical challenges in determining high resolution structures by cryo-EM of proteins embedded in nanodiscs as a result of size variability inherent in an individual preparation of nanodiscs.

The solubilized Tsr could be reconstituted in membranes, though this would be a more circuitous route to obtaining Tsr in membranes than the method for isolating membrane assemblies described in this section. However, reconstitution may offer the advantage of finer control over the composition of the assemblies by first eliminating any contaminating proteins before placing the receptors into membranes. This may be advantageous in minimizing sample heterogeneity, which may improve the possibility of reaching higher resolutions during structure determination by cryo-ET.

Although the detergent solubilized Tsr cannot be used for cryo-EM single particle studies, the protocol developed in this study for Tsr purification could prove useful for other biophysical studies which are performed in solution. For instance, it could be combined with site-directed spin labeling and electron paramagnetic resonance (EPR) spectroscopy. Thus far, such studies have only been performed on intact chemoreceptor homodimers and not on the full trimer of dimers [204]. Prior to pursuing further studies, it would be prudent to assess the functionality of the Tsr purified by the method outlined in this study. This can be accomplished by an

in vitro coupling assay using radiolabeled phospho-CheA in conjunction with the purified Tsr to assess phosphotransfer to CheY [205].

In summary, the protocol for producing Tsr was optimized in multiple ways in order to produce three types of samples that could be used for structural studies. Though the detergent solubilized Tsr proved to be unsuitable for subsequent cryo-EM studies, isolated membranes and whole cells overexpressing Tsr were promising candidates for determination of Tsr structure in the native membrane by cryo-ET.

Chapter 5

Structure of *E. coli* Tsr

Structural studies of Tsr in the presence and absence of its ligand, serine, were pursued in both whole cells overexpressing Tsr and isolated membrane assemblies of Tsr. The samples used for these studies contained only Tsr without the cytoplasmic proteins CheA and CheW in an effort to validate the methods and reduce heterogeneity in the sample. If successful, the approaches used could be applied to native arrays including the cytoplasmic proteins in order to get more physiologically relevant information.

The lack of the CheA and CheW in this study calls for consideration of the role of these proteins in forming and maintaining chemoreceptor array structures. Previous fluorescence microscopy studies in *E. coli* have shown that chemoreceptors natively form compact clusters located at the cell poles. In cells where *cheA* and *cheW* have been knocked out, chemoreceptors still localize to the poles but are generally less tightly clustered [168]. Wild-type clustering can largely be restored in these cells by the expression of CheW, though CheA expression has little effect [170]. The deletion of polar CheA and CheW proteins (from *che*Op_2) in *R. sphaeroides* results in a more drastic change in localization of chemoreceptors. In this case, chemoreceptors are distributed throughout the membrane though they are more concentrated in the polar regions [135]. As is the case in *E. coli*, wild-type clustering in *R. sphaeroides* is restored by the expression of the polar CheW proteins, CheW_2 and CheW_3 [173].

Electron microscopy studies have corroborated the diffuse polar localization of Tsr chemoreceptors when they are expressed at physiological levels in the absence of CheA and CheW [163]. However, overexpression of Tsr relative to the levels of CheA and CheW or in their absence have been shown to form membrane invaginations with receptor-receptor interactions replacing interactions between the receptors and the cytoplasmic array proteins [163, 145]. In both native arrays with CheA and CheW as well as in non-functional "zipper-like" assemblies stabilized by receptor-receptor interactions, the chemoreceptor trimers of dimers form an extended hexagonal lattice in the direction perpendicular to the long axis of the receptors [164, 169]. The plane of hexagonal packing is at the membrane-distal ends of the receptors, either at the point of interactions with the CheA-CheW base layer in native arrays or the region of receptor-receptor interactions in zipper-like assemblies [165, 169]. Although the overall arrangement is conserved in both types of assemblies, the receptors pack more closely in the absence of CheA and CheW, with 9 nm spacing between the centers of adjacent hexagons rather than 12 nm in native arrays [164, 169].

5.1 Different Sample Preparation Strategies for Cryo-ET

The *in situ* nature of these studies made cryo-ET an optimal choice. As was described in **Section 1.2**, thicker samples scatter more electrons than their thinner counterparts during TEM. As a result, a higher fraction of electrons are widely or inelastically scattered and removed by the objective aperture or energy filter, respectively. Therefore, fewer transmitted and elastically scattered electrons are detected. Thus, samples thicker than approximately 1 µm are prohibitively thick for cryo-TEM and thinner samples allow macromolecular complexes to be better resolved. In the case of cryo-ET, the need to minimize sample thickness is more acute since tilting the sample effectively increases the amount of matter in a given area through which an electron beam must pass. Consequently, minimizing the thickness of a sample is an important factor in increasing resolution. For *in situ* studies of Tsr in whole

cells, two ways were attempted to reduce the sample thickness: promoting cell lysis on the EM grid by freeze-thaw and utilizing a thin *E. coli* strain. The isolated membranes containing Tsr assemblies described in the previous chapter also provided an alternative for studying Tsr in its native membrane environment without the loss of signal from the presence of other cellular components.

5.1.1 Cell Lysis on the Grid

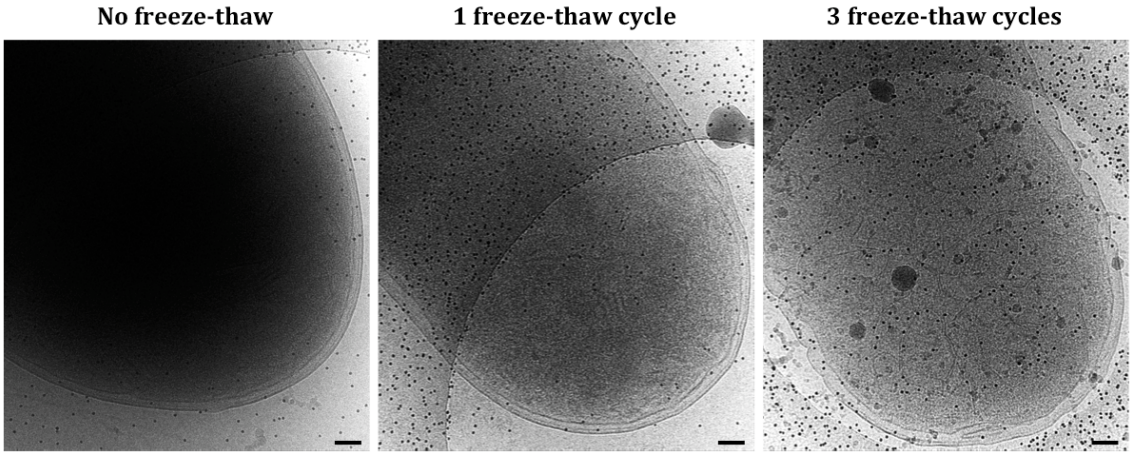


Figure 5.1: Cell lysis on grids induced by freeze-thaw.
Projection images of *E. coli* cells on EM grids demonstrating the loss of cell density that results from one or more cycles of freeze-thaw. Images were taken approximately 9.5 µm underfocus. Scale bars are 100 nm.

A number of previous cryo-ET studies of chemoreceptor arrays in whole *E. coli* cells used incubation with penicillin prior to EM grid preparation in order to partially lyse and flatten cells [164, 165, 206]. Lysis by this mechanism was not tractable in the system used for this study because of the ampicillin resistance conferred by the plasmid used for Tsr expression [185]. As an alternative cell lysis on the grid was attempted by cycles of freeze-thaw after the bacterial cell suspension was applied to the grid. This was accomplished by alternating exposing the grid to liquid nitrogen vapors and bringing it back to room temperature. A single cycle of freeze-thaw demonstrated considerable reduction of the intracellular cell density. Further cycles of freeze-thaw did not significantly reduce cell density, but noticeably disrupted cel-

lular membranes **(Figure 5.1)**. Since proteins are also disrupted by freeze-thaw cycles, this method was not implemented because it had the potential to damage the receptors that were the subject of this study.

5.1.2 Thin *E. coli* cells

Another alternative was to use an *E. coli* mutant strain that was thinner. The strain B/r H266 has been used in prior cryo-ET studies and is reported to have a diameter of approximately 0.6 µm when grown in M9 minimal media [207]. As a B strain derivative, the B/r H266 strain does not have a flagellum and lacks chemoreceptor arrays [208, 209]. The resulting non-motile phenotype was confirmed using soft agar swim plates. Since B/r H266 does not express endogenous chemotaxis proteins, it provides a clean background for the study of assemblies made solely from Tsr.

B/r H266 was transformed with the pHSe5.tsr_{QEQE} construct and Tsr expression was induced during growth in minimal media as described in **Section 3.3.1**. SDS-PAGE analysis of whole cell lysates confirmed the expression of Tsr **(Figure 5.2)**. Cells imaged at mid-log phase revealed receptors in zipper-like assemblies and hexagonally packed crystalline patches that are characteristic of cells overexpressing Tsr **(Figure 5.3)**. Interestingly, the existence of zipper-like assemblies did not result in membrane invaginations as has been previously observed in Tsr overexpressing cells [163]. Instead, these assemblies were observed when receptors in the cytoplasmic membrane interacted with receptors in an intracellular membrane that spans large regions of the cell pole. Both the cytoplasmic membrane invaginations and creation of intracellular membranes seem to stabilize receptor overexpression by providing membranes in close proximity such that receptors can interact at their membrane-distal tips. The reasons why particular cells adopt one technique or another in order to accommodate Tsr overexpression remain unclear, but the lack of membrane invaginations in B/r H266 cells may be a result of the reduced cell thickness of this strain.

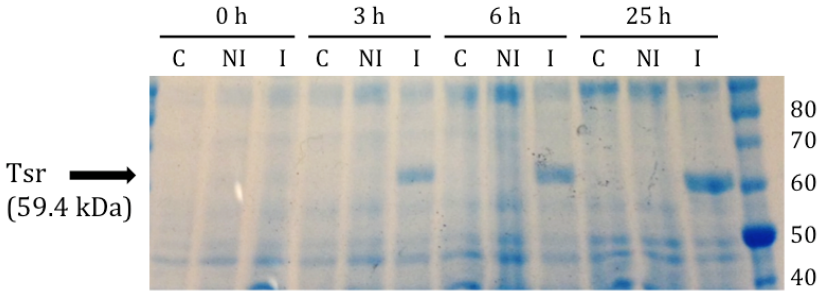


Figure 5.2: SDS-PAGE analysis of Tsr expression in B/r H266.
Analysis of whole cell lysates taken from B/r H266 cells without pHSe5.tsr_{QEQE} plasmid (C), cells with the plasmid but not induced by IPTG (NI), and cells with the plasmid induced with IPTG (I). Cells were initially grown at 37°C and induced at an OD_{600} of ∼0.1. Post-induction growth was at 25°C for the lengths of time specified. Cells used for imaging were collected at 6h after induction, which corresponded to an OD_{600} of 0.68. Cell lysates were normalized by cell density (as determined by OD_{600}) prior to loading.

Immediately proximal to the periplasmic domains of the chemoreceptors in the intracellular membrane, there is a thin region where density is decreased relative to the remainder of the cell body accompanied by a line of stronger density **(Figure 5.3A)**. The latter resembles the width and density of the cytoplasmic membrane in regions lacking receptors. This may indicate that the intracellular membrane curves around to form a continuous thin compartment with minimal internal components, thereby explaining the lack of density between the two putative intracellular membranes. However, the identity of the regions of low and high density immediately proximal to the non-physiological intracellular membrane remains unclear and definitive identification of the biochemical composition of these features will require further studies.

While these results were promising and B/r H266 offers a reasonable choice for structural studies of chemoreceptors in whole cells, membrane assemblies offered the possibility of even thinner samples without cellular background. Thus, isolated membranes seemed more likely to be able to give higher resolution results. In addition, the B/r H266 strain lacks an endogenous chemotaxis system and is genetically distinct from the K-12 strain, which is a precursor for all chemotaxis mutants [210]. Therefore, future studies of native arrays in this system would be complicated by

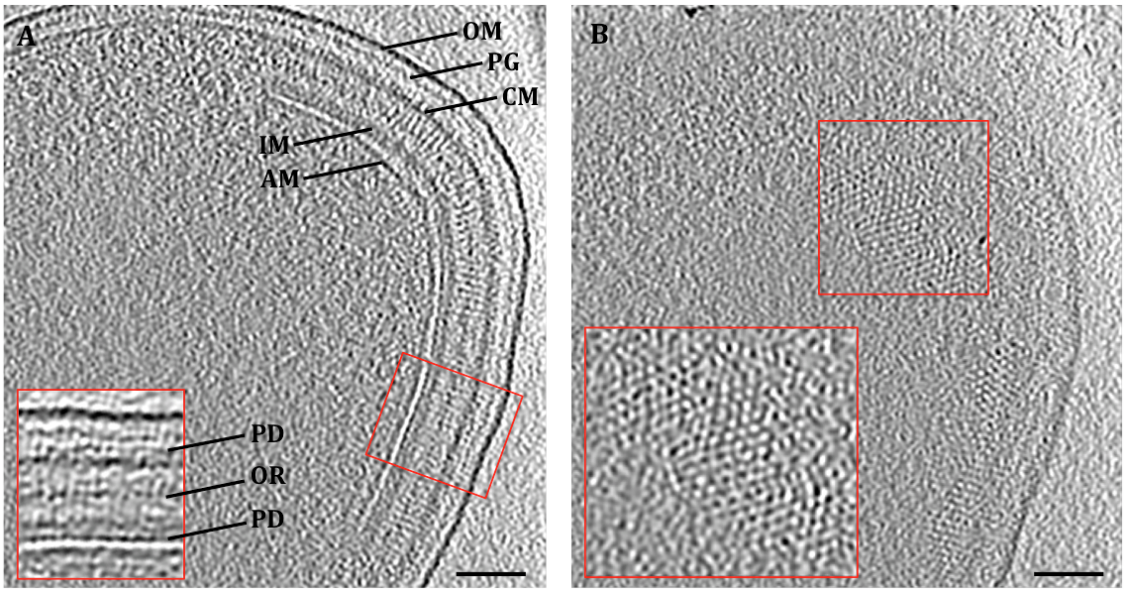


Figure 5.3: Cryo-ET of Tsr assemblies in B/r H266 cells.
A zipper-like assembly (A) and hexagonally packed crystalline patch (B) are evident in tomographic slices of B/r H266 cells overproducing Tsr. Features of cellular morphology are also visible and are denoted in (A). These features include the outer membrane (OM), peptidoglycan layer (PG), cytoplasmic membrane (CM), intracellular membrane containing chemoreceptors (IM), and putative additional membrane (AM) which may be an extension of the intracellular membrane. The insets show selected regions in greater detail. Features of the receptors in the zipper, including the periplasmic domains (PD) and overlapping regions of the membrane distal tips of the chemoreceptors (OR) are indicated in the inset in (A). Scale bars are 50 nm.

the need to introduce and tune the expression of many chemotaxis proteins.

5.1.3 Isolated Membrane Assemblies

A third alternative involved isolating membranes containing Tsr from the remainder of the cell. Previous studies have demonstrated that the zipper-like chemoreceptor assemblies remain intact in isolated membranes and can be viewed from the side as two interdigitated rows of receptors interacting at their cytoplasmic ends or from the top or bottom as a hexagonally packed crystalline patch [185, 169]. As described in **Section 4.3**, membrane assemblies were isolated from Tsr overexpressing HCB721 cells by centrifugation. Both side and top views of zipper-like assemblies were well-

resolved by cryo-ET. Based on prior studies, the periplasmic domains, membranes and overlapping cytoplasmic receptor regions were unambiguously identified in reconstructed tomograms **(Figure 5.4)**. The previously observed hexagonal packing at the region of overlap of the two layers of receptors was also apparent in the top views of several assemblies [185, 169]. Thus, the isolated membranes provided the thinnest type of sample where the receptors were maintained in their native membrane environment but were removed from other cellular components. Therefore, this approach was selected for Tsr structure determination in hopes that it would maximize resolution.

5.2 Tsr Structure in the Presence and Absence of Serine

Cryo-ET was used to image membranes containing Tsr assemblies both in the presence and absence of 1mM serine. Slices through the reconstructed tomograms showed that the overall architecture and characteristic hexagonal packing (as observed by the fast Fourier transform of the tomographic slices) remained the same under both conditions **(Figure 5.4)**. In both cases, the center-to-center hexagonal spacing in top views was 9 nm, which agrees with previously reported values for Tsr zipper-like assemblies [169]. The width of side views between the outer edges of receptor periplasmic domains was 615 $\pm$ 30 Å. The dimensions are consistent with published values for the various parts of the receptor, including an average overlap of 77 Å between interacting receptors [185, 145]. The large range of observed widths is likely a result of the assemblies being tilted differently relative to the 2D tomographic slice in which the width was measured. Though the number of receptors within individual assemblies varied within each of the samples between $\sim$15 and $\sim$200, there were no discernible patterns in the variation.

Following tomographic reconstruction, subtomogram generation, alignment, and averaging were performed by Dr. Xiongwu Wu (NIH Laboratory of Computational

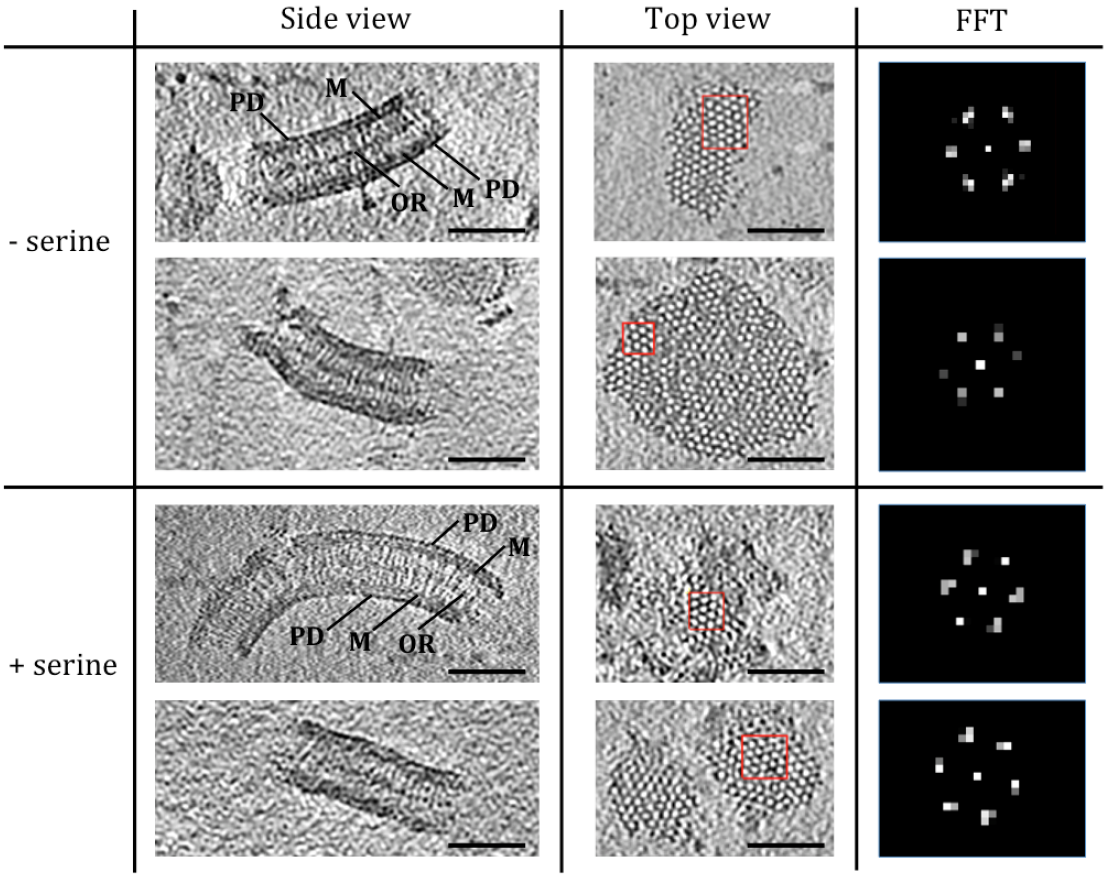


Figure 5.4: Cryo-ET of isolated Tsr membrane assemblies.
Representative tomographic slices showing side views and top views of zipper-like assemblies with and without serine. The periplasmic domains (PD), membranes (M), and regions of overlap at the membrane distal ends of the receptors (OR) are denoted. The fast Fourier transforms of selected areas in top views (boxed in red) were calculated to demonstrate the regular hexagonal packing with 9 nm spacing between the centers of adjacent hexagons. Scale bars are 50 nm.

Biology) as described in **Section 3.4.2**. Average density maps of the Tsr structure in the absence and presence of serine were generated using 1892 and 1663 subtomograms, respectively. The density maps for the two states demonstrate the structural changes in the HAMP and periplasmic domains that occur as a result of serine binding **(Figure 5.5)**. Most notably, the HAMP domain expands when serine is bound, with each of the three dimers being displaced radially outward and forming contacts with neighboring trimers **(Figure 5.5B&E)**. In contrast, the apo structure contains the HAMP domain in a more compact conformation with intra-trimer contacts. These results are consistent with the previously reported structural changes

observed upon serine binding in Tsr overexpressing whole cells [145].

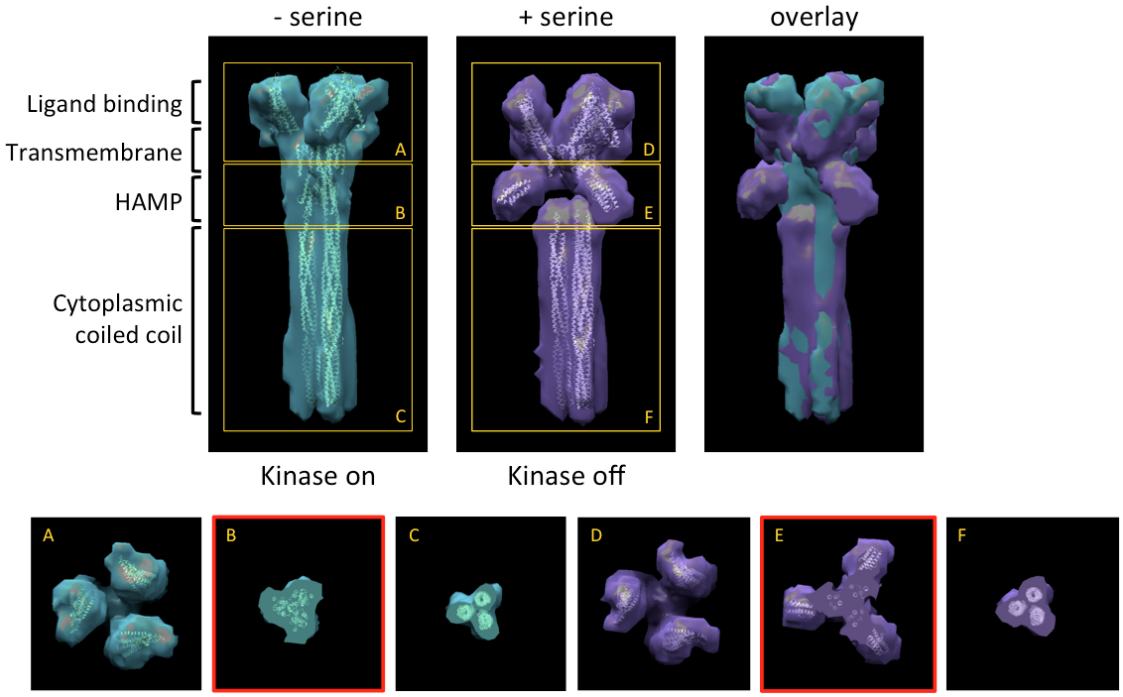


Figure 5.5: Tsr structure in the presence and absence of serine shows expansion of the HAMP domains in response to serine binding.
Average density maps showing the structure of Tsr in the presence (purple) and absence (teal) of serine fitted with a model of a single trimer of dimers. Cross-sectional views correspond to the periplasmic domain and transmembrane helices (A&D), the HAMP signaling domain (B&E), and the cytoplasmic domain (C&F). The expansion of the HAMP domain in the serine bound form as compared to the unliganded form is evident from the comparison of cross-sectional views boxed in red (B&E). The coordinates were taken from the atomic models of the Tsr periplasmic domain (PDB ID code 1VLS), the HAMP domain from an archaeal protein (PDB ID code 2ASW), the Tsr cytoplasmic domain (PDB ID code 1QU7) and a modeled transmembrane domain [144] and were fit in the density map using rigid-body fitting in Chimera.

However, there are differences between previous results and the structures in this study, particularly in the periplasmic domain. As illustrated in **Figure 2.5c**, the previous work by Khursigara *et al.* has shown that in both states, the dimers in the periplasmic domain are separated and slanted in a propeller-like clockwise twist [145]. In this study, neither of the two states show a significant twist in the periplasmic domain and the apo structure shows density connecting the three

dimers **(Figure 5.5)**. Additionally, in both states the dimers in the upper portion of the cytoplasmic domain, which correspond to the methylation helices, are displaced further radially outward by approximately 10 Å with respect to the previously reported structures. The most likely source of these differences is isolation of membrane assemblies from the cell since the imaging conditions and data processing were performed using conditions similar to those used in the study by Khursigara *et al.* Specifically, the aqueous buffer used in this study may not accurately replicate the fluidity of the periplasm, as diffusion coefficients of GFP vary significantly between the *E. coli* periplasm and aqueous buffer [211, 212]. The differences in fluidity of the surrounding media may impact the conformations of the periplasmic domain. In addition, the geometric constraints from curvature at the cell poles where zipper-like assemblies are most often found are not present in isolated assemblies [163]. The relaxation of these constraints may allow for the increased separation of the trimers in the cytoplasmic domain that is observed in this study.

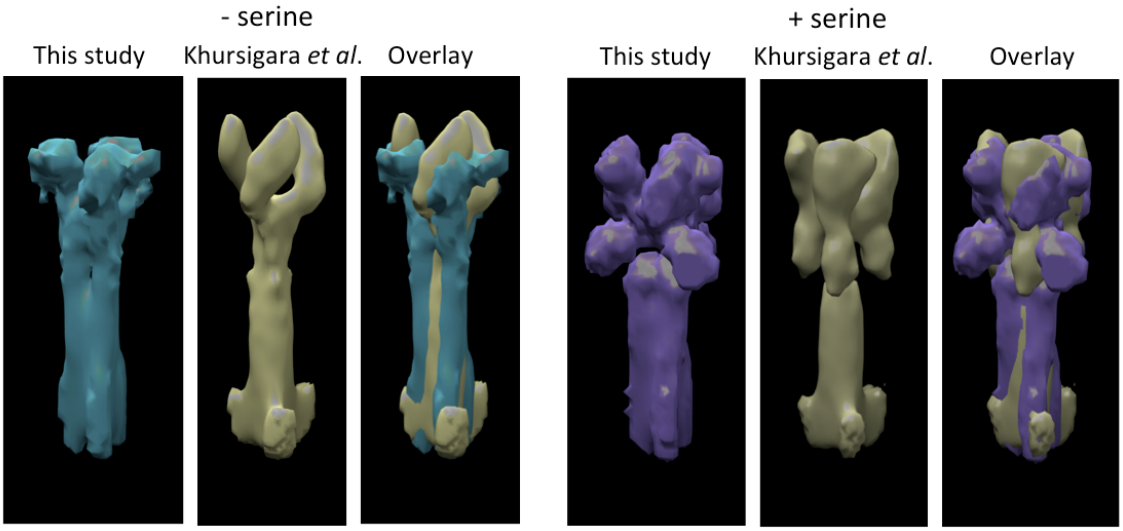


Figure 5.6: Comparison of the structures from this study with previous study by Khusigara *et al.* [145].
Side by side and overlay comparisons for the average density maps from this study compared to the previous electron microscopy study. Structures from this study are shown in teal and purple for the states in the absence and presence of serine, respectively. The density maps from the Khusigara *et al.* study are shown in yellow.

5.3 Discussion

Although there have been studies of assemblies before, this study is novel in its attempt to determine the full-length Tsr structure from isolated Tsr membrane assemblies [185, 169]. It is also the first time that isolated assemblies have been studied outside of the cellular context in the presence and absence of serine. The results show that serine binding causes expansion of the HAMP domain as was previously reported [145]. However, there are differences in the structures of the periplasmic and cytoplasmic domains in relation to the prior work by Khursigara *et al* **(Figure 5.6)**. These inconsistencies are likely a consequence of the variations in surrounding media and the relaxed geometric constraints that result from removing Tsr assemblies from whole cells. Further structural work on full length Tsr will be needed not only to reconcile these discrepancies but also to study the structural states of the receptor in functional arrays with the cytoplasmic proteins, CheA and CheW. Both this study and the previous work by Khursigara *et al.* were done on non-functional arrays containing only Tsr, so it still must be determined if the observed expansion of the HAMP domain occurs in functional arrays or whether this conformational change is an artifact of zipper-like assemblies [26].

The formation of Tsr assemblies in the absence of CheA and CheW both in this study and in previous work affirms that these receptors have a propensity to interact at their membrane-distal tips, either with each other or with their cytoplasmic proteins [163, 145, 169]. However, based on previous fluorescence and electron microscopy studies, it appears that chemoreceptors are only able to cluster in arrays when they are overexpressed even though their physiological expression levels are high relative to many other proteins [168, 135, 163]. Thus, there appears to be a threshold expression level when Tsr clusters into zipper-like assemblies in the absence of CheA and CheW. Imaging of receptors expressed at different levels in a *cheA-cheW* knockout strain could be used to empirically determine the receptor expression level that corresponds to the change in receptor clustering. It also remains unknown whether this switch in clustering behavior is specific to Tsr in *E.*

coli or whether it is a general characteristic of methyl accepting chemotaxis proteins (MCPs) in various bacterial species. The structure of native arrays containing MCPs, CheA and CheW has been shown to be conserved across bacterial species, but overexpression studies of MCPs in other species besides *E. coli* have not been carried out to test for the formation of zipper-like assemblies [164]. However, the zipper-like architecture of these assemblies is reminiscent of the native cytoplasmic arrays found in *R. sphaeroides* and *Vibrio cholerae* [167]. When compared to the zipper-like assemblies of Tsr, the receptors of the cytoplasmic arrays are oriented in the opposite direction in the zipper sandwich since the kinase activating ends (which correspond to the cytoplasmic ends of the membrane receptors) face outward. The kinase activating ends interact with CheA and CheW in ordered hexagonal arrangements on both sides of the zippered structure. Interestingly, cytoplasmic fragments of the *E. coli* chemoreceptor Tar also form assemblies with CheA and CheW that resemble the cytoplasmic arrays. These assemblies contain the kinase activating ends facing outward in an ordered hexagonal array containing CheA and CheW [167]. These results in conjunction with the work on Tsr zipper-like assemblies done previously and in this study show an emerging theme of the need for the kinase activating ends of the receptors to organize into ordered hexagonal arrays as well as the varied use of zipper-like assemblies to achieve this end.

The need for the cytoplasmic ends of Tsr to form ordered arrays using zipper-like assemblies was further demonstrated in the B/r H266 strain. Unlike previously observed Tsr overexpressing cells where the cytoplasmic membrane was invaginated, this strain produced an intracellular membrane in order to accommodate receptor-receptor interactions in zipper-like assemblies [163]. In both cases, it appears that the cells increased the available membrane area into which overexpressed receptors could be inserted rather than allowing the receptors aggregate in inclusion bodies. Therefore, co-expression of Tsr along with a membrane protein of interest in *E. coli* may provide a mechanism for increasing the production of functional membrane proteins that may otherwise end up in inclusion bodies. This type of approach has been used in the past with calveolin-1 co-expression to display both eukaryotic and

prokaryotic membrane proteins on intracellular vesicles [213, 214]. Since calveolin-1 is of eukaryotic origin, Tsr may be a preferable alternative since it is a native protein in *E. coli* and therefore is less likely to adversely affect cell growth and metabolism.

The differences between the structures in this study as compared to those from previous whole cell studies show that the environment used to study the arrays has a significant effect on the organization. In addition to using the thin cell strain, B/r H266, as was described in this chapter, alternative methods for reducing cell thickness while maintaining the cellular context need to be explored. For example, one approach has been the creation of minicells, round anucleate cells often less than 0.5 µm which result from asymmetric division near the poles of small diameter cells [166]. Previous cryo-ET studies of chemoreceptor arrays on minicells have provided structural insights, but the maximum achievable resolution remains limited by the remaining density of the minicells [206, 166]. Another elegant solution has been the creation of *E. coli* cell ghosts by lysis mediated by the phage ϕX174 E gene [215]. This approach involves the release of cytoplasmic contents through a small lesion while the majority of the inner and outer membranes remain intact. Cell ghosts retain their rod shape until the point of vitrification, at which point the cell membranes collapse to yield just two layers of cell membranes. This technique has been used to perform cryo-ET native chemoreceptor arrays with superior signal-to-noise relative to other studies [215].

In this study, the resolution of the determined structures is in the range of 25-30 Å. Though large-scale quaternary structural changes are evident at this resolution, resolving the movements of individual helices would require structures at a resolution below 10 Å. Other recent cryo-ET studies have achieved such resolutions, largely by the implementation of direct electron detectors and associated advances in image processing that were unavailable in this lab when this study was performed [44, 45]. This study was performed using a 2K charge-coupled device (CCD) for detection which limited the ability to use a total electron dose of less than 100 $e^-/Å^2$ and limited the accuracy of CTF estimation in individual tilted images. Furthermore,

the inability to use a CCD to collect in 'movie mode' with intermediate frames for this study made drift correction in individual tilted images impossible. The resolution of the structures, and consequently the level of biological information obtained, could have been improved by using better detection technology and related image processing advances.

In summary, this study sought to utilize cryo-ET imaging in the presence and absence of serine to determine structural changes that accompany ligand binding. In the process, techniques for reducing the thickness of the samples were examined and two viable options (expression in the B/r H266 strain and isolation of membrane assemblies) were identified. Though the B/r H266 strain was not selected for further structural study, overexpression of Tsr in this strain revealed novel features. These included the formation of an intracellular membrane to accommodate zipper-like assemblies of highly expressed Tsr as well as additional variations in density that may indicate a thin membrane compartment formed from the receptor-containing intracellular membrane looping around. Both the sources of these additional regions of low and high density and the underlying mechanisms behind the formation of these structures remain to be determined.

Subvolume averaging of the receptor densities in isolated Tsr membrane assemblies in the presence and absence of serine was utilized as a novel approach for determining the full length receptor structure. However, the technical limitations on resolution and the structural differences from previous whole cell studies made it difficult to draw novel physiologically relevant conclusions. Further studies with newly developed direct electron detectors, in the presence of CheA and CheW, and in the cellular context will be required to determine the conformational changes in full-length receptors in native arrays.

Chapter 6

Organization of the *R. sphaeroides* Cytoplasmic Chemosensory Partitioning Protein, PpfA

The *R. sphaeroides* cytoplasmic chemosensory cluster is segregated by the ParA-like ATPase, PpfA, in order to ensure that each daughter cell receives a cluster. Previous studies using fluorescent protein fusions have shown that wild-type PpfA distributes across the nucleoid with bright foci corresponding to the cytoplasmic cluster. When PpfA is locked in the ATP-bound state and cannot hydrolyze ATP (D39A mutant), it produces brighter foci than WT PpfA [178]. This suggests that the D39A mutant accumulates more protein at these foci because it cannot dissociate once it binds to the cluster or nucleoid. In order to further probe this difference in localization between the WT and D39A forms of PpfA, both forms of the protein were tagged so that they could be selectively labeled and visualized by electron microscopy.

6.1 Creation of Plasmids and Strains

As described in **Section 1.2**, there are currently two commonly used electron dense tags used to specifically label subcellular protein for electron microscopy. The first, immuno-gold labeling, utilizes gold to visualize an antibody that binds to the pro-

tein of interest in sections of cells that have been chemically fixed and stained. The second involves a genetically encodable tag that produces a singlet oxygen that reacts with diaminobenzidine (DAB) to generate an osmophillic precipitate around the protein of interest. Subsequent staining with osmium tetraoxide generates electron dense regions where the precipitate was deposited.

Constructs were generated with PpfA and PpfAD39A such that either of these two approaches could be used for visualization [36]. Since there was no direct antibody against PpfA that could be used for immuno-labeling, PpfA and PpfAD39A were tagged with two HA epitopes that are readily detectable using immuno-gold EM. In addition, both forms of PpfA were tagged with APXII in order to allow visualization with DAB and osmium tetraoxide [216]. For convenience, the proteins were tagged with both the APXII and HA tags on the same construct. Both the WT and D39A mutant forms were tagged N- and C-terminally such that the APXII tag was closest to the protein and the the two tandem HA epitopes were separated from the protein coding sequence by the APXII tag. The tagged variants of PpfA and PpfAD39A were all created in pIND, the IPTG-inducible expression expression vector for *R. sphaeroides* [186]. The vectors were conjugated into tho the WT *R. sphaeroides* strain WS8N as well a strain containing a genomic TlpT-YFP fusion (JPA1558) to allow visualization of the cytoplasmic cluster.

6.2 Fluorescence Microscopy for Phenotypic Validation

Fluorescence microscopy was performed in order to determine whether cells expressing the different levels of tagged PpfA and PpfAD39A were able to segregate their fluorescently-tagged cytoplasmic clusters. Since PpfA functions to segregate cytoplasmic clusters, cells with functional PpfA would be expected to display either one or two clusters, depending on the phase of the cell cycle during which a given cell was imaged. Cells nearing division would display two clusters while others would display

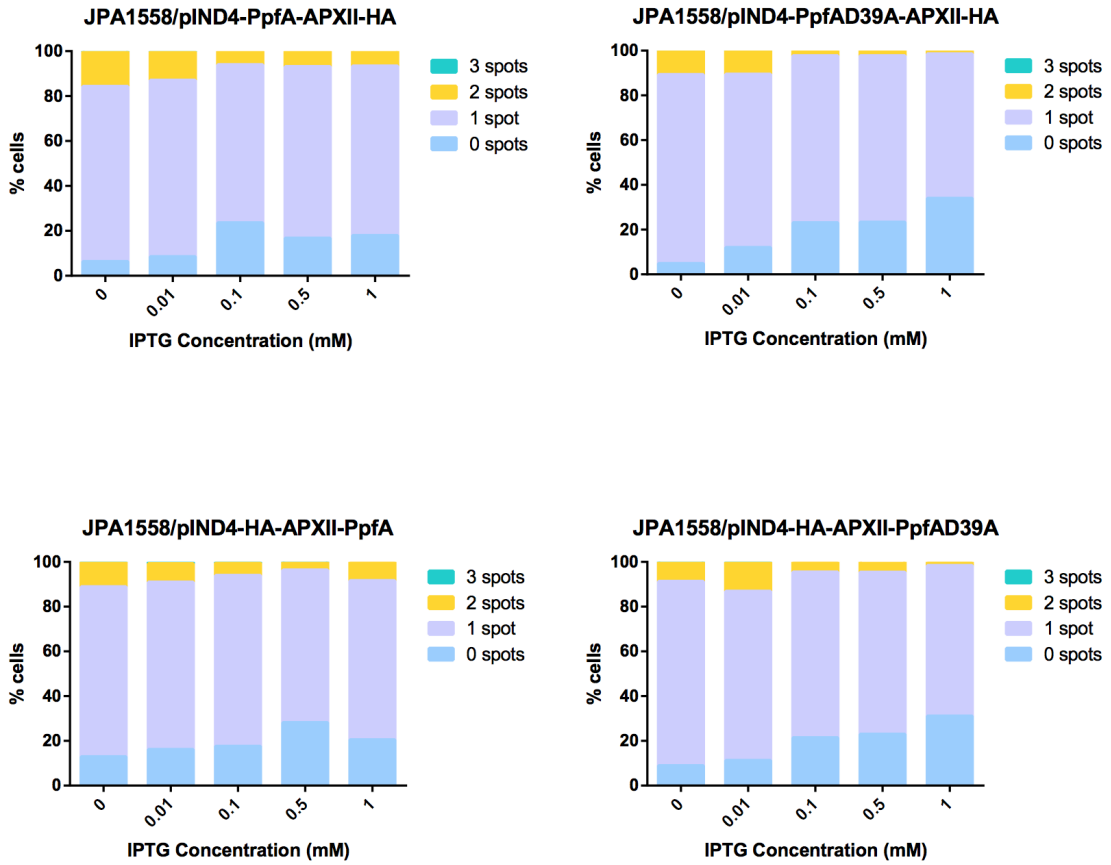


Figure 6.1: Extent of cluster segregation by tagged variants of PpfA. Quantification of the number of fluorescent foci observed in TlpT-YFP cells containing tagged PpfA constructs induced at different levels. For all four constructs, there appeared to be a sharp drop in cluster segregation efficiency (as indicated by a significant drop in the number of cells with 2 spots) between induction levels with 0.01mM and 0.1mM IPTG.

one cluster. However, anomalies of either three or zero clusters would occasionally be observed. In cells where PpfA had reduced or no functionality, either zero or one cluster would be observed. Since cluster segregation was disrupted, one of the two daughter cells during each division would not receive a cytoplasmic chemosensory cluster and would need to synthesize it *de novo* and thus may not display an observable cluster. Even cells nearing division would have only a single cluster rather than two because of the lack of segregation.

R. sphaeroides cells expressing YFP-tagged TlpT from the genome were used as the

base strain for assays since they phenotypically display a fluorescently tagged cluster. Tagged PpfA and PpfAD39A constructs were introduced into these cells. They were subsequently grown and induced as described in **Section 3.3.2** before being places on agarose pad slides for fluorescence imaging. Phase and YFP images were collected as described in **Section 3.4.1** and were analyzed for the number of visible foci using MicrobeTracker [187]. In the case of all four constructs, there appeared to be a sharp drop in cluster segregation efficiency between induction levels with 0.01mM and 0.1mM IPTG **(Figure 6.1)**. Thus, it was determined that the best initial induction level to be used for EM imaging was 0.01mM since the phenotypic activity of PpfA was nearly unaffected at that level. Additionally, the C-terminally tagged constructs (pIND-PpfA-APXII-HA and pIND-PpfAD39A-APXII-HA) appeared to have a more consistent pattern of progressively reducing cluster segregation as the induction level was increased. Previous fluorescent fusions of PpfA had also utilized C-terminal tags [178]. Therefore, the C-terminally tagged constructs were used for EM imaging.

6.3 Comparison of Organization of Wild-Type and ATP-locked PpfA by Immuno-gold Labeling

Immuno-gold labeling rather than DAB and osmophillic staining was selected as the primary method for visualizing the distribution of the WT and mutant forms of PpfA. Instead of detecting a precipitate formed near the protein, this would allow detection of the HA epitopes with which the protein has been tagged.

R. sphaeroides cells containing the C-terminally tagged PpfA constructs were grown and imaged as described in **Sections 3.3.2 & 3.4.3**. Imaging was initially performed on WS8N cells containing the pIND-PpfA-APXII-HA construct which had been induced with 0.01mM IPTG. Unfortunately, the signal in these images was too low to draw any conclusions **(Figure 6.2)**. In order to have sufficient signal to see a labeling pattern, cells that were induced with 0.5mM IPTG needed to be imaged.

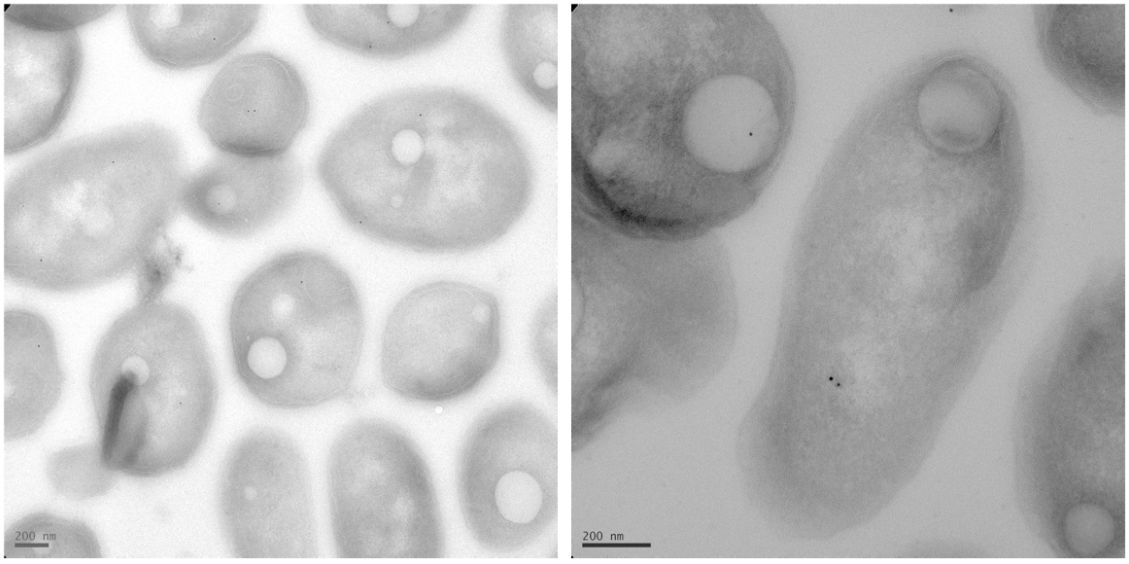


Figure 6.2: Immuno-gold labeling of HA-tagged PpfA induced with 0.01mM IPTG.
Labeling of WS8N cells containing the pIND-PpfA-APXII-HA construct which had been induced with 0.01mM IPTG.

At this expression level, cells WS8N cells containing either pIND-PpfA-APXII-HA or pIND-PpfAD39A-APXII-HA show some instances of a labeling pattern where there are foci that appear to have more dense labeling than the remainder of the cell body **(Figure 6.3)** However, these results are still preliminary as they need to be validated by statistical analysis to analyze the distribution of the gold labeling relative to the cell body and determine the propensity for clustering.

6.4 Discussion

The results show that this approach can be used to label PpfA, but that statistical analysis must be employed in order to be able to make definitive conclusions about the degree of clustering and how it varies between the wild-type and mutant form of PpfA. Additional validation from immunofluorescence experiments using the same antibody would help to confirm sufficient accessibility of the HA tag to determine if this was the cause of low signal. This is particularly relevant because very high induction levels of tagged PpfA were required in order to have sufficient signal to

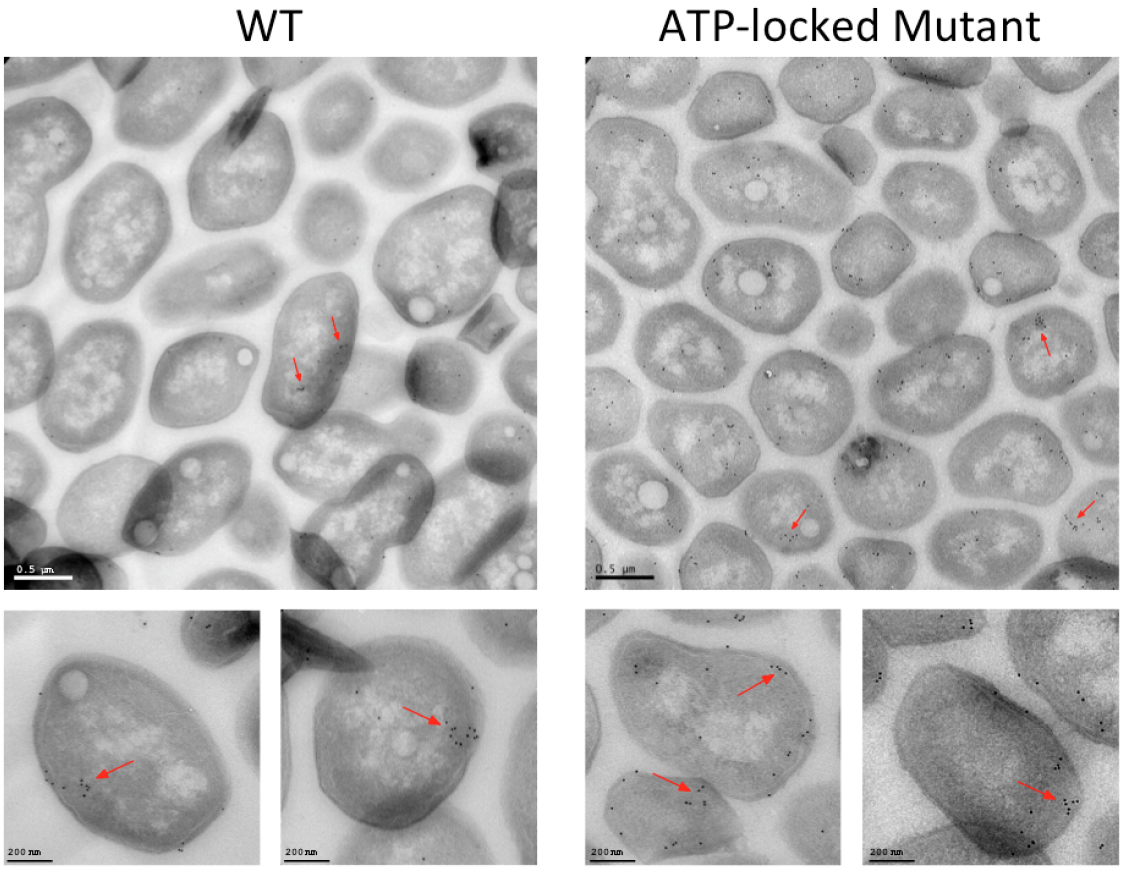


Figure 6.3: Immuno-gold labeling of wild type and mutant HA-tagged PpfA induced with 0.5mM IPTG.
Comparison of immuno-gold labeling of cells expressing the wild-type and ATP-locked mutant forms of PpfA tagged with HA. Though there does not appear to be a significant difference in the pattern of distribution, the level of labeling appears to be significantly higher in the cells expressing mutant PpfA. There is some evidence of clustering (denoted by red arrows) observed in both samples. There appears to be more visible clusters in the mutant sample, but that is likely due to higher overall levels of signal.

recognize a suggested labeling pattern. Immunofluorescence will also provide higher signal and therefore be easier to interpret since it detects epitopes throughout the whole cell body while immuno-EM only detects epitopes that are exposed on the surface of each section.

The low observed signal may also be a result of the fact that untagged WT PpfA is being expressed from the genome in these cells and may be outcompeting tagged PpfA in its interactions with the cluster. Moreover, the lack of a polyclonal antibody

raised directly against PpfA makes it more difficult to obtain a signal. As previously mentioned, the antibody only binds to exposed epitopes at the surface of a section. Therefore, a small epitope such as HA is less likely to be detected than the larger protein to which it is tagged. However, this can be alleviated to some extent by having 5 or 6 HA tags in tandem. Another possibility for labeling without a direct antibody may be using a polyclonal antibody raised against a close homologue such as the ParA-like ATPase Soj [178]. If cross-reactivity of the antibody with PpfA is confirmed using immunofluorescence, then it can provide a viable substitute for an antibody raised directly against PpfA.

One consistent observation when comparing the images of WT and mutant PpfA was that the mutant always appeared to have a higher degree of labeling **(Figure 6.3)**. The source of the higher signal in the mutant is unclear but it may result from differential protein expression between the WT and mutant constructs. This can be evaluated by analyzing protein expression using SDS-PAGE. If the D39A mutation results in a structural change of PpfA, this could increase accessibility of the HA tag and thus cause higher signal. Further studies will be necessary in order to distinguish between the possible reasons for the higher signal in the mutant samples.

While immuno-gold labeling of cryo-sections offers a possible avenue for probing the cellular localization of PpfA, there are other technologies that can also contribute to this understanding. Among these is super-resolution microscopy which has the potential to allow for superior spatial and temporal resolution when compared to traditional light microscopy [36]. Since these techniques are based on fluorescence, they offer the advantage that the protein of interest, PpfA in this case, can be specifically tagged by a fluorescence marker which allows it to be definitively identified [35]. Previous unpublished work in the Armitage lab has utilized photo-activated localized microscopy (PALM) to observed the spatial distribution of PpfA. Based on these studies, PpfA is believed to loosely associate in a non-specific manner with the chromosome, which does not compact in *R. sphaeroides* cells. Thus, while there is a higher local concentration of PpfA near the cytoplasmic cluster, it is also found

throughout the cell.

One remaining question is whether PpfA associates solely with the outer surface of the bacterial nucleoid or whether it is embedded throughout. The z resolution of PALM is insufficient to determine whether the nucleoid occludes PpfA since fluorescence above and below the plane of focus is often still visible. While cryo-sections offer an option to visualize z slices, there are other techniques that also have sufficient z resolution to be able to address this question. Amongst them is focused ion beam - scanning electron microscopy or FIB-SEM that has reported resolutions of up to 3nm in the x, y, and z directions [40]. This technique, based on iterative abrading of the sample with a focused ion beam followed by imaging of the exposed surface offers the opportunity to get a 3-D volume of the sample from the stack of images of the exposed surface. By imaging a sample where PpfA was tagged with APEX or a similar entity to produce an electron dense stain, the 3-D distribution of PpfA could have been determined. In addition to being able to localize the specific protein of interest, this technique also offers that advantage that it can be correlated with fluorescence microscopy. In the case of PpfA, this may be useful for attempts to correlate the localization of PpfA to the fluorescently tagged cytoplasmic cluster [40]. However, this technique has the caveat that it requires specimens to be fixed, dehydrated, stained, and embedded in resin prior to imaging. Thus, caution must be taken in interpreting the physiological meaning of the results obtained from a sample that has undergone preparation through a series of highly non-physiological conditions.

An alternative method that more closely mimics native conditions is imaging lamellae produced from frozen-hydrated specimens milled by a cryo-FIB. This technique involves preserving the sample in a near-native state by either plunge-freezing or high pressure freezing depending on the thickness of the sample (high pressure freezing more rapidly freezes thicker samples) [217]. Subsequently, a FIB is used as a precise mechanism to cut a thin slice of the frozen sample, which can then be subsequently studied by cryo-EM or cryo-ET [218, 219]. This offers the advantage of being able to

circumvent complications relating to sample thickness for whole cell studies. In the case of PpfA, this technique may have utility because it maintains the protein and its cellular context in a near-native state. However, it may be difficult to specifically localize PpfA using this technique because there is no known evidence that it forms recognizable higher order structures. While there are hypotheses in the field that PpfA may polymerize into filaments, none have been observed in PALM studies performed in the Armitage lab.

There are certainly a number of methods that can be used to probe intracellular protein localization, Though it may require some optimization in order to observe sufficient labeling, immuno-gold labeling of cryo-sections remains one of the most sensitive techniques to specifically identify a target protein of interest in individual z slices with high resolution in x and y. In the context of this study, further optimization of labeling as well as statistical analysis of the propensity for clustering would be required in order to draw biological conclusions.

Chapter 7

Discussion

The structures of chemoreceptor arrays are largely conserved across many species and repeating organizational themes have arisen. For instance, hexagonal packing has been observed in all arrays studied, though the packing densities vary based on the specific components of the array. Endogenous arrays containing appropriate CheAs and CheWs have 12 nm hexagonal center-to-center spacing, regardless of whether they are membrane-associated or cytoplasmic [164, 167]. However, the absence of certain components can result in altered spacing. As shown in **Chapter 5** and previously reported by Briegel *et al.*, Tsr arrays without the partner cytoplasmic proteins are more closely packed and demonstrate 9 nm center-to-center hexagonal spacing [169].

While there have been numerous studies on the receptor packing and arrangement of the membrane-associated chemoreceptor arrays, the overall architecture of the *R. sphaeroides* cytoplasmic cluster was only recently elucidated [167]. A number of questions about the structure and organization of the components of the *R. sphaeroides* cytoplasmic cluster remain unanswered. Since there are two soluble chemoreceptors (TlpT and TlpC), there may be distinctions in the way that they arrange and interact with other components of the cytoplasmic chemosensory cluster. Recent unpublished work has shown that the cytoplasmic CheAs ($CheA_3$ and $CheA_4$) interact differently with TlpT and TlpC. It has also been observed that the cluster packs more loosely when the Tlps are attached to the P5 domain of $CheA_3$

and more tightly when the Tlps are attached to the P5 domain of $CheA_4$. The fact that $CheA_3$ and $CheA_4$ are different lengths further supports the hypothesis that there are structural distinctions in arrangement between clusters involving one of the two cytoplasmic CheAs. The packing of the cytoplasmic cluster in the absence of the CheAs also remains unknown. The presence of TlpT and $CheW_4$ is sufficient to form a cytoplasmic array, so it is possible to compare the array architecture in the absence of CheAs to the architecture when each of the CheAs is involved alone.

Careful comparison of structures of arrays where the molecular components are varied may provide valuable insights into the functions of the different chemotaxis proteins. This is particularly relevant in cases such as those of the cytoplasmic CheAs where there is a duplication of homologous components but evidence that they function differently and are not completely interchangeable. The relatively small amount of structural knowledge on the cytoplasmic chemoreceptor arrays makes it a very promising, though challenging, subject of study.

Cryo-EM and cryo-ET provide powerful tools for elucidating the architecture of chemoreceptor arrays. Since chemoreceptor arrays have a very distinct structure, they are readily identifiable in frozen hydrated samples that are not labeled for a specific target protein. Recent developments in detectors have made it possible to determine higher resolution information about these arrays than ever before. In addition, the emergence of technologies such as cryo-focused ion beam (FIB) milling have provided new mechanisms for studying samples, such as whole bacterial cells, that were previously prohibitively thick for high resolution cryo-TEM studies.

Part III

Human Cancer Metabolism

Chapter 8

Introduction to PKM2 and its Role in Cancer Metabolism

8.1 Cancer-related Modifications to Cellular Metabolism

Human cancer cells arise from the acquisition of multiple alterations in normal cells in order to be able to proliferate unnaturally. The changes relate to a variety of processes including cell growth, regulation of gene expression and cell metabolism [220]. The metabolic requirements of cancer cells are determined by the specific genetic and epigenetic state of the cells in addition to their microenvironment. Thus, in order to continue to grow and divide with varying levels of nutrients and oxygen, cancer cells must carefully regulate their metabolic machinery [221].

There is growing body of evidence that metabolic reprogramming plays an important role in cancer and that, in some cases, it may also play a part in the initiation of tumorigenesis [222, 223, 224, 225]. The importance of metabolic changes in cancer cells was first recognized by the biochemist Otto Warburg in the 1920s. He demonstrated that cancer cells, unlike normal cells, always upregulate glycolysis even when oxygen is readily available [226, 227]. This metabolic reprogramming to utilize glycolysis is known as the Warburg effect [228].

Glycolysis in an oxygen-independent metabolic pathway that converts glucose into

two molecules of pyruvate with a net gain of two molecules of adenosine triphosphate (ATP) and two molecules of reduced nicotinamide adenine dinucleotide (NADH) [229]. Although oxygen is not used in the glycolysis pathway itself, the metabolism of pyruvate through the mitochondrial citric acid cycle is an aerobic process. The NADH generated by the citric acid cycle is then used for another aerobic process, oxidative phosphorylation (electron transport), which uses a proton gradient to produce large amounts of ATP [230]. When oxygen is absent or in short supply, pyruvate is converted to lactate by the enzyme lactate dehydrogenase (LDH) [231]. Thus, glycolysis can take place either in the presence or absence of oxygen, though the treatment of pyruvate produced by glycolysis varies based upon the level of oxygen present.

The majority of ATP production in healthy cells results from mitochondrial oxidative phosphorylation. Warburg proposed that cancer cells switch to glycolysis mechanism for ATP production as a result of mitochondrial dysfunction [232]. Since glycolysis is less efficient than oxidative phosphorylation for producing ATP, it must be greatly upregulated in cancer cells if it is the sole source of ATP synthesis. However, Warburg's hypothesis has been challenged since recent studies have found that many cancers that display upregulated glycolysis have no disruption of oxidative phosphorylation. If exposed to the same oxygen pressures, cancer cells produce similar amounts of ATP by oxidative phosphorylation when compared to their normal counterparts [233, 234]. New studies have also revealed that upregulated glycolysis contributes to the synthesis of biomasses such as ribonucleotides and amino acids in addition to producing ATP [235, 236]. Glycolysis also reduces the production of nicotinamide adenine dinucleotide phosphate (NADPH), which results in the reduction of reactive oxygen species (ROS) which are generated by increased metabolism under hypoxic conditions [237, 238]. This suggests that the Warburg effect may be a strategy for cancer cells to not only to cope with requirements for growth and proliferation in a dynamic microenvironment where oxygen and nutrients may be limited, but also to reduce ROS and therefore oxidative stress [239].

The regulation of ROS at levels elevated with respect to healthy cells is essential for cancer cell survival and proliferation. Elevated ROS levels are required for cancer cell growth and proliferation, tumor angiogenesis, and promoting invasion and metastasis [240, 241, 242]. However, ROS also contribute to oxidative damage that can adversely affect cells and even cause cell death [243]. Thus, the level of ROS in cancer cells must be maintained within a range that is high enough to promote growth and proliferation but low enough to avoid severe oxidative damage. Thus, the Warburg effect and its regulation of redox status through the production of NADPH is an important cancer cell survival mechanism [239].

Although the Warburg effect refers to the metabolic reprogramming used by cancer cells, the phenomenon of aerobic glycolysis has been observed in other normal cells undergoing rapid proliferation (such as stimulated lymphocytes and mitotic and proliferating fibroblasts) [244, 245, 246]. This metabolic pattern has been attributed to the temporarily high demands for metabolic material and energy required for rapid cell division. *E. coli* and yeasts also employ aerobic glycolysis, suggesting that this might be an evolutionarily-conserved method for cells that need to meet special needs during periods of rapid cell proliferation [247, 248]. It is believed that cancer cells exploit this strategy in order to cope with their constant need for rapid growth and proliferation. However, the interpretation of the functional roles of the Warburg effect continues to evolve as new information becomes available [239].

8.2 The Role of PKM2 in Cancer

An important step in the metabolic reprogramming of cancer cells is a switch in the isoform of pyruvate kinase, which catalyzes the conversion from phosphoenolpyruvate (PEP) and ADP to ATP and pyruvate in the last step of glycolysis. Most cancer cells use the M2 isoform of pyruvate kinase (PKM2) in place of the M1 isoform of the enzyme (PKM1) which is used in most healthy adult tissues [249]. PKM1 and PKM2 are alternative splice variants of the *PKM* gene that differ solely by one exon

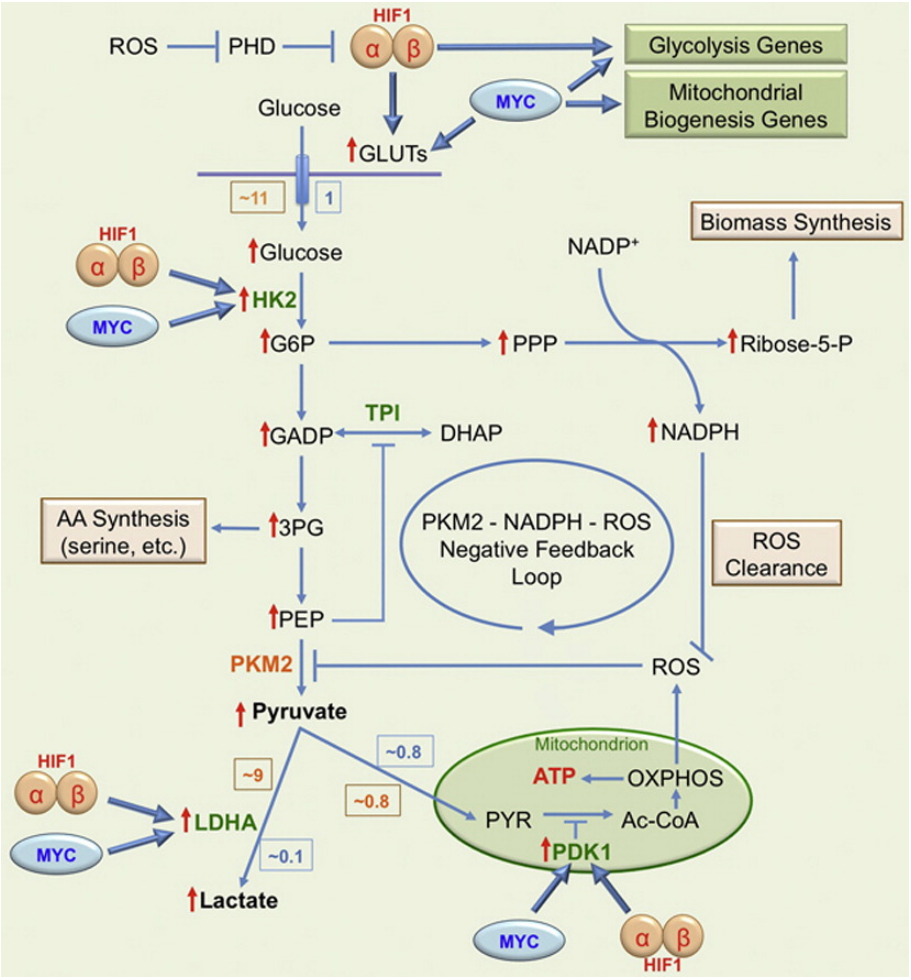


Figure 8.1: Schematic representation of the Warburg effect.
The glycolytic pathway is upregulated and metabolic reprogramming results in the use of PKM2 rather than PKM1 to catalyze the final step. This results in a channeling of upstream products to the pentose phosphate pathway (PPP) which increases NADPH levels and clears ROS. PKM2 is involved in this feedback loop to regulate ROS levels. The enzymes in orange function mainly in cancer cells and the enzymes in green function in both normal and cancer cells. Red arrows indicate elevated levels in cancer cells. The relative amount of consumption of glucose and its metabolic products is shown in blue and orange for normal and cancer cells, respectively. Figure taken from [239].

[250]. PKM2 is less active than PKM1 because of its tendency to dissociate from an active tetramer to a dimer that inefficiently catalyzes PEP [251, 252]. Since cancer cells upregulate glycolysis not only for the synthesis of ATP, but also for the production of biomass and NADPH, the use of PKM2 allows cancer cells to upregulate the early steps of glycolysis without producing an excess of pyruvate. This leads to the

accumulation of early glycolytic intermediates, which can then be diverted to other pathways such as the amino acid biosynthesis, hexosamine, and pentose phosphate pathways [237, 238]. The use of PKM2 results in the accumulation of PEP, which activates the pentose phosphate pathway (PPP) by inhibiting the glycolytic enzyme triosephosphate isomerase (TPI). This increases antioxidative metabolism by increasing NADPH production, reducing ROS and amplifying the inhibitory effect of PKM2. In addition, ROS and PKM2 form a negative feedback loop to regulate ROS levels since PKM2 activity is decreased by oxidization on cysteine 358 by the ROS hydrogen peroxide **(Figure 8.1)** [237, 238]. The combination of these regulatory mechanisms involving PKM2 allow cancer cells to achieve a metabolic balance between ATP production, biomass synthesis, and regulating oxidative stress resulting from ROS generation [239].

8.3 PKM2 Structure and Function

All isoforms of mammalian pyruvate kinase are tetrameric protein consisting of four identical subunits arranged as a dimer-of-dimers **(Figure 8.2)**. Each monomer consists of three main domains (A, B and C) plus a small N-terminal domain [253, 254]. The A domain is a symmetric α_8/β_8 barrel tertiary structure with the active site located in the cleft between the A and B domains at one end of the barrel. The B domain is mobile and restricts access to the active site upon binding the Mg^{2+}-ADP substrate complex [255]. The C domain is located on the opposite side of the A domain and contains the binding pocket for the upstream glycolytic intermediate fructose-1,6-bisphosphate (FBP) that serves as an allosteric activator for the PKM2 isoform [256, 257]. Individual subunits come together at the A domains in order to form dimers and at C domains to form the dimer-of-dimers. The amino acids encoded by the alternatively spliced exon that distinguished PKM1 and PKM2 are located in the C domain, which forms the dimer-dimer interface **(Figure 8.3)**. Thus, the differences in the amino acids in this region are responsible for the distinctions in tetramer stability, FBP binding, and allosteric regulation between the PKM1 and

PKM2 isoforms [257].

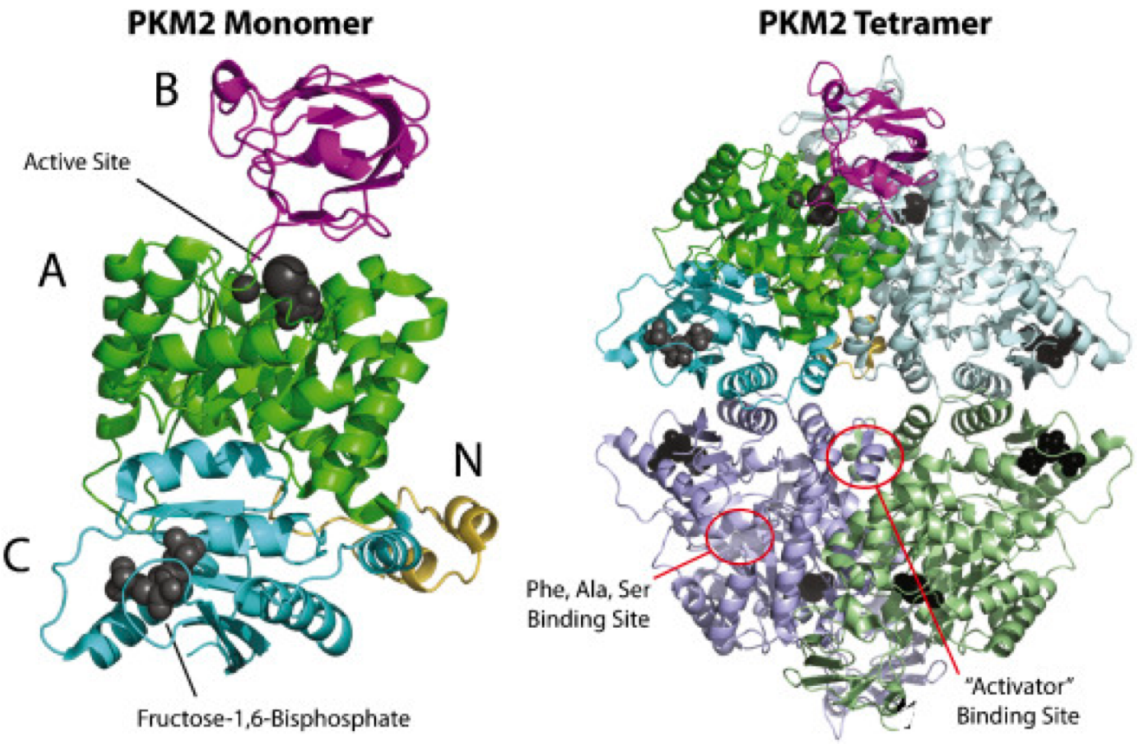


Figure 8.2: Crystal structure of PKM2 with bound ligands.
Structure of tetrameric PKM2 (PDB ID: 3BJF [249]) with one monomer colored to show the A (green), B (magenta), C (blue) , and N (yellow) domains. FBP and the molecules at the catalytic site (K+,Mg^{2+}, and oxalate) are shown as grey spheres. Figure taken from [258].

8.3.1 Allosteric Regulation

Mammalian pyruvate kinases catalyze the phosphate transfer from PEP and ADP (which is complexed with Mg^{2+}) to ATP and pyruvate using a conserved lysine residue found at position 270 [259, 260]. During catalysis, the catalytic lysine residue stabilizes the pentacoordinate transition state that exists as the phosphate is transferred from PEP to ADP. Different pyruvate kinase isoforms have similar kinetic parameters with respect to ADP, but differ in their kinetics with respect to PEP [261, 262, 263]. While ADP binding is non-cooperative and insensitive to allosteric regulation, most pyruvate kinase regulation occurs through changes in PEP binding affinity. In the absence of allosteric activators, PKM2 has low affinity for PEP

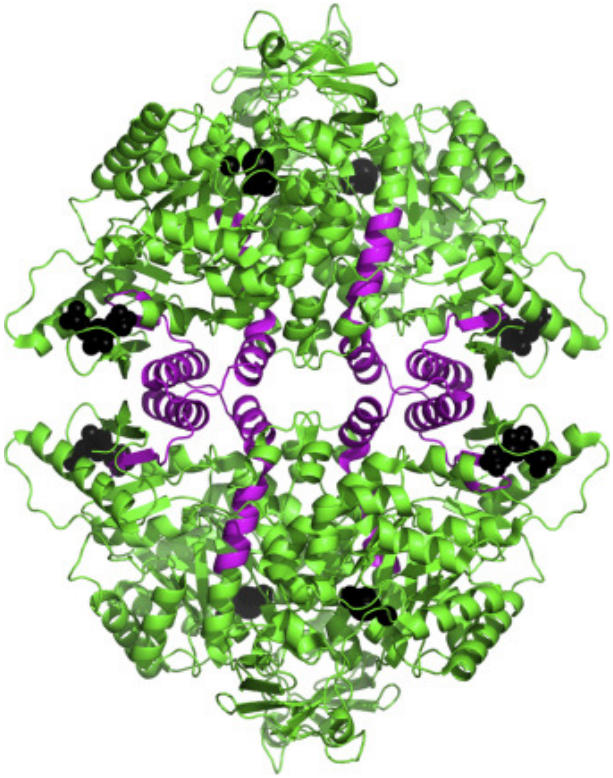


Figure 8.3: Isoform-specific region of PKM2.
Ribbon structure of tetrameric PKM2 with the portions specific to PKM2 colored in magenta. The PKM2-specific residues form the majority of the dimer-dimer interface and a portion of the FBP binding site. Figure taken from [258].

and exhibits positive cooperativity for PEP binding. Affinity for PEP can be increased by FBP binding in a specific pocket in the C domain [264] **(Figure 8.2)**. FBP binding also promotes the tetramerization of PKM2 and stabilizes it in the active tetrameric state. When PKM2 is diluted in the absence of FBP, it reversibly dissociates into dimers and its enzyme activity drops to ∼4% of that found in the active tetramer [265]. PKM1 does not bind FBP due to structural differences in the binding pocket [256]. However, since it is constitutively in the active tetrameric state, PKM1 is insensitive to most of the allosteric regulators of PKM2 [262]. PKM1 and FBP-bound PKM2 have very similar kinetic parameters and are in the same three-dimensional conformation with the exception of the FBP binding site and dimer-dimer interface **(Figure 8.3)** [262, 263].

PKM2 is allosterically regulated by a number of other small molecules and metabo-

lites, which allow coordination of PKM2 activity with metabolic fluxes in pathways not associated with glycolysis. Amino acids including phenylalanine, alanine, and serine are important regulators of PKM2 and bind at a site that is distinct from both the active site and the FBP binding site **(Figure 8.2)**. Structural studies suggest that phenylalanine inhibits kinase activity by stabilizing the tetramer in the inactive T-state, where the active site cleft is locked in an open conformation [262]. In contrast, alanine appears to inhibit PKM2 by promoting dissociation into dimers [266]. Therefore, FBP can only partially activate phenylalanine-bound PKM2 while it can completely overcome alanine inhibition [267, 268]. Serine binds in the same pocket as alanine and phenylalanine, but acts as an allosteric activator of PKM2 [269].

Other allosteric effectors include both endogenous and synthetic small molecules. The thyroid hormone triido-L-thyronine (T3) is an allosteric inhibitor of PKM2 that binds to and stabilizes the inactive monomeric form of the protein [265, 262]. Succinylaminoimidazolecarboxamide ribose-5' phosphate (SAICAR), an intermediate metabolite in the *de novo* purine synthesis pathway, can activate PKM2 with an EC_{50} of 300 μM [270, 271]. PKM2 affinity for SAICAR can be enhanced by ERK-dependant phosphorylation of S37, but both the binding site of SAICAR and how PKM2 conformation is influenced by S37 phosphorylation are unknown [272]. A number of synthetic small molecule allosteric activators have been co-crystallized with PKM2 and show that the activators bind at the subunit interface of the tetramer [273, 274, 84]. This site is distinct from the binding sites of endogenous allosteric activators. Oxalate can modulate PKM2 activity by interacting at the catalytic pocket, either as a competitive inhibitor or as an activator contributing to positive cooperativity of PEP binding at sub-saturating PEP concentrations [256, 275].

PKM2 tetramers can adopt one of two conformations, either the active R-state of the inactive T-state. Allosteric effectors can stabilize the tetramer in one of the two states. For instance, FBP stabilizes PKM2 in the active R-state, while the inhibitor

phenylalanine locks it in the inactive T-state [262]. The two states differ in the position of the monomers such that there is a 13° rotation of the subunits in the T-state with respect to the R-state. Both states are stabilized by salt bridges across the dimer-dimer interface between the C domains, but they differ in the position of the FBP binding loop [262].

8.3.2 Post-translational Modifications and Mutations

A number of post-translational modifications and mutations have also been shown to alter PKM2 activity. For instance, acetylation of K305 inhibits PKM2 kinase activity *in vitro* and acetylation of K433 prevents PKM2 activation by affecting FBP binding [276, 277]. Phosphorylation of Y105 inhibits PKM2 tetramer formation and thus, kinase activity [252]. The R399E mutation has been shown to disrupt tetramer formation, and thus kinase activity, by interfering with one of the dimer interfaces [208]. The mutations H391Y and K422R reduce pyruvate kinase activity in cancer-prone Bloom Syndrome (BS) patients [261, 278, 279, 280].

The H391Y and K422R mutations are of particular interest as they are patient-derived mutations, and thus have clinical implications [278]. *In vitro* characterization of PKM2 carrying these two missense mutations showed that both mutants retained their homotetrameric structure in spite of the fact that these mutations are located in the intersubunit contact domain. However, both mutants show altered kinase activity with respect to the wild-type PKM2 tetramer [261]. For the K422R mutant, it is has been shown that the loss of enzymatic activity results from adoption of the inactive T-state [281]. However, no structural characterization has been performed on H391Y mutant to attempt to elucidate the underlying mechanisms for its biochemical behavior. As compared to wild-type PKM2, the H391Y mutant has a 20% loss in kinase activity thought it shows a 6-fold increase in affinity for its substrate PEP [261]. Additionally, H391Y appears to have reduced cooperativity between subunits and behaves like a non-allosteric protein. The H391Y homodimer also displays enhanced stability and resistance toward structural alteration upon

binding either FBP or phenylalanine [261].

In summary, regulation of PKM2 activity is very complex and involves a variety of different mechanisms including binding of allosteric effectors and post-translational modifications. This suggests that the *in vivo* control of PKM2 activity is extremely dynamic and responds to changes in a number of difference cellular pathways. This is not surprising because of its central role in redirecting intermediate metabolites from the glycolytic pathway to other pathways that promote the synthesis of biomass and ultimately, cell proliferation.

8.4 Aims of this Section of the Thesis

Considering the prevalence of PKM2 in various types of cancer cells and its importance in the Warburg effect, it is highly desirable to understand the molecular mechanisms governing PKM2 activity and regulation. Thus far, high-resolution structural studies have proved invaluable in identifying substrate and regulatory binding sites as well as determining conformational changes that accompany allosteric regulation or post-translational modification.

This section focuses on structural studies of the H391Y patient-derived mutant PKM2 in order to build on the existing structural knowledge about PKM2 while probing a novel target. The wild-type protein was also studied as a comparative reference. The specific aims of this section are:

1. Characterize the effects of the allosteric regulator fructose 1,6-bisphosphate (FBP) and the reducing agent Tris(2-carboxyethyl)phosphine (TCEP) on the oligomeric states of wild-type PKM2.

2. Determine and compare the structures of the wild-type and H391Y mutant forms of PKM2 using single particle cryo-electron microscopy.

3. Analyze protein flexibility based on resolution variations, relaxing symmetry constraints, and comparison to previously reported crystal structures.

Chapter 9

Methods

9.1 Strains, Plasmids, and Growth Conditions

The plasmid for PKM2 expression, pET28aLIC-PKM2 was a gift from Dr. Cheryl Arrowsmith (Addgene plasmid #25360) and was amplified in the *E. coli* general cloning strain NEB 5-alpha (New England Biolabs). pET28aLIC-PKM2 is a high copy plasmid that confers kanamycin resistance and includes PKM2 under the control of the IPTG-inducible T7 promoter. *E. coli* One Shot BL21(DE3) cells (Invitrogen) were used for high-level inducible PKM2 expression from the T7 promoter.

The H391Y point mutation was introduced into pET28aLIC-PKM2 using the appropriate primers listed in **Appendix B** and the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies). The procedure was done as per the manufacturer's instructions. The modified plasmid was retrieved by growing transformed cell colonies in liquid culture **(Section 3.1.2)** and performing a plasmid purification **(Section 3.2.1)**. The presence of the desired mutation was confirmed by sequencing **(Section 3.2.1)**.

E. coli growth conditions, antibiotic concentrations and storage of cell strains were all performed as described in **Section 3.1**.

9.2 Protein Expression and Purification

9.2.1 Overexpression in *E. coli*

BL21(DE3) cells transformed with the expression plasmid for either wild type or mutant PKM2 were scraped off of plates as described in **Section 3.3.1** and diluted in Super Broth (KD Medical) supplemented with kanamycin. The dilution was done such that the initial OD_{600} was ∼0.1. Cells were grown at 37°C with 225 rpm shaking to an OD_{600} of ∼0.8 and were subsequently cooled with 225 rpm shaking to 17°C. Cells were induced with 0.5mM IPTG and the induction was allowed to proceed for 16-18 hours. Cells were harvested by centrifugation for 10 min at 6,000 x *g* at 4°C and cell pellets were stored at -80°C.

9.2.2 Protein Purification

PKM2 protein purification was implemented in collaboration with Dr. Soojay Banerjee (Subramaniam laboratory) based on the previously described protocol [262]. Dr. Banerjee performed affinity chromatography and size exclusion chromatography while all other steps were performed independently. Cell pellets for each 1L of culture were thawed on ice and resuspended in 30mL of Buffer A containing 1 complete ULTRA EDTA-free tablet (Roche) and lysed using high pressure homogenization. The cell lysate was centrifuged at 23,000 x *g* at 10°C for 45 min and was passed through a 0.2 µm syringe filter to remove cell debris and particulates. A cobalt affinity chromatography column was prepared by placing 1 mL of HisPur Cobalt Resin (Thermo Scientific) into a chromatography column and allowing to settle for 10 min. The column was equilibrated with 5 mL of buffer A and the cleared cell lysate was applied to the column and run through. The flow through was collected and analyzed by SDS-PAGE to check for the loss of protein. The column was washed with 60 mL of of Buffer A and protein was eluted using 4 mL of Buffer B, collected in 1 mL fractions. PKM2 usually eluted in the first 3 fractions, with the second fraction having the highest protein concentration.

Size exclusion chromatography (SEC) was performed to remove further impurities and analyze the oligomeric state of the protein. Protein eluted from the Co^{2+} column was loaded onto a either a Superdex 200 10/300 GL column (GE Healthcare Life Sciences) or a Superose 6 10/300 GL column (GE Healthcare Life Sciences) attached to an ÄKTA system. The column was pre-equilibrated with Dulbecco's PBS without calcium or magnesium (Sigma), sometimes containing 1mM Fructose 1,6-bisphosphate (FBP) and 5mM Tris(2-carboxyethyl)phosphine (TCEP) as described in **Section 10.1**. The protein solution was run through the column with a flow rate of 0.25mL/min (Superdex 200) or 0.3 mL/min (Superose 6) and 0.5mL fractions were collected. Peak fractions for the tetramer (B2 - B4 for Superdex 200 and C4 - C6 for Superose 6) were collected and analyzed by SDS-PAGE and NativePAGE. Fractions B2, B3, and B4 correspond to elution volumes of 11 - 11.5mL, 10.5 - 11mL, and 10 - 10.5mL, respectively. Fractions C4, C5, and C6 correspond to elution volumes of 13.5 - 14mL, 14 - 14.5mL, and 14.5 - 15mL, respectively.Protein concentration of the selected fractions was measured using absorbance at 280nm as measured by a NanoDrop 2000 Spectrophotometer (Thermo Scientific). The C6 fraction was split into 5-10 μL aliquots, flash frozen in liquid nitrogen and stored at -80°C.

9.2.3 Validating Protein Expression and Quality

The presence of PKM2 in whole cell lysates and purified protein samples was indicated by denaturing or native polyacrylamide gel electrophoresis, performed as described in **Section 3.3.4**.

9.3 Microscopy and Image Analysis

9.3.1 Negative Staining

In order to prepare grids with a continuous carbon film, a thin layer of carbon was deposited onto a clean surface of freshly cleaved mica using a carbon evaporator. The carbon was subsequently floated off of the mica in MilliQ water and deposited on copper mesh grids. Carbon coated grids were dried overnight under a heat lamp and stored in a dessicator until use. Prior to use, grids were cleaned with a glow discharge for 15 seconds. Grids were held with anti-capillary tweezers and blotted from the side with Whatman filter paper after each of the following steps. Care was taken not to allow the grids to dry between steps. Grids were incubated for 2 min with 2 μL protein solution (thawed on ice and diluted to a concentration of 0.01 mg/mL in Dulbecco's PBS with 1mM FBP), washed twice with 5 μL NanoW stain (pH 6.8), and then incubated for 2 min with 5 μL NanoW. After this final incubation, the grid was blotted and air dried for at least 5 minutes before imaging. Grids were imaged on a FEI Tecnai 12 electron microscope as described in **Section 3.4.3**.

9.3.2 Preparation of Vitreous Specimens

Prior to sample preparation, protein aliquots were thawed on ice and were combined with solutions of NaCl, TCEP, and octyl glucoside (OG) in order to obtain the final concentrations listed in **Table 9.1**.

Table 9.1: Components in PKM2 cryo-EM samples. All components except OG were in buffer (Dulbecco's PBS with 1mM FBP). OG was added from a 1% stock solution dissolved in MilliQ water.

Component	Quantity
Protein	1.5 mg/mL
TCEP	5mM
NaCl	150mM
OG	0.025%

Specimens were prepared using a Mark IV Vitrobot (FEI Company) as described in **Section 3.4.2** with the following exceptions. Samples were prepared using 200 mesh Quantifoil R1.2/1.3 holey carbon grids (Quantifoil Micro Tools) that were cleaned by glow discharge for 3 cycles of 15 seconds each.

9.3.3 Cryo-electron Microscopy

Specimens were imaged using a Titan Krios electron microscope (FEI Company) aligned for parallel illumination and operated at 300 kV. Grids were maintained at liquid nitrogen temperatures throughout transfer to the microscope and imaging. The microscope was equipped with a K2 Summit camera (Gatan) placed at the end of a GIF Quantam energy filter (Gatan), operated in zero-energy-loss mode with a slit width of 20 eV. Images were collected in super-resolution counting mode with a physical pixel size of 0.637Å, a dose rate of ∼2.4 e^-/pixel/s (∼5.9 $e^-/Å^2$/s at the specimen plane), and a total dose of ∼45 $e^-/Å^2$ spread over 38 intermediate frames.

9.3.4 Image Processing

The datasets for the PKM2 wild-type ($PKM2^{WT}$) and H391Y mutant ($PKM2^{H391Y}$) consisted of 940 and 1,654 micrographs, respectively. The images were collected 0.75 to 2.5 µm underfocus and displayed Thon rings extending beyond ∼3Å. In order to correct for drift and beam induced motion, individual movie frames were aligned to the cumulative average of the previous frames as previously described [42, 282]. Contrast transfer function (CTF) estimation was performed on each micrograph using radially averaged power spectra from periodogram averaging as described previously [282]. The maximum resolution used to estimate the defocus with TOMOCTFFIND was 2.75Å.

9.3.5 Structure Refinement and Analysis

Initial Model

A 9Å resolution structure of $PKM2^{WT}$ was generated and used an a reference for subsequent particle picking and assignment of orientations. This structure was determined from 1,432 images collected on a FEI Titan Krios equipped with high-brightness XFEG and Cs corrector. The microscope was aligned for parallel illumination and operated at at 300kV. Images were using EPU automated data-acquisition software using a Falcon II CMOS detector with a pixel size of 1.406Å in the specimen plane. Images were collected 2.0 to 3.5 µm underfocus with a dose rate of 20 $e^-/Å^2/s$ and exposure time of 3.5 seconds split over 7 frames. Particles (62,207) were picked automatically using a Gaussian disk of 80Å radius and extracted from the aligned images without binning with a box size of 256 x 256 pixels. Using the RELION version 1.3 package (Scheres laboratory), particles were iteratively classified for 25 iterations in both 2D and 3D with D2 symmetry imposed [283]. Particles in low-resolution or uninterpretable classes were removed at each step. The remaining 24,843 particles were processed using the RELION auto-refine routine until convergence and were B-factor corrected with a value of -25 [73]. The 3D reference used for orientation assignment was the crystal structure 4B2D low-pass filtered to 60Å.

Structure Determination and Visualization

Table 9.2: Data Collection and Image Processing Statistics.

Sample	Micrographs	Number of Auto-picked Particles	Symmetry Imposed	Particles in 3D Reconstruction
$PKM2^{WT}$	940	77,786	C1	39,758
			D2	38,938
$PKM2^{H391Y}$	1,654	131,471	C1	80,359
			D2	56,407

Particles were picked automatically from the $PKM2^{WT}$ and $PKM2^{H391Y}$ datasets us-

ing template matching with the projections of the 9Å initial model. Particles were extracted with a binning factor of 2 and a box size of 384 x 384. Particles were then subjected to 3D classification for 25 iterations with either C1 or D2 symmetry using the RELION version 1.4 package [283]. The best class for each was selected and subsequently processed with the same specified symmetry as the previous step using the RELION auto-refine procedure until convergence [73]. The intermediate frames for each particle were then weighted using RELION's movie processing and particle polishing. Subsequently, the particles were subjected to another round of 3D classification in 10 or 20 classes without alignment and the subset of classes that yielded the best resolution refinement was selected. After auto-refinement in RELION, maps were masked, B-factor corrected by an automatically determined value, and sharpened using the modulation transfer function (MTF) of the detector as implemented in the RELION package post-processing.The 9Å initial model was low-pass filtered to 50Å and 20Å respectively for 3D classification and auto-refinement. The EMAN2 program suite (Ludke laboratory) was used for reference-free 2D classification of the particle stack before and after 3D classification, although this was not used to eliminate any particles [284]. The number of particles included in each 3D structural reconstruction is listed in **Table 9.2**. Structures were visualized and figures were produced using UCSF Chimera (Resource for Biocomputing, Visualization, and Informatics) [285] and COOT (Paul Emsley) [286]. Models were generated from automated real space refinement of the crystal structure coordinates 4B2D in the electron density maps. The real space refinement was performed using PHENIX [287]. Local resolution estimation was performed using ResMap (Alp Kucukelbir) [288].

Chapter 10

Biochemical and Structural Characterization of Pyruvate Kinase M2 (PKM2)

In an effort to gain mechanistic insights into the differences in activity between the wild-type and H391Y patient mutant forms of PKM2, the structures of both forms were determined by single particle cryo-EM. The main goal was to compare the wild-type and mutant structures to identify possible differences in the vicinity of the mutation site and in the overall architecture of the complex. In addition, the effects of the presence of the allosteric regulator fructose 1,6-bisphosphate (FBP) and the reducing agent Tris(2-carboxyethyl)phosphine (TCEP) were analyzed in the context of the proteins oligomeric state.

10.1 PKM2 Purification and Biochemical Characterization of Oligomeric State

Wild-type PKM2 ($PKM2^{WT}$) was expressed and purified from pET28aLIC-PKM2 in BL21(DE3) cells as described in **Sections 9.2.1 and 9.2.2**. For the production of the H391Y mutant ($PKM2^{H391Y}$), the required point mutation was introduced

into pET28aLIC-PKM2 by site-directed mutagenesis. Subsequently, the same procedures for expression and purification were used. Briefly, protein expression was induced by the addition of IPTG at an OD_{600} of ∼0.8 and the induction was allowed to proceed overnight at 17°C. After cells were harvested by centrifugation and homogenized, cell debris and membranes were removed by centrifugation and cell lysates were applied to a Co^{2+} column for His-tag affinity purification. Size exclusion chromatography (SEC) was then used to isolate tetramer peak fractions and to analyze the oligomeric states of PKM2 under the given purification conditions.

Effect of FBP and TCEP on Oligomeric State of the PKM2

SEC chromatography of PKM2 was performed under different conditions that varied by the presence or absence of the allosteric activator FBP and the reducing agent TCEP. The presence of 1mM FBP during SEC reduced the prevalence of dimers with a proportional increase in tetrameric species **(Figure 10.1a)**. This is consistent with previous evidence that PKM2 undergoes reversible dissociation into dimers when diluted in the absence of FBP [265]. This dissociation and resulting inactivation of PKM2 are believed to allow dynamic regulation of its enzymatic activity in the cellular context [258]. After storage at 4°C for 3 days in the presence of FBP and the absence of TCEP, PKM2 spontaneously forms higher order oligomers **(Figure 10.1b)**. PKM2 has a number of exposed cysteine residues which may result in oligomerization if they are oxidized to form disulfide bonds. Addition of 5mM TCEP to samples that show evidence of oxidation restores the protein to the expected oligomeric state **(Figure 10.1b)**. Though oxidation wasn't a concern on the timescale of cryo-EM sample preparation, subsequent purifications and imaging were performed in the presence of 1mM FBP and 5mM TCEP in order to stabilize tetramers and prevent the formation of higher order oligomers or aggregates. Additional characterization was performed by step-wise removal of FBP and TCEP after purification in the presence of both compounds. As shown in **Figure 10.1c**, removal of FBP alone by dialysis was not sufficient to lead to dissociation into dimers. Subsequent removal of TCEP and incubation in the absence of the reducing agent

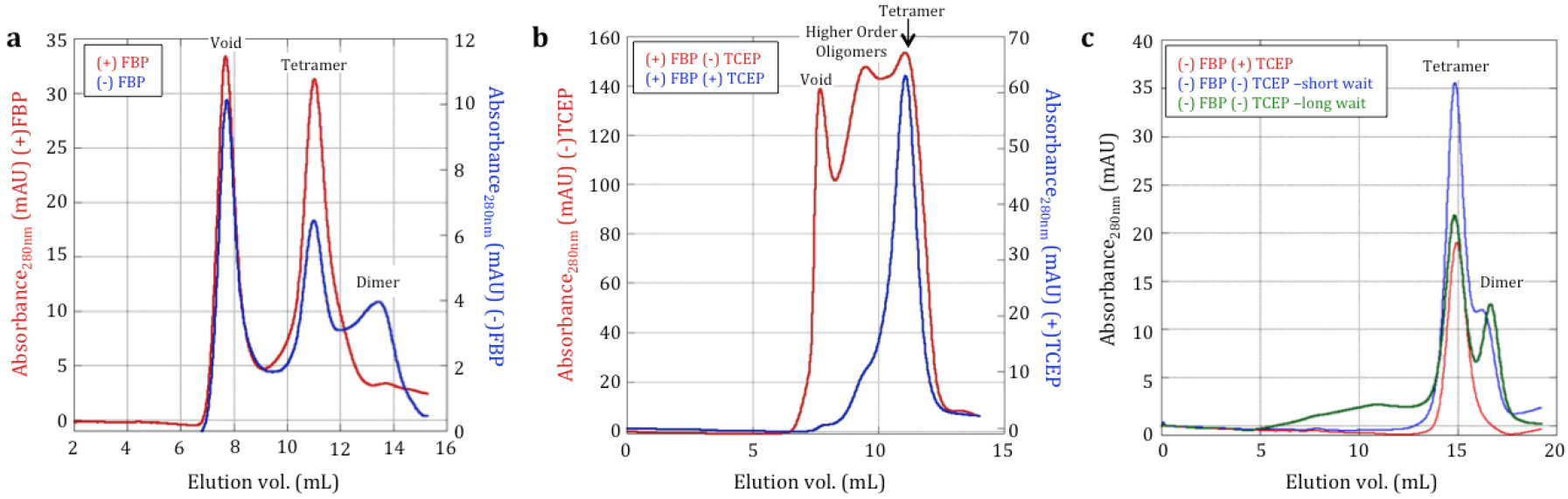


Figure 10.1: SEC analysis of the effect of FBP and TCEP on PKM2 oligomeric state.
a) SEC profile of PKM2 indicating that the presence of FBP stabilizes the tetrameric form and prevents dissociation into dimers. These samples were run immediately after affinity purification was performed. b) In the presence of FBP, the SEC profile shows that the addition of TCEP prevents, and in fact can reverse, the formation of higher order oligomers. Both samples in this panel were stored for 3 days at 4°C and differ in the fact that the sample with TCEP was incubated for an additional day at 4°C in the presence of 5mM TCEP. c) SEC profile showing the effect of step-wise removal of FBP and TCEP. Dissociation into dimers required both the absence of the the allosteric activator as well as incubation of the protein in an oxidizing environment. Incubations were ~2 hours (blue) and overnight (green). The elution volume is different from panels A & B because a Superose 6 column was used instead of a Superdex 200 column.

was required in order for dissociation to occur. This agrees with previous studies which have shown that in the absence of an allosteric activator, oxidation of the reactive cysteine at position 358 induces dissociation [237]. *In vivo*, inactivation of PKM2 by oxidation-induced dissociation serves a feedback mechanism in order to allow cancer cells to regulate the levels of reactive oxygen species [238].

10.2 Determination of Structure of Wild-Type and H391Y Mutant Forms of PKM2

Cryo-electron microscopy was used to image vitrified specimens of $PKM2^{WT}$ and $PKM2^{H391Y}$ as described in **Section 9.3.3 (Figure 10.2)**. Intermediate image frames were aligned and averaged in order to correct for drift and beam induced motion of the sample **(Section 9.3.4)**. Particles were identified automatically in aligned frame averages using template matching with a 9Å initial model. Extracted particles were subjected to classification, refinement, masking and sharpening as described in **Section 9.3.5** in order to obtain D2 symmetrized structures of $PKM2^{WT}$ and $PKM2^{H391Y}$ with reported resolutions of 3.4Å **(Figure 10.3)**.

Atomic models for both structures were generated using real space refinement of the coordinates from 4B2D in the electron density maps [289]. Since previous biochemical studies studies demonstrate significant differences in the regulation by FBP of the wild-type and mutant, the differences in FBP binding in both structures were first examined [261]. The density for FBP was readily resolvable in both maps **(Figure 10.4)**. In both maps, the FBP appears to bind in the same position and same orientation. Although the real space refined atomic models demonstrate some differences in the puckering of the sugar ring and the rotation of the phosphate, the current resolution of the maps is not sufficient to determine whether these are real differences of simply artifacts of the automated real space refinement procedure.

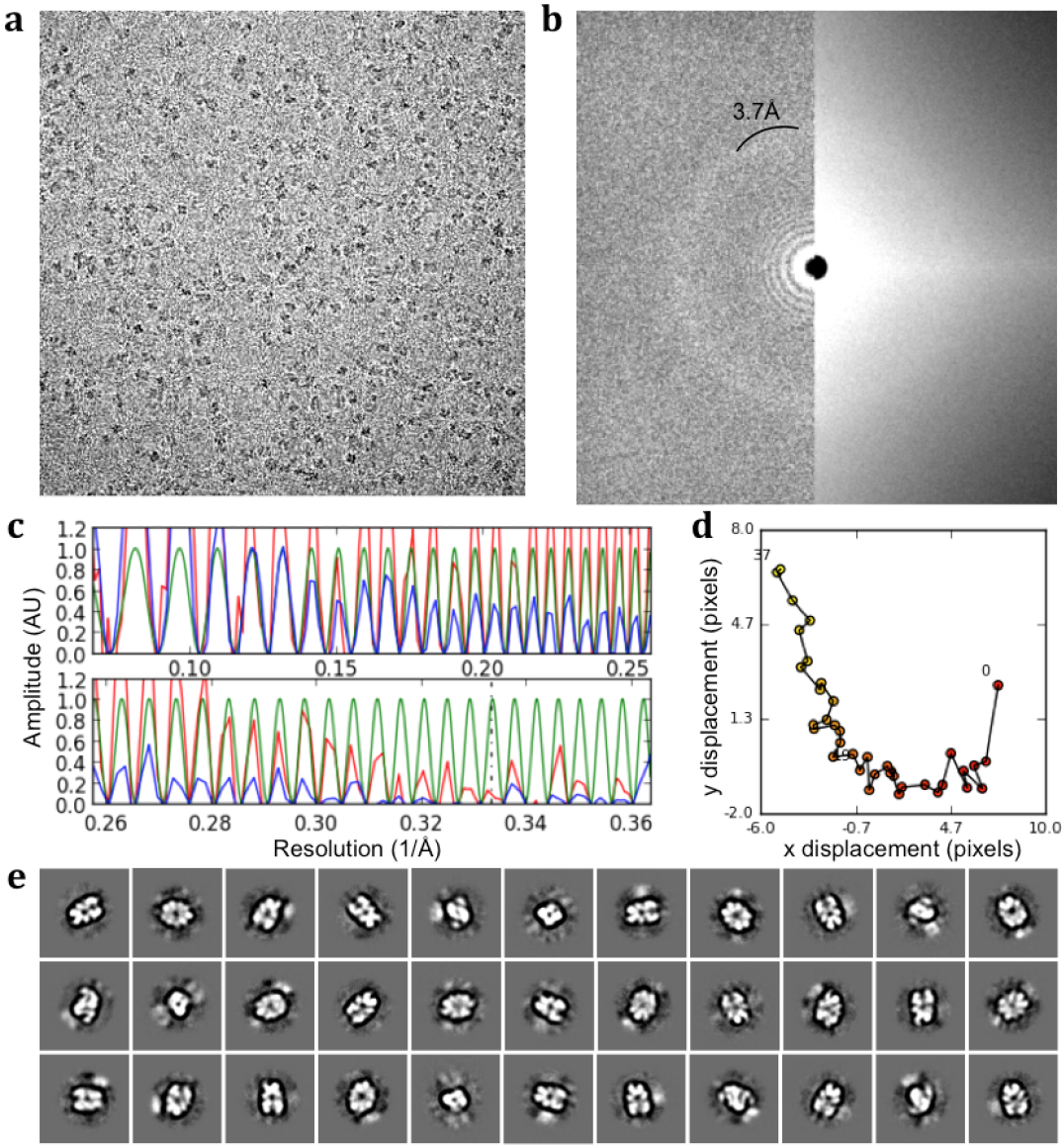


Figure 10.2: Cryo-EM of vitrified specimens of PKM2^{H391Y}.
a) Representative average of 38 aligned intermediate image frames acquired over a 7.6 second exposure with a dose rate of 5.9 $e^-/Å^2/s$. Images were collected at 300kV in super-resolution mode with a physical pixel size of 0.637Å. b) FFTs of the image in (a) showing the Thon rings obtained from the average of the aligned frames (left) and the raw individual frames (right). c) 1D profile of the power spectrum showing signal extending past 3Å, indicated by the dashed line. Green is indicative of the CTF at the estimated defocus of the image (1.9 µm), red is indicative of the radial CTF profile of the frame average, and blue is indicative of the radial CTF profile of the individual frames. d) Plot of the movement of the 38 frames during exposure used to correct drift for the image in (a). Displacements are measured in terms of the super-resolution pixel. e) Averages from reference-free 2D classification showing various views of PKM2.

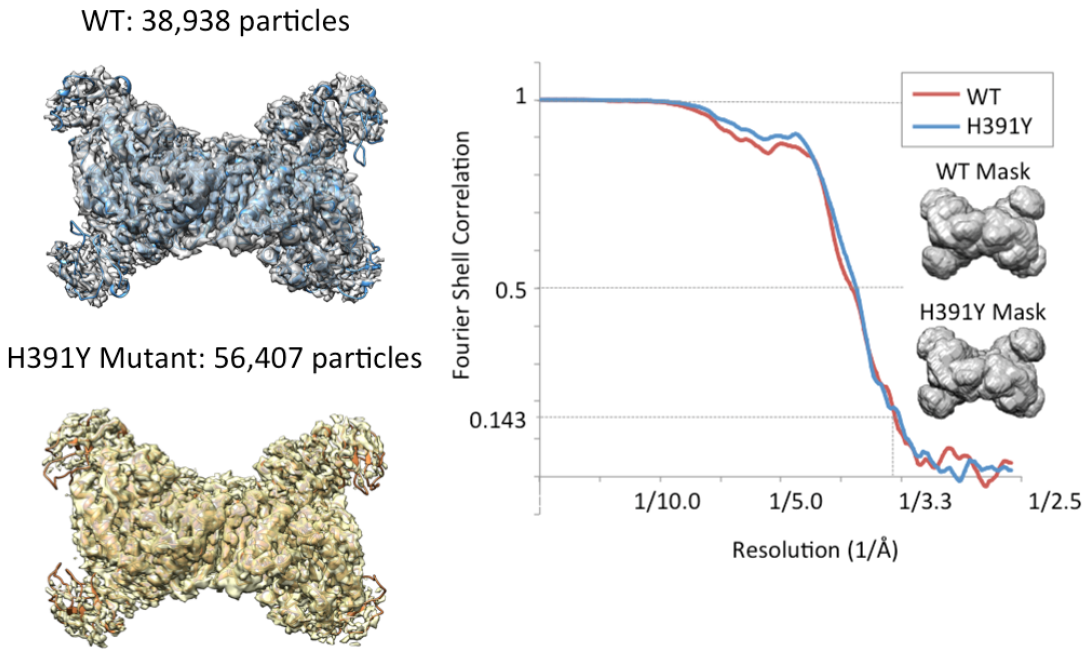


Figure 10.3: Symmetrized cryo-EM maps of PKM2^{WT} and PKM2^{H391Y}. Density maps of the WT and H391Y mutant of PKM2 determined at resolutions of 3.4Å. Molecular models generated automatically in PHENIX from PDB ID 4B2D were placed in the densities. Also shown is the Fourier Shell Correlation (FSC) curve between two semi-independently refined halves of the data for the WT (red) and H391Y mutant (blue) showing their reported resolutions of 3.4Å. FSC calculations were done using the soft shape maps showed as insets.

Since there were no significant differences in the FBP position or orientation, the site of the mutated residue was examined in order to determine potential structural mechanisms for the previously reported biochemical differences the the wild-type and H391Y mutant forms of PKM2. The difference between the densities for residue 391 in both maps, including the hydroxyl density for the mutant tyrosine, were clearly resolvable **(Figure 10.5)**. Close examination of atomic models in the immediate environment of the mutation site revealed a difference in the position of arginine 447. Based on this difference, there may be a cation-π interaction that forms with the mutant tyrosine, but not with the mutant histidine. A stabilizing interaction of this sort near the subunit interface could potentially be propagated in order to explain the losses of allosteric regulation and cooperativity demonstrated by the H391Y mutant. However, based on the density for arginine 447, the orientation of the

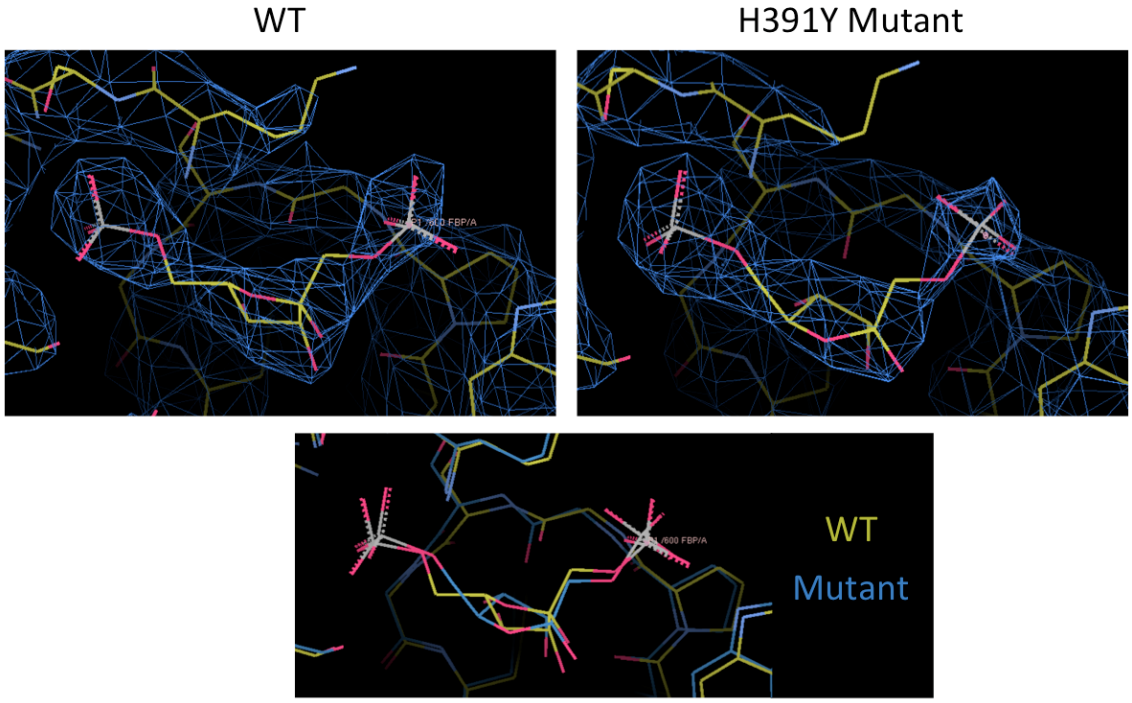


Figure 10.4: Comparison of FBP density in wild-type and mutant PKM2. FBP binds density is found in the same orientation in both the wild-type and mutant protein. The resolution of the map is not sufficient to conclude that the differences between the tow atomic models are actual differences between the two structures.

side chain at this position cannot be definitively assigned. Thus, further structural and mutational studies must be performed in order to either prove or disprove the existence of a stacking interaction between arginine 447 and the mutant tyrosine 391.

There were no global changes in architecture observed when comparing the maps of the wild-type and H391Y mutant forms of PKM2. Notably, this does deviate from previous circular dichroism (CD) spectroscopy results that found a significant increase in the α-helical content of $PKM2^{H391Y}$ relative to $PKM2^{WT}$ (67% and 38%, respectively) [261]. Since the protein purification procedure was similar between the two studies, the most likely source of this discrepancy is the intrinsic limitations of CD spectroscopy in accurately determining secondary structure content for proteins that are not all α-helical or whose β-strand geometry is not canonical [290].

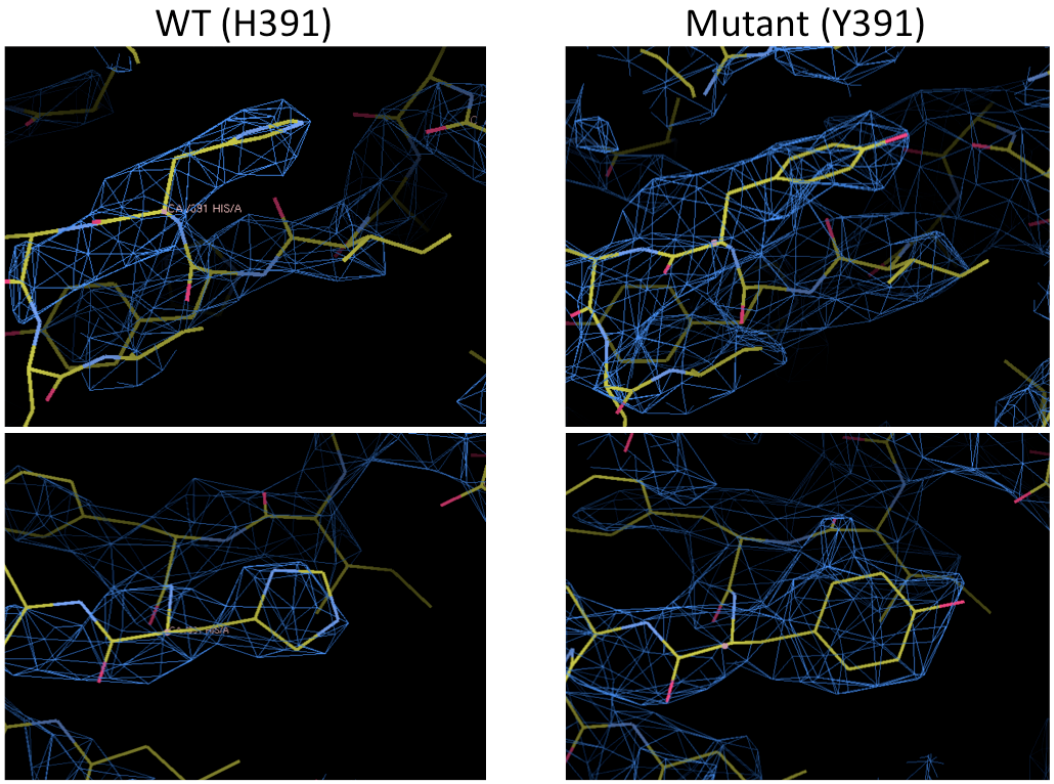


Figure 10.5: Comparison of electron density at residue 391 in wild-type and mutant PKM2.
Electron densities show the difference between the wild-type histidine and mutant tyrosine at the site of the mutation.

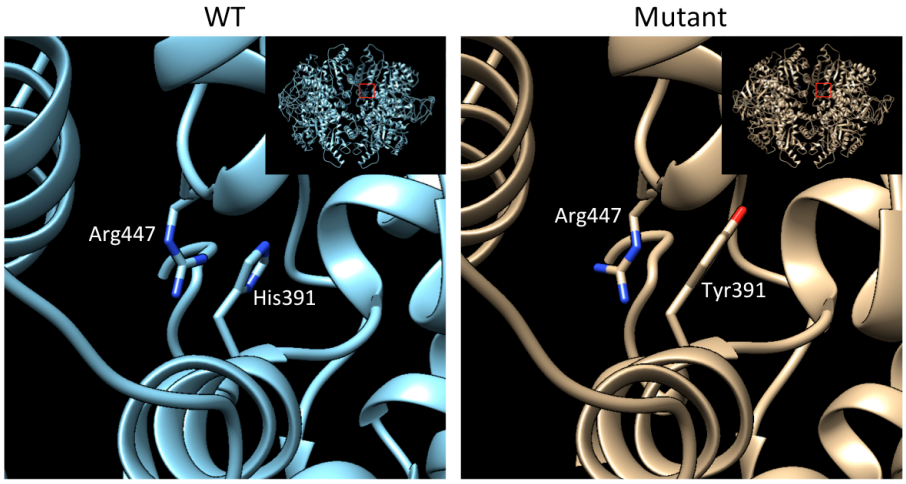


Figure 10.6: Potential stacking interaction formed in the mutant.
The differences in the orientations of arginine 447 in the atomic models suggest that the mutant may be forming a cation-π stacking interaction between the mutant tyrosine at position 391 and the arginine at position 447. The insets show the position of the pictured interaction in one subunit of the full protein.

10.3 Flexibility of the Domains with Monomers

Local resolution estimation of the electron density maps using ResMap demonstrated that the B domains were resolved at lower resolution than other portions of the structure **(Figure 10.7)**. This was indicative that these regions may may have a greater range of motion than the remainder of the protein. This is consistent with previous studies that have demonstrated movement of the B domain in order to restrict access to the active site once the Mg^{2+}-ADP substrate complex is bound [255]. In order to further probe the potential flexibility of these regions and differences that may exist

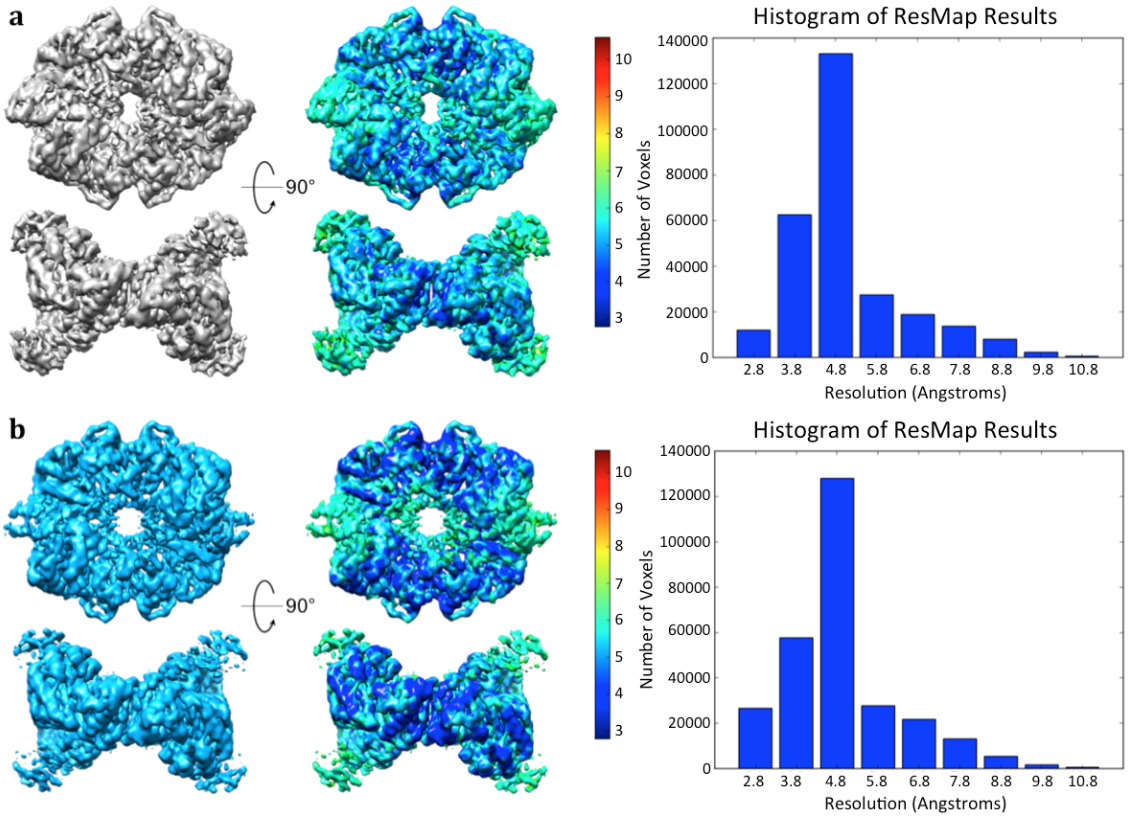


Figure 10.7: Local resolution estimation for symmetrized cryo-EM maps of $PKM2^{WT}$ and $PKM2^{H391Y}$.
Density maps colored by estimated local resolution and corresponding resolution histograms for $PKM2^{WT}$ (a) and $PKM2^{H391Y}$ (b) show a decrease in resolution in the B domains relative to the remainder of the protein. Estimations were done over a range of 2.8 to 10.8Å with a step size of 1Å. As has been previously noted in the literature, the absolute resolution assignments from the local resolution estimation reported by ResMap should be conservatively interpreted [59, 291].

in their position between the four monomers, PKM2^{WT} and PKM2^{H391Y} structures were processed as described in **Section 9.3.5** without symmetry imposed. The resulting structures were determined at reported resolutions of 4.4Å and 5.1Å for the wild-type and H391Y mutant, respectively **(Figure 10.8)**. The local resolution estimates once again showed that the lowest resolution areas were concentrated at the periphery of the protein and particularly in the B domains **(Figure 10.9)**.

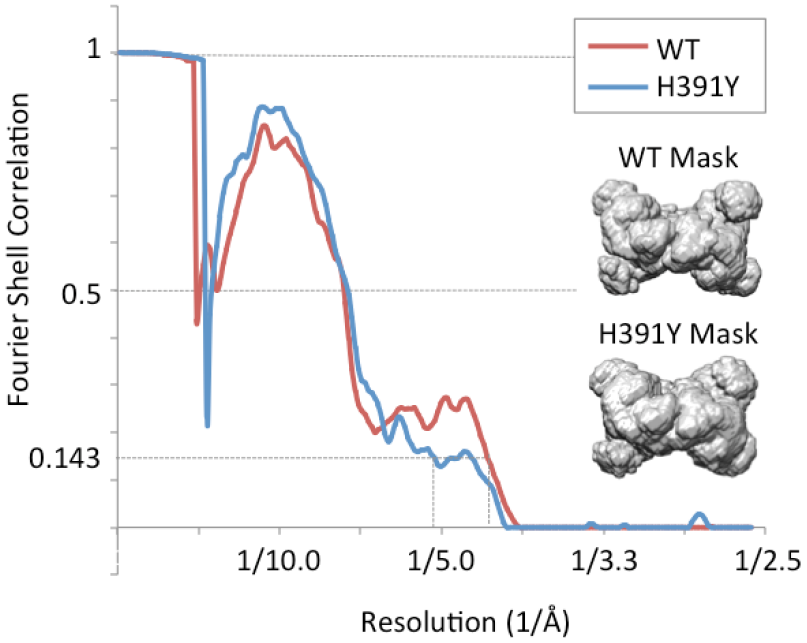


Figure 10.8: Resolution estimation for asymmetric cryo-EM maps of PKM2^{WT} and PKM2^{H391Y}.
Fourier Shell Correlation (FSC) curve between two semi-independently refined halves of the data for the WT (red) and H391Y mutant (blue) without applying symmetry. The shapes of the curves indicate that the resolutions may be lower than the reported values of 4.4Å and 5.1Å for the WT and H391Y structures, respectively. FSC calculations were done using the soft shape maps showed as insets.

In order to visualize the degree of flexibility between different domains, the coordinates for each individual monomer were split into three regions as follows: A domain (residues 44-116 & 219-389), B domain (residues 117-218), and C & N domains (residues 1-43 & 390-531) [292]. For a single monomer, each region was placed into the corresponding electron density using a rigid body fit in Chimera.

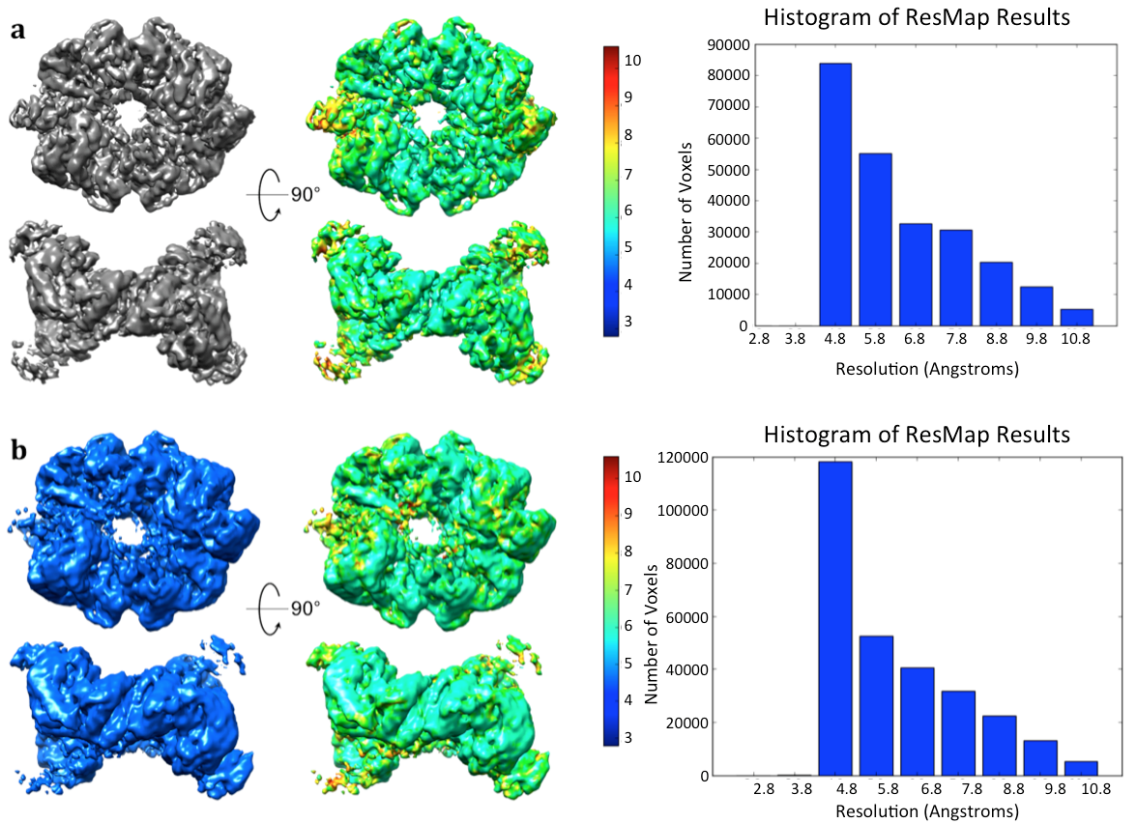


Figure 10.9: Local resolution estimation for asymmetric cryo-EM maps of PKM2$^{\mathbf{WT}}$ and PKM2$^{\mathbf{H391Y}}$.
Density maps colored by estimated local resolution and corresponding resolution histograms for PKM2$^{\mathrm{WT}}$ (a) and PKM2$^{\mathrm{H391Y}}$ (b) show a decrease in resolution in the B domains relative to the remainder of the protein. Estimations were done over a range of 2.8 to 10.8Å with a step size of 1Å.

This was repeated for each of the four asymmetric monomers. A model for each complete monomer was then generated by combining the three independently fitted regions. These models were subsequently aligned to each other using the density map of a single monomer as a reference. This analysis was performed for both the wild-type and H391Y mutant in order to visualize the degree of asymmetry between the monomers of each structure. While both the wild-type and the H391Y mutant demonstrate flexibility, this analysis suggests that the mutant has a greater range of motion between the domains in each monomer **(Figure 10.10)**. Since residue 391 is at the interface of the A and C domains, it has previously been proposed that the H391Y mutant may alter the dynamics of subunit movement [261]. However, further structural studies at higher resolution are needed to quantify the degree of

flexibility within monomers of the mutant protein.

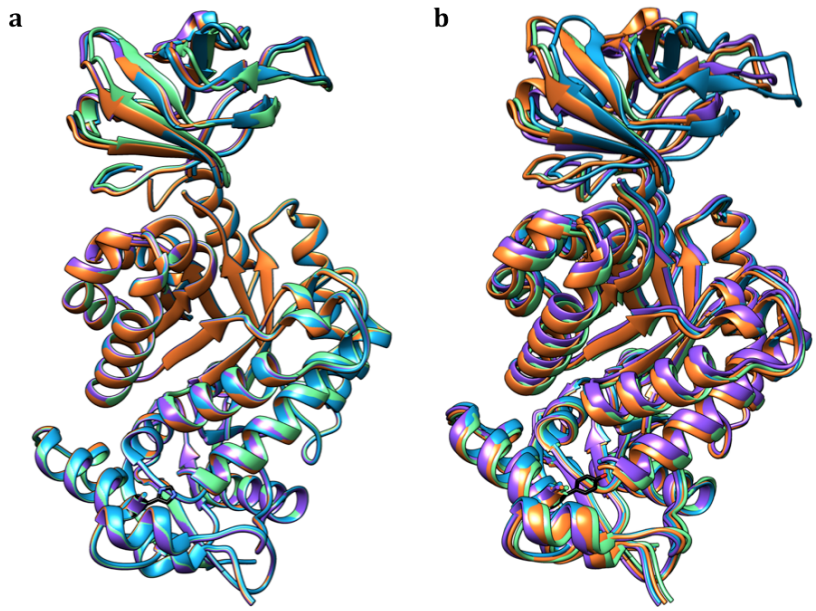


Figure 10.10: Domain flexibility in asymmetric structures of PKM2^{WT} and PKM2^{H391Y}.
Comparison of the aligned models created for the four monomers in the wild-type (a) and H391Y mutant (b) asymmetric structures showing greater flexibility of the mutant domains. Models for each monomer were created by splitting the coordinates into three sections (A: 44-116 & 219-389, B: 117-218, C & N: 1-43 & 390-531), fitting them into the corresponding electron density for each monomer, and then combining them. Each of the four chains from the four subunits is labeled in a different color.

10.4 Discussion

This study utilizes structural studies to characterize PKM2 and probe structural differences between the wild-type and H391Y mutant forms of the protein. SEC chromatography was used to show that the allosteric activator FBP prevents dissociation into dimers and that a reducing environment prevents or reverses the formation of higher order oligomers. Interestingly, under reducing conditions, the removal of FBP alone was not sufficient to cause tetramer dissociation. Subsequent removal of the reducing agent TCEP and overnight incubation under oxidizing conditions was required in order to observe dissociation into dimers. Thus, reducing

conditions can preserve the tetrameric state of PKM2 even in the absence of an allosteric regulator. However, upon extended exposure to oxidizing conditions without FBP, PKM2 dissociates into dimers. This most likely is a result of oxidation at C358, a phenomenon that has been previously described as a mechanism for PKM2 inactivation both *in vitro* and *in vivo* [237, 238].

In order to investigate the differences between the wild-type and H391Y patient-derived mutant forms of PKM2, cryo-EM structures were determined with D2 symmetry at a resolution of 3.4Å. The electron density maps allowed us to resolve the position of the bound allosteric regulator FBP as well as the density difference between the two structures in the mutation site. Since PKM2 is only 240kDa, it is a notable achievement that its structure can be solved to near-atomic resolution by cryo-EM. To date, only one other structure of a protein less than 250kDa has been solved at a resolution better than 3.5Å [293]. The difficulties in solving cryo-EM structures from small particles arise because smaller particles contribute less signal for each individual particle image, resulting in greater difficulty in accurate particle alignment and orientation assignment [41]. Though iterating through the process with a large number of particles can improve the outcome by improving the signal to noise ratios of the references, reaching near-atomic resolution for small proteins still remains challenging. In the case of PKM2, the problem is further exacerbated by flexibility of the outer domains. Variations in the positions of these external domains can interfere with accurate alignment if different portions of the protein align slightly differently with the reference. In spite of these methodological challenges, structures of both the wild-type and H391Y mutant were determined with sufficient information to resolve both the small molecule regulator, FBP, and the difference in a single residue mutation.

Examination of the structures reveals that there is no significant difference in the position of orientation of the FBP that can be conclusively determined from the maps at their current resolution. This fails to explain the previous biochemical observation that FBP fails to regulate affinity for the substrate in the mutant [261].

Further comparison of the atomic models for the wild-type and mutant proteins show that although the backbone position at the mutation site is nearly unchanged, there might be a change in the orientation of arginine 447. This may be indicative of the arginine forming a cation-π stacking interaction with the mutant tyrosine residue at position 391. Although the density for the arginine is not at a high enough resolution to definitively identify the orientation of the side chain, this potential stacking interaction may explain the loss of cooperativity previously observed in the mutant if it it is propagated to the subunit interface [261].

Another potential functional implication of a stacking interaction with the mutant tyrosine is that it may affect the communication between domains since it is located at the interface of the A and C domains. The A domain is directly connected to the active site while the C domain contains the regulatory binding site for FBP [258]. Thus, a stabilizing stacking interaction at the mutation site may be able to explain the lack of FBP regulation on substrate binding observed in the mutant [261].

In order to be able to definitively identify the orientation of the arginine residue that may be forming the stacking interaction or be able to conclusively determine other potential interactions, the resolution of both maps must be at least higher than 3Å. As explained above, this is extremely difficult for a small protein like PKM2 that contains flexible domains. However, the presence of this proposed stacking interaction could be probed by mutational analysis and subsequent biochemical testing determine the effects on FBP regulation and substrate binding. For example, a mutation of residue 447 from an arginine to a leucine (similar in size, but neutral in charge) in the mutant protein can be tested. If wild-type activity is restored by the incorporation of this additional mutation, it will be indicative that the proposed stacking interaction is indeed the reason for the H391Y mutant's differential biochemistry.

Another factor that must be considered is that the mutation from histidine to tyrosine at position 391 may also be involved in disrupting existing interactions that

occur with the wild-type histidine. One such interaction that has been previously proposed is with glutamate 386 [261]. Therefore, an additional mutant that may be useful to study is one where the 391 position is changed to a phenylalanine since that can simultaneously allow formation of a cation-π interaction with arginine 447 without disrupting the interaction with glutamate 386.

Since the active site is located in the cleft between the A and B domains, the position of the B domain is likely of great functional importance as it controls access to the active site [258, 256]. Local resolution estimates performed on the symmetrized maps of both the wild-type and the mutant revealed that the B domains were the lowest resolution portions of the map. This was an indication that there may be greater flexibility in these domains, which results in lower resolution when averaged. This is consistent with previous studies which have shown that the B domains have a large range of motion compared to the remainder of the protein in order to regulate access to the active site based on its occupancy [255, 262]. In order to further examine variations between monomers that may result from flexibility, wild-type and mutant maps were determined without imposing any symmetry constraints. Though the maps were reported to have resolutions of 4.4Å and 5.1Å for the WT and H391Y structures, respectively, the map features and FSC curves indicated that the resolution may be lower.

Analysis of the flexibility of the regions within each monomer suggested that the mutant protein may have a greater range of motion than the wild-type. Greater flexibility of the mutant is consistent with the fact that when symmetry was imposed, the mutant density map required a much larger number of particles in the final reconstruction than the wild-type map in order to reach the same resolution. In addition, the asymmetric map was lower in resolution than the wild-type map despite having a larger number of particles included in the final reconstruction. However, care must be taken in interpreting these observations as they may also be indicative that there are contaminating particles contributing to the mutant maps. These particles may give the perception of greater flexibility as noise or other contaminating

images are less likely to conform to the expected shape of the protein. Thus, noisy images or non-real particles can decrease the resolution of maps and increase the perception of variability between subunits.

As a result of the limited resolution of the asymmetric maps and the variations in resolution between the different monomers and the caveats mentioned above, further analysis needs to be done to probe the degree of flexibility. Such analysis includes variations on 3D classification, including focused classification on flexible regions of the protein such as the B domains. This method has been recently utilized in order to more carefully sample the conformational space of flexible regions of proteins [294]. The flexibility of the protein should also be independently validated by other biophysical techniques such as fluorescence resonance energy transfer (FRET) to determine if probes attached to the B domains of mutant proteins come in closer contact than those attached to wild-type proteins. Since such techniques are performed in solution, they can more readily determine the inherent flexibility over time in dynamic samples as compared to techniques like cryo-EM which classifying snapshots frozen in time [295].

In summary, the determination of near-atomic structures of PKM2 in the wild-type and H391Y mutant forms is not only a technical achievement because of the protein's small size and flexibility, but also offers some preliminary biological insights. While the resolution of the maps are not high enough to conclusively determine molecular interactions, and thereby mechanisms, they offer the ability to generate and compare atomic models of the two forms. Differences between the models can be used to propose a stacking interaction that occurs at the domain interface and close to the subunit interface, which may explain some aspects of the differential activity of the H391Y mutant with respect to the wild-type.

Part IV

Concluding Remarks

Chapter 11

Summary of Imaging Methods Utilized in This Study

The studies in this thesis utilized a wide range of TEM-based methods in order to characterize key components of several diverse biological systems. The imaging techniques utilized vary in terms of both their practical resolution limits and the types of samples studied. Immuno-gold labeling of cryo-sections yields the lowest resolution (∼40Å) but compensates for this by providing the most cellular context. Thus it served as an appropriate choice for the PpfA studies since they were focused on cellular localization of the protein/complex rather than resolving high-resolution features of its atomic structure. In this case, the protein of interest, PpfA, was visualized in the context of intact cells with specific labeling directed towards the HA epitope with which the protein was tagged. This technique differed from the other imaging methods that were utilized in that the sample was stained and dried and thus was not imaged in its native state.

The cryo-electron tomography (cryo-ET) studies of Tsr provided an intermediate resolution range (∼20-25Å) as well as the possibility of visualizing the protein either in the cellular context or in purified membrane assemblies. Since Tsr becomes unstable when removed from its native membrane environment, cryo-ET was an optimal choice as it collects sufficient signal for subsequent macromolecular structure determination without requiring full solubilization of the target protein. It also al-

lowed for the preservation of near-native state conditions by using frozen, hydrated samples. Unlike PpfA in the samples used for immuno-labeling, Tsr was untagged and was not specifically labeled.

The highest resolution studies (∼3.5-5Å) focused on PKM2 and utilized single particle cryo-electron microscopy (cryo-EM). Here, the protein was completely removed from the cellular context in order to prevent the scattering of electrons by other cellular components. As a result, the signal in the collected images comes solely from the particles and the thin layer of vitrified ice in which they are embedded, thereby preserving the greatest amount of high resolution information about the particles. In both cases where macromolecular structure determination was required (Tsr and PKM2), subregions of the images or volumes were aligned, classified, and averaged in order to strengthen the signal to noise ratio.

The studies used in all of the experiments demonstrate a trade-off between the achievable resolution of the protein/complex and the level of information about its cellular context. Studies that are geared towards spatial localization of specific cellular contents tend to require more context and less resolution while macromolecular structure determination attempts to maximize resolution by minimizing signal from any source that is not the protein of interest. In all cases, the goal is to gain mechanistic insights based on the structures of macromolecules or cellular components.

Chapter 12

Appendices

A Media and Reagents

Lysogeny broth (LB)

Ingredient	Quantity
Yeast extract	5 g
Tryptone	10 g
NaCl	5 g

Make up to 1 liter with MilliQ water. Adjusted to pH 7.0. Autoclave.

Succinate media

Ingredient	Quantity
1M Phosphate buffer pH 7.0	20 mL
Concentrated base (see below)	20 mL
Growth factors (see below)	2 mL
$(NH_4)_2$ SO_4	0.5 g
NaCl	0.5 g
Sodium succinate	2 g
Casamino acids	1 g

Make up to 1 liter with MilliQ water, pH adjust to 7.2 with KOH, autoclave.

1M Phosphate buffer pH 7.0

Ingredient	Quantity
K_2HPO_4 1M	174.18 g
KH_2PO_4 1M	136.09 g

Make up to 1 liter with MilliQ water. Adjust to ph 7.0 with KOH, autoclave and store at 4°C.

Concentrated base

Ingredient	Quantity
Nitrilotiracetic acid (diNa salt)	5.94 g
Metals 44 solution (see below)	25 mL
$MgSO_4$ $7H_2O$	14.5 g
$CaCl_2$ $2H_2O$	1.67 g
$FeSO_4$ $7H_2O$	50 mg
Ammonium molybdate $4H_2O$	4.6 mg or 0.2 ml of 23 mg/mL

Dissolve diNa first and pH with KOH to pH 5.0, it won't dissolve until very near pH 5. This allows $FeSO_4$ to dissolve. Don't worry if you over pH it, adding the other ingredients with bring the pH back down. Add all the other ingredients. pH to 6.8 with KOH. Autoclave and store a 4°C.

Metals 44 solution

Ingredient	Quantity per litre
Tetrasodium versenate (EDTA)	2.5 g
$ZnSO_4$ $7H_2O$	11 g
$FeSO_4$ $7H_2O$	5 g
3M sulphuric acid	1.5 mL
$MnSO_4$ $4H_2O$	3 g
$CuSO_4$ $5H_2O$	0.39 g
H_3BO_3	0.12 g
$CoCl_2$ $6H_2O$	0.2 g

Add all of the dry ingredients and dissolve in 900 milliliters of MilliQ water. Add the sulphuric acid. Make up to 1 liter with MilliQ water. Filter sterilize and store at 4°C.

Growth factor

Ingredient	Quantity
Biotin	20 mg
$NaHCO_3$	500 mg
Niacin	1 g
Thiamine HCl	500 mg

Make up to 1 liter with MilliQ water. Autoclave and store at 4°C.

10x M9 Minimal Media

Ingredients	Quantity
$NaHPO_4$ $2H_2O$	75.2 g
KH_2PO_4	30 g
NaCl	5 g
NH_4Cl	10 g

Make up to 1 liter with MilliQ water.

SOC media

Ingredient	Concentration
Glucose	20 mM
MgCl	10 mM
$MgSO_4$	10 mM
NaCl	10 mM
Tryptone	20 g/l
Yeast extract	5 g/l

5x DNA loading dye

Ingredient	Quantity (W/v)
Glycerol	70%
Bromophenol blue	0.25%
Xylene cyanol	0.25&
MilliQ water	29.5&

10 x TBE

Ingredient	Quantity
Tris base	108 g
Boric acid	55 g
EDTA pH 8.0	7.4 g

Make up to 1 liter with MilliQ water. pH adjusted to pH 8.3.

TFB I

Ingredient	Concentration
Potassiun acetate	30 mM
$RbCl_2$	100 mM
$CaCl_2$ $2H_2O$	10mM
$MnCl_2$ $4H_2O$	50 mM
Glycerol	15% (v/v)

Adjust pH to 5.8 with 0.2 M acetic acid. Filter sterilize.

TFB II

Ingredient	Concentration
PIPES	10 mM
$CaCl_2$ $2H_2O$	75mM
$RbCl_2$	10 mM
Glycerol	15 % (v/v)

Adjust pH to 6.8 with HCl. Filter sterilize.

TSS

Ingredient	Quantity
LB	850mL
PEG-3350	100 g in 50 mL MilliQ water
DMSO	50 mL
1M $MgCl_2$	20 mL

Autoclave LB and PEG-3350 solutions. Allow to cool, then combine, filter sterilize, and store at 4°C.

5x KCM

Ingredient	Quantity
KCl	37.78 g
$CaCl_2$	32.86 g
$MgCl_2$	50.83 g

Make up to 1 liter with MilliQ water. Filter sterilize or autoclave.

Western Blot Transfer buffer

Ingredient	Quantity
20x NuPAGE Transfer Buffer (Invitrogen)	50 mL
Methanol	100 mL
NuPAGE Antioxidant (Invitrogen)	1 mL
MilliQ water	849 mL

2x Reducing Sample Buffer for SDS-PAGE

Ingredient	Quantity
2x Sample Reducing Buffer (Invitrogen)	700 μL
10x Reducing Agent (Invitrogen)	280 μL
MilliQ water	420 μL

Buffer A

Ingredient	Concentration
NaH_2PO_4	50 mM
NaCl	300 mM
Imidazole	20 mM

Adjust to pH 8.0. Filter sterilize.

Buffer B

Ingredient	Concentration
NaH_2PO_4	50 mM
NaCl	300 mM
Imidazole	250 mM

Adjust to pH 8.0. Filter sterilize.

B Primers

Restriction sites and mutation sites are indicated in bold.

Name	Sequence (5' - 3')	Feature
mv001	CGC**GGATCC**AGCGCATGGAGCCATCCGCAGTTTG	Forward strep tag (*bamhI* site)
mv002	CCC**AAGCTT**TTATTTTTCAAACTGCGGATGGCTCC	Reverse strep tag (*hindIII* site)
mv005	CATG**CCATGG**GAATGTTAAAACGTATCAAAATTGTG	Forward Tsr (*ncoI* site)
mv007	CGC**GGATCC**GCTGCCGCGCGGCACCAGAAATGTTTCCCAGTT CTCC	Reverse Tsr with cleavage site (*bamhI* site)
mv008	GGGAATTC**CATATG**CAGGGAGAGCTGATGCG	Forward mid-Tsr (*ndeI* site)
mv009	CGGGCATTTTCATGGCGATTATTTTTCAAACTGCGGATGGC	Reverse overlapping strep tag and YjiZ
mv010	GCCATCCGCAGTTTGAAAAATAATCGCCATGAAAATGCCCG	Forward overlapping strep tag and YjiZ
mv011	CCC**AAGCTT**ATCTCAATGCAGGTAGCGTC	Reverse YjiZ (*hindIII* site)
mv012	CCGTCAGCGATAACAATGC	Mid-Tsr sequencing
mv013	ATGCA**CCATGG**CCTACCCCTACGACGTGCCGGACTACGCCTCG GCCTATCCCTAC	Forward double HA N-term tag (*ncoI* site)
mv021	GACTTTCCCGACCCGGCGTAGTCCGGCACGTCGTAGGGATAGG CCGAGGCGTAG	Reverse overlapping double HA-APXII N-term tag
mv022	GACTACGCCGGGTCGGGAAAGTCTTACCCAACTGTGAGTGC	Forward overlapping double HA-APXII N-term tag
mv023	CAAGT**GGATCC**GGCCGAGGCATCAGCAAACCCAAGCTCGGAC AG	Reverse APXII N-term tag (*bamhI* site)
mv024	CAAGT**GGATCC**GCCTCGGGAAAGTCTTACCCAACTGTGAGTGC	Forward APXII C-term tag (*bamhI* site)
mv025	GTCGTAGGGGTACGACCCGGCATCAGCAAACCCAAGCTCGGAC AG	Reverse overlapping APXII-double HA C-term tag
mv026	GCCGGGTCGTACCCCTACGACGTGCCGGACTACGCCTCCGCCTA TCCCTAC	Forward overlapping APXII-double HA C-term tag
mv027	CAAGT**GGATCC**GCCTTGCGGGCGATGATCTGC	Forward PpfA for N-term tag (*bamhI* site)
mv028	CTAGT**AAGCTT**ATGTCCTGGTTCCCTTTCGTC	Reverse PpfA for N-term tag (*hindIII* site)
mv029	CCCAAGCTCGGA**CAG**CTTTTGGTG	Reverse APXII to remove *hindIII* site
mv048	CTGCCATCTAC**T**ACTTGCAATTATTTG	Forward PKM2 for introducing H391Y mutation
mv049	CAAATAATTGCAAGT**A**GTAGATGGCAG	Reverse PKM2 for introducing H391Y mutation

Bibliography

[1] B. Alberts, A. Johnson, Lewis. J., D. Morgan, M. Raff, K. Roberts, P. Water, J. Wilson, and T. Hunt. *Molecular Biology of the Cell*. Garland Science, New York, 6th edition, Nov 2014.

[2] Protein function. *Scitable by Nature Education website*, Accessed: 10 May 2016.

[3] K. A. Dill. Dominant forces in protein folding. *Biochemistry*, 29(31):7133–55, Aug 1990.

[4] P. S. Kim and R. L. Baldwin. Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.*, 59:631–60, Jul 1990.

[5] W. Kauzmann. Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.*, 14:1–63, 1959.

[6] R. S. Spolar, J. H. Ha, and M. T. Record, Jr. Hydrophobic effect in protein folding and other noncovalent processes involving proteins. *Proc. Natl. Acad. Sci. U. S. A.*, 86(21):8382–5, Nov 1989.

[7] S. Raina and D. Missiakas. Making and breaking disulfide bonds. *Annu. Rev. Microbiol.*, 51:179–202, Oct 1997.

[8] L. G. Presta and G. D. Rose. Helix signals in proteins. *Science*, 240(4859):1632–41, Jun 1988.

[9] G. E. Tusnády and I. Simon. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J. Mol. Biol.*, 283(2):489–506, Oct 1998.

[10] X. Feng and P. Barth. A topological and conformational stability alphabet for multipass membrane proteins. *Nat. Chem. Biol.*, 12(3):167–73, Mar 2016.

[11] C. Cohen and D. A. Parry. Alpha-helical coiled coils and bundles: how to design an alpha-helical protein. *Proteins*, 7(1):1–15, 1990.

[12] K. K. Kim, H. Yokota, and S. H. Kim. Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature*, 400(6746):787–92, Aug 1999.

[13] C. Chothia. Conformation of twisted beta-pleated sheets in proteins. *J. Mol. Biol.*, 75(2):295–302, Apr 1973.

[14] J. S. Richardson. The anatomy and taxonomy of protein structure. *Adv. Protein Chem.*, 34:167–339, 1981.

[15] M. H. Zehfus and G. D. Rose. Compact units in proteins. *Biochemistry*, 25(19):5759–65, Sep 1986.

[16] J. E. Rothman. Polypeptide chain binding proteins: catalysts of protein folding and related processes in cells. *Cell*, 59(4):591–601, Nov 1989.

[17] N. Tokuriki and D. S. Tawfik. Protein dynamism and evolvability. *Science*, 324(5924):203–7, Apr 2009.

[18] I. S. Moreira, P. A. Fernandes, and M. J. Ramos. Hot spots–a review of the protein-protein interface determinant amino-acid residues. *Proteins*, 68(4):803–12, Sep 2007.

[19] S. Soga, H. Shirai, M. Kobori, and N. Hirayama. Use of amino acid composition to predict ligand-binding sites. *J. Chem. Inf. Model.*, 47(2):400–6, Jan 2007.

[20] C. Tanford. Chemical potential of bound ligand, an important parameter for free energy transduction. *Proc. Natl. Acad. Sci. U. S. A.*, 78(1):270–3, Jan 1981.

[21] J. A. Wells. Additivity of mutational effects in proteins. *Biochemistry*, 29(37):8509–17, Sep 1990.

[22] S. J. Benkovic and S. Hammes-Schiffer. A perspective on enzyme catalysis. *Science*, 301(5637):1196–202, Aug 2003.

[23] S. Grisolia. The catalytic environment and its biological implications. *Physiol. Rev.*, 44:657–712, Oct 1964.

[24] E. R. Stadtman. Allosteric regulation of enzyme activity. *Adv. Enzymol. Relat. Areas Mol. Biol.*, 28:41–154, 1966.

[25] E. Whitehead. The regulation of enzyme activity and allosteric transition. *Prog. Biophys. Mol. Biol.*, 21:321–97, 1970.

[26] M. D. Manson. Bacterial chemoreceptors as membrane-spanning allosteric enzymes. *Sensory Mechanisms in Bacteria*, pages 107–149, 2010.

[27] D. L. Garland. Kinetics and mechanism of colchicine binding to tubulin: evidence for ligand-induced conformational change. *Biochemistry*, 17(20):4266–72, Oct 1978.

[28] S. Highsmith and D. Eden. Ligand-induced myosin subfragment 1 global conformational change. *Biochemistry*, 29(17):4087–93, May 1990.

[29] G. Weber. Energetics of ligand binding to proteins. *Adv. Protein Chem.*, 29:1–83, 1975.

[30] P. De Meyts and J. Roth. Cooperativity in ligand binding: a new graphic analysis. *Biochem. Biophys. Res. Commun.*, 66(4):1118–26, Oct 1975.

[31] P. Cohen. The regulation of protein function by multisite phosphorylation–a 25 year update. *Trends Biochem. Sci.*, 25(12):596–601, Dec 2000.

[32] T. Hunter. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell*, 80(2):225–36, Jan 1995.

[33] J. B. Stock, A. J. Ninfa, and A. M. Stock. Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.*, 53(4):450–90, Dec 1989.

[34] D. T. Denhardt. Signal-transducing protein phosphorylation cascades mediated by ras/rho proteins in the mammalian cell: the potential for multiplex signalling. *Biochem. J.*, 318 (Pt 3):729–47, Sep 1996.

[35] R. S. Fischer, Y. Wu, P. Kanchanawong, H. Shroff, and C. M. Waterman. Microscopy in 3D: a biologist's toolbox. *Trends Cell Biol.*, 21(12):682–91, Dec 2011.

[36] D.S. Lidke and K.A. Lidke. Advances in high-resolution imaging - techniques for three-dimensional imaging of cellular structures. *J. Cell Sci.*, 125(11):2571–80, Jun 2012.

[37] B. Glancy, L. M. Hartnell, D. Malide, Z. X. Yu, C. A. Combs, P. S. Connelly, S. Subramaniam, and R. S. Balaban. Mitochondrial reticulum for cellular energy distribution in muscle. *Nature*, 523(7562):617–20, Jul 2015.

[38] T. Do, G. Murphy, L. A. Earl, G. Q. Del Prete, G. Grandinetti, G. Li, J. D. Estes, P. Rao, C. M. Trubey, J. Thomas, J. Spector, D. Bliss, A. Nath, J. D. Lifson, and S. Subramaniam. Three-dimensional imaging of HIV-1 virological synapses reveals membrane architectures involved in virus transmission. *J. Virol.*, 88(18):10327–39, Sep 2014.

[39] J. A. Heymann, D. Shi, S. Kim, D. Bliss, J. L. S. Milne, and S. Subramaniam. 3D imaging of mammalian cells with ion-abrasion scanning electron microscopy. *J. Struct. Biol.*, 166(1):1–7, Apr 2009.

[40] K. Narayan, C. M. Danielson, K. Lagarec, B. C. Lowekamp, P. Coffman, A. Laquerre, M. W. Phaneuf, T. J. Hope, and S. Subramaniam. Multi-resolution correlative focused ion beam scanning electron microscopy: applications to cell biology. *J. Struct. Biol.*, 185(3):278–84, Mar 2014.

[41] E. Nogales and S. H. W. Scheres. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Mol. Cell*, 58(4):677–89, May 2015.

[42] A. Bartesaghi, A. Merk, D. Banerjee, S.and Matthies, X. Wu, J. L. S. Milne, and S. Subramaniam. 2.2 Å resolution cryo-EM structure of β-galactosidase in complex with a cell-permeant inhibitor. *Science*, 348(6239):1147–51, Jun 2015.

[43] T. A. M. Bharat, L. R. Castillo Menendez, W. J. H. Hagen, V. Lux, S. Igonet, M. Schorb, F. K. M. Schur, H. G. Kräusslich, and J. A. G. Briggs. Cryo-electron microscopy of tubular arrays of HIV-1 Gag resolves structures essential for immature virus assembly. *Proc. Natl. Acad. Sci. U.S.A.*, 111(22):8233–8, Jun 2014.

[44] S. Asano, Y. Fukuda, F. Beck, A. Aufderheide, F. Förster, R. Danev, and W. Baumeister. Proteasomes. A molecular census of 26S proteasomes in intact neurons. *Science*, 347(6220):439–42, Jan 2015.

[45] S. Pfeffer, L. Burbaum, P. Unverdorben, M. Pech, Y. Chen, R. Zimmermann, R. Beckmann, and F. Förster. Structure of the native Sec61 protein-conducting channel. *Nat. Commun.*, 6:8403, Sep 2015.

[46] A. Cavalli, X. Salvatella, C. M. Dobson, and M. Vendruscolo. Protein structure determination from NMR chemical shifts. *Proc. Natl. Acad. Sci. U.S.A.*, 104(23):9615–20, Jun 2007.

[47] P. R. L. Markwick, T. Malliavin, and M. Nilges. Structural biology by NMR: structure, dynamics, and interactions. *PLoS Comput. Biol.*, 4(9):e1000168, Sep 2008.

[48] D. B. Williams and C. B. Carter. *Transmission Electron Microscopy: A Textbook for Materials Science*. Springer Science+Business Media, LLC, New York, 2nd edition, 2009.

[49] J. Frank. *Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their Native State*. Oxford University Press, Inc., New York, 2nd edition, Mar 2006.

[50] Y. Cheng, N. Grigorieff, P. A. Penczek, and T. Walz. A primer to single-particle cryo-electron microscopy. *Cell*, 161(3):438–49, Apr 2015.

[51] R. Danev, B. Buijsee, M. Khoshouei, J. M. Plitzko, and W. Baumeister. Volta potential phase plate for in-focus phase contrast transmission electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.*, 111(44):15635–40, Sep 2014.

[52] R. Danev and W. Baumeister. Cryo-EM single particle analysis with the Volta phase plate. *eLife*, 5:e13046, Mar 2016.

[53] J. Dubochet, J. Lepault, R. Freeman, Berriman J. A., and J.-C. Homo. Electron microscopy of frozen water and aqueous solutions. *J. Microscop.*, 128(3):219–37, Dec 1982.

[54] T. Hrabe and F. Förster. Structure Determination by Single Particle Cryo-electron Tomography. *eLS*, Feb 2011.

[55] W. Kühlbrandt. Biochemistry. The resolution revolution. *Science*, 343(6178):1443–4, Mar 2014.

[56] A. F. Brilot, J. Z. Chen, A. Cheng, J. Pan, S. C. Harrison, C. S. Potter, B. Carragher, R. Henderson, and N. Grigorieff. Beam-induced motion of vitrified specimen on holey carbon film. *J. Struct. Biol.*, 177(3):630–7, Mar 2012.

[57] M. G. Campbell, A. Cheng, A. F. Brilot, A. Moeller, D. Lyumkis, D. Veesler, S. C. Pan, J.and Harrison, C. S. Potter, B. Carragher, and N. Grigorieff. Movies of ice-embedded particles enhance resolution in electron cryomicroscopy. *Structure*, 20(11):1823–8, Nov 2012.

[58] M. Liao, E. Cao, D. Julius, and Y. Cheng. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature*, 504(7478):107–12, Dec 2013.

[59] M. G. Campbell, D. Veesler, A. Cheng, C. S. Potter, and B. Carragher. 2.8 Å resolution reconstruction of the *Thermoplasma acidophilum* 20S proteasome using cryo-electron microscopy. *eLife*, 4:e06380, Mar 2015.

[60] K. J. Weissman. Uncovering the structures of modular polyketide synthases. *Nat. Prod. Rep.*, 32(3):436–53, Mar 2015.

[61] P. A. Penczek. Image restoration in cryo-electron microscopy. *Methods Enzymol.*, 482:35–72, 2010.

[62] J. Zhu, P. A. Penczek, R. Schröder, and J. Frank. Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: procedure and application to the 70s escherichia coli ribosome. *J. Struct. Biol.*, 118(3):197–219, Apr 1997.

[63] J. A. Mindell and N. Grigorieff. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.*, 142(3):334–47, Jun 2003.

[64] M. van Heel and M. Stöffler-Meilicke. Characteristic views of *E. coli* and *B. stearothermophilus* 30S ribosomal subunits in the electron microscope. *EMBO J.*, 4(9):2389–95, Sep 1985.

[65] L. Joyeux and P. A. Penczek. Efficiency of 2D alignment methods. *Ultramicroscopy*, 92(2):33–46, Jul 2002.

[66] P. A. Penczek, R. A. Grassucci, and J. Frank. The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy*, 53(3):251–70, Mar 1994.

[67] M. Radermacher, T. Wagenknecht, A. Verschoor, and J. Frank. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J. Microsc.*, 146(Pt 2):113–36, May 1987.

[68] A. E. Leschziner and E. Nogales. The orthogonal tilt reconstruction method: an approach to generating single-class volumes with no missing cone for ab initio reconstruction of asymmetric particles. *J. Struct. Biol.*, 153(3):284–99, Mar 2006.

[69] V. Lučić, F. Förster, and W. Baumeister. Structural studies by electron tomography: from cells to molecules. *Annu. Rev. Biochem.*, 74:833–65, Jul 2005.

[70] M. van Heel. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy*, 21(2):111–23, 1987.

[71] S. H. W. Scheres, H. Gao, M. Valle, G. T. Herman, P. P. B. Eggermont, J. Frank, and J. M. Carazo. Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat. Methods*, 4(1):27–9, Jan 2007.

[72] S. H. W. Scheres, R. Núñez Ramírez, Y. Gómez-Llorente, C. San Martín, P. P. B. Eggermont, and J. M. Carazo. Modeling experimental image formation for likelihood-based classification of electron microscopy data. *Structure*, 15(10):1167–77, Oct 2007.

[73] S. H. W. Scheres. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.*, 180(3):519–30, Dec 2012.

[74] S. Asano, B. D. Engel, and W. Baumeister. In Situ Cryo-Electron Tomography: A Post-Reductionist Approach to Structural Biology. *J. Mol. Biol.*, Oct 2015.

[75] A. Dubrovsky, S. Sorrentino, J. Harapin, K. T. Sapra, and O. Medalia. Developments in cryo-electron tomography for in situ structural analysis. *Arch. Biochem. Biophys.*, 581:78–85, Sep 2015.

[76] J. Frank, editor. *Electron Tomography: Methods for Three-Dimensional Visualization of Structures in the Cell*. Springer-Verlag, New York, 2nd edition, Dec 2006.

[77] R. A. Crowther, D. J. DeRosier, and A. Klug. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. *Proc. Roc. Soc. Lond.*, 317(1530):319–40, Jun 1920.

[78] R. Hovden, P. Ercius, Y. Jiang, D. Wang, Y. Yu, H. D. Abruña, V. Elser, and D. A. Muller. Breaking the Crowther limit: combining depth-sectioning and tilt tomography for high-resolution, wide-field 3D reconstructions. *Ultramicroscopy*, 140:26–31, May 2014.

[79] J. L. S. Milne and S. Subramaniam. Cryo-electron tomography of bacteria: progress, challenges and future prospects. *Nat. Rev. Microbiol.*, 7(9):666–75, Sep 2009.

[80] Q. Xiong, M. K. Morphew, C. L. Schwartz, A. H. Hoenger, and D. N. Mastronarde. CTF determination and correction for low dose tomographic tilt series. *J. Struct. Biol.*, 168(3):378–87, Dec 2009.

[81] O. Kuybeda, G. A. Frank, A. Bartesaghi, M. Borgnia, S. Subramaniam, and G. Sapiro. A collaborative framework for 3D alignment and classification of heterogeneous subvolumes in cryo-electron tomography. *J. Struct. Biol.*, 181(2):116–27, Feb 2013.

[82] A. Bartesaghi, P. Sprechmann, J. Liu, G. Randall, G. Sapiro, and S. Subramaniam. Classification and 3D averaging with missing wedge correction in biological electron tomography. *J. Struct. Biol.*, 162(3):436–50, Jun 2008.

[83] F. Förster, S. Pruggnaller, A. Seybert, and A. S. Frangakis. Classification of cryo-electron sub-tomograms using constrained correlation. *J. Struct. Biol.*, 161(3):276–86, Mar 2008.

[84] D. Anastasiou, Y. Yu, W. J. Israelsen, J. K. Jiang, M. B. Boxer, B. S. Hong, W. Tempel, S. Dimov, M. Shen, A. Jha, H. Yang, K. R. Mattaini, C. M. Metallo, B. P. Fiske, K. D. Courtney, S. Malstrom, T. M. Khan, C. Kung, A. P. Skoumbourdis, H. Veith, N. Southall, M. J. Walsh, K. R. Brimacombe, W. Leister, S. Y. Lunt, Z. R. Johnson, K. E. Yen, K. Kunii, S. M. Davidson, H. R. Christofk, C. P. Austin, J. Inglese, M. H. Harris, J. M. Asara, G. Stephanopoulos, F. G. Salituro, S. Jin, L. Dang, D. S. Auld, H. W. Park, L. C. Cantley, C. J. Thomas, and M. G. Vander Heiden. Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis. *Nat. Chem. Biol.*, 8(10):839–47, Oct 2012.

[85] K. F. Jarrell and M. J. McBride. The surprisingly diverse ways that prokaryotes move. *Nat. Rev. Microbiol.*, 6(6):466–76, Jun 2008.

[86] D. Nakane, K. Sato, H. Wada, M. J. McBride, and K. Nakayama. Helical flow of surface protein required for bacterial gliding motility. *Proc. Natl. Acad. Sci. U.S.A.*, 110(27):11145–50, Jul 2013.

[87] I. Bulyha, C. Schmidt, P. Lenz, V. Jakovljevic, A. Höne, B. Maier, M. Hoppert, and L. S-Andersen. Regulation of the type IV pili molecular machine by dynamic localization of two motor proteins. *Mol. Microbiol.*, 74(3):691–706, Nov 2009.

[88] T. Mignot, J. W. Shaevitz, P. L. Hartzell, and D. R. Zusman. Evidence that focal adhesion complexes power bacterial gliding motility. *Science*, 315(5813):853–6, Feb 2007.

[89] D. Kaiser. Bacterial swarming: a re-examination of cell-movement patterns. *Curr. Biol.*, 17(14):R561–70, Jul 2007.

[90] H. C. Berg and R. A. Anderson. Bacteria swim by rotating their flagellar filaments. *Nature*, 245(5425):380–2, Oct 1973.

[91] J. K. Anderson, T. G. Smith, and T. R. Hoover. Sense and sensibility: flagellum-mediated gene regulation. *Trends Microbiol.*, 18(1):30–7, Jan 2010.

[92] S. L. Porter, G. H. Wadhams, and J. P. Armitage. *Rhodobacter sphaeroides*: complexity in chemotactic signalling. *Trends Microbiol.*, 16(6):251–60, Jun 2008.

[93] A. J. Wolfe, D. S. Millikan, J. M. Campbell, and K.. L Visick. *Vibrio fischeri* σ54 controls motility, biofilm formation, luminescence, and colonization. *Appl. Environ. Microbiol.*, 70(4):2520–4, Apr 2004.

[94] N. W. Charon, E. P. Greenberg, M. B. Koopman, and R. J. Limberger. Spirochete chemotaxis, motility, and the structure of the spirochetal periplasmic flagella. *Res. Microbiol.*, 143(6):597–603, 1992.

[95] B. L. Taylor and D. E. Koshland, Jr. Reversal of flagellar rotation in monotrichous and peritrichous bacteria: generation of changes in direction. *J. Bacteriol.*, 119(2):640–2, Aug 1974.

[96] J. P. Armitage and R. M. Macnab. Unidirectional, intermittent rotation of the flagellum of *Rhodobacter sphaeroides*. *J. Bacteriol.*, 169(2):514–8, Feb 1987.

[97] T. Pilizota, M. T. Brown, M. C. Leake, R. W. Branch, R. M. Berry, and J. P. Armitage. A molecular brake, not a clutch, stops the *Rhodobacter sphaeroides* flagellar motor. *Proc. Natl. Acad. Sci. U.S.A.*, 106(28):11582–7, Jul 2009.

[98] H. C. Berg and D. A. Brown. Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature*, 239(5374):500–4, Oct 1972.

[99] C. T. Lefèvre and D. A. Bazylinski. Ecology, diversity, and evolution of magnetotactic bacteria. *Microbiol. Mol. Biol. Rev.*, 77(3):497–526, Sep 2013.

[100] S. Romagnoli, H. L. Packer, and J. P. Armitage. Tactic responses to oxygen in the phototrophic bacterium *Rhodobacter sphaeroides* WS8N. *J. Bacteriol.*, 184(20):5590–8, Oct 2002.

[101] M. J. Sackett, J. P. Armitage, E. E. Sherwood, and T. P. Pitta. Photoresponses of the purple nonsulfur bacteria *Rhodospirillum centenum* and *Rhodobacter sphaeroides*. *J. Bacteriol.*, 179(21):6764–8, Nov 1997.

[102] M. F. Gluch, D. Typke, and W. Baumeister. Motility and thermotactic responses of *Thermotoga maritima*. *J. Bacteriol.*, 177(19):5473–9, Oct 1995.

[103] J. P. Armitage. Bacterial motility and chemotaxis. *Sc.i Prog.*, 76(301-302 Pt 3-4):451–77, 1992.

[104] G. Alexandre. Coupling metabolism and chemotaxis-dependent behaviours by energy taxis receptors. *Microbiology*, 156(Pt 8):2283–93, Aug 2010.

[105] S. M. Kirov, M. Castrisios, and J. G. Shaw. Aeromonas flagella (polar and lateral) are enterocyte adhesins that contribute to biofilm formation on surfaces. *Infect. Immun.*, 72(4):1939–45, Apr 2004.

[106] S. Moorthy and P. I. Watnick. Genetic evidence that the *Vibrio cholerae* monolayer is a distinct stage in biofilm development. *Mol. Microbiol.*, 52(2):573–87, Apr 2004.

[107] S. Pandya, P. Iyer, V. Gaitonde, T. Parekh, and A. Desai. Chemotaxis of rhizobium sp.s2 towards cajanus cajan root exudate and its major components. *Curr. Microbiol.*, 38(4):205–9, Apr 1999.

[108] D. S. Millikan and E. G. Ruby. Flra, a σ54-dependent transcriptional activator in *Vibrio fischeri*, is required for motility and symbiotic light-organ colonization. *J. Bacteriol.*, 185(12):3547–57, Jun 2003.

[109] C. Josenhans and S. Suerbaum. The role of motility as a virulence factor in bacteria. *Int. J. Med. Microbiol.*, 291(8):605–14, Mar 2002.

[110] S. M. Williams, Y.-T. Chen, T. M. Andermann, J. E. Carter, D. J. McGee, and K. M. Ottemann. *Helicobacter pylori* chemotaxis modulates inflammation and bacterium-gastric epithelium interactions in infected mice. *Infect. Immun.*, 75(8):3747–57, Aug 2007.

[111] S. M. Butler and A. Camilli. Both chemotaxis and net motility greatly influence the infectivity of *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.*, 101(14):5018–23, Apr 2004.

[112] R. B. Bourret and R. E. Silversmith. Two-component signal transduction. *Curr. Opin. Microbiol.*, 13(2):113–5, Apr 2010.

[113] G. H. Wadhams and J. P. Armitage. Making sense of it all: bacterial chemotaxis. *Nat. Rev. Mol. Cell Biol.*, 5(12):1024–37, Dec 2004.

[114] J. P. Armitage. Bacterial tactic responses. *Adv. Microb. Physiol.*, 41:229–89, 1999.

[115] T. Boldog, S. Grimme, M. Li, S. G. Sligar, and G. L. Hazelbauer. Nanodiscs separate chemoreceptor oligomeric states and reveal their signaling properties. *Proc. Natl. Acad. Sci. U.S.A.*, 103(31):11509–14, Aug 2006.

[116] J. J. Falke and G. L. Hazelbauer. Transmembrane signaling in bacterial chemoreceptors. *Trends Biochem. Sci.*, 26(4):257–65, Apr 2001.

[117] V. Sourjik and J. P. Armitage. Spatial organization in bacterial chemotaxis. *EMBO J.*, 29(16):2724–33, Aug 2010.

[118] S. L. Mowbray and D. E. Koshland, Jr. Additive and independent responses in a single receptor: aspartate and maltose stimuli on the tar protein. *Cell*, 50(2):171–80, Jul 1987.

[119] S. A. Chervitz and J. J. Falke. Molecular mechanism of transmembrane signaling by the aspartate receptor: a model. *Proc. Natl. Acad. Sci. U.S.A.*, 93(6):2545–50, Mar 1996.

[120] G. L. Hazelbauer and W.-C. Lai. Bacterial chemoreceptors: providing enhanced features to two-component signaling. *Curr. Opin. Microbiol.*, 13(2):124–32, Apr 2010.

[121] M. Welch, K. Oosawa, S. Aizawa, and M. Eisenbach. Phosphorylation-dependent binding of a signal molecule to the flagellar switch of bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, 90(19):8787–91, Oct 1993.

[122] A S Toker and R M Macnab. Distinct regions of bacterial flagellar switch protein flim interact with flig, flin and chey. *J Mol Biol*, 273(3):623–34, Oct 1997.

[123] M. M. McEvoy, A. Bren, M. Eisenbach, and F. W. Dahlquist. Identification of the binding interfaces on CheY for two of its targets, the phosphatase CheZ and the flagellar switch protein fliM. *J. Mol. Biol.*, 289(5):1423–33, Jun 1999.

[124] J. F. Hess, K. Oosawa, N. Kaplan, and M. I. Simon. Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell*, 53(1):79–87, Apr 1988.

[125] G. S. Anand, P. N. Goudreau, and A. M. Stock. Activation of methylesterase CheB: evidence of a dual role for the regulatory domain. *Biochemistry*, 37(40):14038–47, Oct 1998.

[126] N. Barkai and S. Leibler. Robustness in simple biochemical networks. *Nature*, 387(6636):913–7, Jun 1997.

[127] U. Alon, M. G. Surette, N. Barkai, and S. Leibler. Robustness in bacterial chemotaxis. *Nature*, 397(6715):168–71, Jan 1999.

[128] G. L. Hazelbauer, J. J. Falke, and J. S. Parkinson. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem. Sci.*, 33(1):9–19, Jan 2008.

[129] R. Hamer, P.-Y. Chen, J. P. Armitage, G. Reinert, and C. M. Deane. Deciphering chemotaxis pathways using cross species comparisons. *BMC Syst. Biol.*, 4:3, 2010.

[130] G. H. Wadhams, A. V. Warren, A. C. Martin, and J. P. Armitage. Targeting of two signal transduction pathways to different regions of the bacterial cell. *Mol. Microbiol.*, 50(3):763–70, Nov 2003.

[131] G. H. Wadhams, A. C. Martin, A. V. Warren, and J. P. Armitage. Requirements for chemotaxis protein localization in *Rhodobacter sphaeroides*. *Mol. Microbiol.*, 58(3):895–902, Nov 2005.

[132] P. A. Hamblin, B. A. Maguire, R. N. Grishanin, and J. P. Armitage. Evidence for two chemosensory pathways in *Rhodobacter sphaeroides*. *Mol. Microbiol.*, 26(5):1083–96, Dec 1997.

[133] S. Poggio, C. Abreu-Goodger, S. Fabela, A. Osorio, G. Dreyfus, P. Vinuesa, and L. Camarena. A complete set of flagellar genes acquired by horizontal transfer coexists with the endogenous flagellar system in *Rhodobacter sphaeroides*. *J. Bacteriol.*, 189(8):3208–16, Apr 2007.

[134] A. M. del Campo, T. Ballado, J. de la Mora, S. Poggio, L. Camarena, and G. Dreyfus. Chemotactic control of the two flagellar systems of *Rhodobacter sphaeroides* is mediated by different sets of CheY and FliM proteins. *J. Bacteriol.*, 189(22):8397–401, Nov 2007.

[135] G. H. Wadhams, A. C. Martin, and J. P. Armitage. Identification and localization of a methyl-accepting chemotaxis protein in *Rhodobacter sphaeroides*. *Mol. Microbiol.*, 36(6):1222–33, Jun 2000.

[136] G. H. Wadhams, A. C. Martin, S. L. Porter, J. R. Maddock, J. C. Mantotta, H. M. King, and J. P. Armitage. TlpC, a novel chemotaxis protein in *Rhodobacter sphaeroides*, localizes to a discrete region in the cytoplasm. *Mol. Microbiol.*, 46(5):1211–21, Dec 2002.

[137] S. L. Porter, M. A. J. Roberts, C. S. Manning, and J. P. Armitage. A bifunctional kinase-phosphatase in bacterial chemotaxis. *Proc. Natl. Acad. Sci. U.S.A.*, 105(47):18531–6, Nov 2008.

[138] S. L. Porter, G. H. Wadhams, and J. P. Armitage. Signal processing in complex chemotaxis pathways. *Nat. Rev. Microbiol.*, 9(3):153–65, Mar 2011.

[139] K. A. Scott, S. L. Porter, E. A. L. Bagg, R. Hamer, J. L. Hill, D. A. Wilkinson, and J. P. Armitage. Specificity of localization and phosphotransfer in the CheA proteins of *Rhodobacter sphaeroides*. *Mol. Microbiol.*, 76(2):318–30, Apr 2010.

[140] S. L. Porter, G. H. Wadhams, A. C. Martin, E. D. Byles, D. E. Lancaster, and J. P. Armitage. The CheYs of *Rhodobacter sphaeroides*. *J. Biol. Chem.*, 281(43):32694–704, Oct 2006.

[141] C. A. Studdert and J. S. Parkinson. Insights into the organization and dynamics of bacterial chemoreceptor clusters through *in vivo* crosslinking studies. *Proc. Natl. Acad. Sci. U.S.A.*, 102(43):15623–8, Oct 2005.

[142] M. K. Herrera Seitz, V. Frank, D. A. Massazza, A. Vaknin, and C. A. Studdert. Bacterial chemoreceptors of different length classes signal independently. *Mol. Microbiol.*, 93(4):814–22, Aug 2014.

[143] S.-Y. Park, P. P. Borbat, G. Gonzalez-Bonet, J. Bhatnagar, A. M. Pollard, J. H. Freed, A. M. Bilwes, and B. R. Crane. Reconstruction of the chemotaxis receptor-kinase assembly. *Nat. Struct. Mol. Biol.*, 13(5):400–7, May 2006.

[144] J. Lefman, P. Zhang, T. Hirai, R. M. Weis, J. Juliani, D. Bliss, M. Kessel, E. Bos, P. J. Peters, and S. Subramaniam. Three-dimensional electron microscopic imaging of membrane invaginations in *Escherichia coli* overproducing the chemotaxis receptor Tsr. *J. Bacteriol.*, 186(15):5052–61, Aug 2004.

[145] C. M. Khursigara, X. Wu, P. Zhang, J. Lefman, and S. Subramaniam. Role of HAMP domains in chemotaxis signaling by bacterial chemoreceptors. *Proc. Natl. Acad. Sci. U.S.A.*, 105(43):16555–60, Oct 2008.

[146] H. Tajima, K. Imada, M. Sakuma, F. Hattori, T. Nara, N. Kamo, M. Homma, and I. Kawagishi. Ligand specificity determined by differentially arranged common ligand-binding residues in bacterial amino acid chemoreceptors Tsr and Tar. *J. Biol. Chem.*, 286(49):42200–10, Dec 2011.

[147] R. R. Draheim, A. F. Bormans, R.-Z. Lai, and M. D. Manson. Tuning a bacterial chemoreceptor with protein-membrane interactions. *Biochemistry*, 45(49):14655–64, Dec 2006.

[148] C. A. Adase, R. R. Draheim, and M. D. Manson. The residue composition of the aromatic anchor of the second transmembrane helix determines the signaling properties of the aspartate/maltose chemoreceptor Tar of *Escherichia coli*. *Biochemistry*, 51(9):1925–32, Mar 2012.

[149] G. A. Wright, R. L. Crowder, R. R. Draheim, and M. D. Manson. Mutational analysis of the transmembrane helix 2-HAMP domain connection in the *Escherichia coli* aspartate chemoreceptor Tar. *J. Bacteriol.*, 193(1):82–90, Jan 2011.

[150] S. Kitanovic, P. Ames, and J. S. Parkinson. A trigger residue for transmembrane signaling in the *Escherichia coli* serine chemoreceptor. *J. Bacteriol.*, 197(15):2568–79, Aug 2015.

[151] J. I. Yeh, H. P. Biemann, G. G. Privé, J. Pandit, D. E. Koshland, Jr, and S. H. Kim. High-resolution structures of the ligand binding domain of the wild-type bacterial aspartate receptor. *J. Mol. Biol.*, 262(2):186–201, Sep 1996.

[152] M. Hulko, F. Berndt, M. Gruber, J. U. Linder, V. Truffault, A. Schultz, J. Martin, J. E. Schultz, A. N. Lupas, and M. Coles. The HAMP domain structure implies helix rotation in transmembrane signaling. *Cell*, 126(5):929–40, Sep 2006.

[153] K. E. Swain and J. J. Falke. Structure of the conserved HAMP domain in an intact, membrane-bound chemoreceptor: a disulfide mapping study. *Biochemistry*, 46(48):13684–95, Dec 2007.

[154] J. S. Parkinson. Signaling mechanisms of HAMP domains in chemoreceptors and sensor kinases. *Annu. Rev. Microbiol.*, 64:101–22, 2010.

[155] M. V. Airola, K. J. Watts, A. M. Bilwes, and B. R. Crane. Structure of concatenated HAMP domains provides a mechanism for signal transduction. *Structure*, 18(4):436–48, Mar 2010.

[156] R. P. Alexander and I. B. Zhulin. Evolutionary genomics reveals conserved structural determinants of signaling and adaptation in microbial chemoreceptors. *Proc. Natl. Acad. Sci. U.S.A.*, 104(8):2885–90, Feb 2007.

[157] S. Djordjevic and A. M. Stock. Chemotaxis receptor recognition by protein methyltransferase CheR. *Nat. Struct. Biol.*, 5(6):446–50, Jun 1998.

[158] E. Perez and A. M. Stock. Characterization of the *Thermotoga maritima* chemotaxis methylation system that lacks pentapeptide-dependent methyltransferase CheR:MCP tethering. *Mol. Microbiol.*, 63(2):363–78, Jan 2007.

[159] M. Li and G. L. Hazelbauer. Cellular stoichiometry of the components of the chemotaxis signaling complex. *J. Bacteriol.*, 186(12):3687–94, Jun 2004.

[160] J. R. Maddock and L. Shapiro. Polar location of the chemoreceptor complex in the *Escherichia coli* cell. *Science*, 259(5102):1717–23, Mar 1993.

[161] V. Sourjik and H. C. Berg. Localization of components of the chemotaxis machinery of *Escherichia coli* using fluorescent protein fusions. *Mol. Microbiol.*, 37(4):740–51, Aug 2000.

[162] S. Thiem, D. Kentner, and V. Sourjik. Positioning of chemosensory clusters in *E. coli* and its relation to cell division. *EMBO J.*, 26(6):1615–23, Mar 2007.

[163] P. Zhang, C. M. Khursigara, L. M. Hartnell, and S. Subramaniam. Direct visualization of *Escherichia coli* chemotaxis receptor arrays using cryo-electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.*, 104(10):3777–81, Mar 2007.

[164] A. Briegel, D. R. Ortega, E. I. Tocheva, K. Wuichet, Z. Li, S. Chen, A. Müller, C. V. Iancu, G. E. Murphy, M. J. Dobro, I. B. Zhulin, and G. J. Jensen. Universal architecture of bacterial chemoreceptor arrays. *Proc. Natl. Acad. Sci. U.S.A.*, 106(40):17181–6, Oct 2009.

[165] A. Briegel, X. Li, A. M. Bilwes, K. T. Hughes, G. J. Jensen, and B. R. Crane. Bacterial chemoreceptor arrays are hexagonally packed trimers of receptor dimers networked by rings of kinase and coupling proteins. *Proc. Natl. Acad. Sci. U.S.A.*, 109(10):3766–71, Mar 2012.

[166] J. Liu, B. Hu, D. R. Morado, S. Jani, M. D. Manson, and W. Margolin. Molecular architecture of chemoreceptor arrays revealed by cryoelectron tomography of *Escherichia coli* minicells. *Proc. Natl. Acad. Sci. U.S.A.*, 109(23):E1481–8, Jun 2012.

[167] A. Briegel, M. S. Ladinsky, C. Oikonomou, C. W. Jones, M. J. Harris, D. J. Fowler, Y. Chang, L. K. Thompson, J. P. Armitage, and G. J. Jensen. Structure of bacterial cytoplasmic chemoreceptor arrays and implications for chemotactic signaling. *eLife*, 3:e02151, Mar 2014.

[168] V. Sourjik and H. C. Berg. Localization of components of the chemotaxis machinery of *Escherichia coli* using fluorescent protein fusions. *Mol. Microbiol.*, 37(4):740–51, Aug 2000.

[169] A. Briegel, M. L. Wong, H. L. Hodges, C. M. Oikonomou, K. N. Piasta, M. J. Harris, D. J. Fowler, L. K. Thompson, J. J. Falke, L. L. Kiessling, and

G. J. Jensen. New insights into bacterial chemoreceptor array structure and assembly from electron cryotomography. *Biochemistry*, 53(10):1575–85, Mar 2014.

[170] D. Kentner, S. Thiem, M. Hildenbeutel, and V. Sourjik. Determinants of chemoreceptor cluster formation in *Escherichia coli*. *Mol. Microbiol.*, 61(2):407–17, Jul 2006.

[171] S. Thiem and V. Sourjik. Stochastic assembly of chemoreceptor clusters in *Escherichia coli*. *Mol. Microbiol.*, 68(5):1228–36, Jun 2008.

[172] D. Greenfield, A. L. McEvoy, H. Shroff, G. E. Crooks, N. S. Wingreen, E. Betzig, and J. Liphardt. Self-organization of the *Escherichia coli* chemotaxis network imaged with super-resolution light microscopy. *PLoS Biol.*, 7(6):e1000137, Jun 2009.

[173] S.-W. Chiu, M. A. J. Roberts, M. C. Leake, and J. P. Armitage. Positioning of chemosensory proteins and ftsz through the *Rhodobacter sphaeroides* cell cycle. *Mol. Microbiol.*, 90(2):322–37, Oct 2013.

[174] S. R. Thompson, G. H. Wadhams, and J. P. Armitage. The positioning of cytoplasmic protein clusters in bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, 103(21):8209–14, May 2006.

[175] T. A. Leonard, P. J. Butler, and J. Löwe. Bacterial chromosome segregation: structure and DNA binding of the Soj dimer–a conserved biological switch. *EMBO J.*, 24(2):270–82, Jan 2005.

[176] A. G. Vecchiarelli, Y.-W. Han, X. Tan, M. Mizuuchi, R. Ghirlando, C. Biertümpfel, B. E. Funnell, and K. Mizuuchi. ATP control of dynamic P1 ParA-DNA interactions: a key role for the nucleoid in plasmid partition. *Mol. Microbiol.*, 78(1):78–91, Oct 2010.

[177] K. Gerdes, M. Howard, and F. Szardenings. Pushing and pulling in prokaryotic DNA segregation. *Cell*, 141(6):927–42, Jun 2010.

[178] M. A. J. Roberts, G. H. Wadhams, K. A. Hadfield, S. Tickner, and J. P. Armitage. ParA-like protein uses nonspecific chromosomal DNA binding to partition protein complexes. *Proc. Natl. Acad. Sci. U.S.A.*, 109(17):6698–703, Apr 2012.

[179] D. Barillà, E. Carmelo, and F. Hayes. The tail of the ParG DNA segregation protein remodels ParF polymers and enhances ATP hydrolysis via an arginine finger-like motif. *Proc. Natl. Acad. Sci. U.S.A.*, 104(6):1811–6, Feb 2007.

[180] S. Ringgaard, J. van Zon, M. Howard, and K. Gerdes. Movement and equipositioning of plasmids by ParA filament disassembly. *Proc. Natl. Acad. Sci. U.S.A.*, 106(46):19369–74, Nov 2009.

[181] D. Mattheis, O. Dalmas, M. J. Borgnia, P. K. Dominik, A. Merk, P. Rao, B. G. Reddy, S. Islam, A. Bartesaghi, E. Perozo, and S. Subramaniam. Cryo-EM structures of the magnesium channel CorA reveal symmetry break upon gating. *Cell*, 164(4):747–56, Feb 2016.

[182] M. P. Conley, A. J. Wolfe, D. F. Blair, and H. C. Berg. Both CheA and CheW are required for reconstitution of chemotactic signaling in *Escherichia coli*. *J. Bacteriol.*, 171(9):5190–3, Sep 1989.

[183] C. L. Woldringh. Morphological analysis of nuclear separation and cell division during the life cycle of *Escherichia coli*. *J. Bacteriol.*, 125(1):248–57, Jan 1976.

[184] R. E. Sockett, J. C. A Foster, and J. P. Armitage. Molecular biology of the *Rhodobacter sphaeroides* flagellum. *FEMS Symposium*, 53:473–9, 1990.

[185] R. M. Weis, T. Hirai, A. Chalah, M. Kessel, P. J. Peters, and S. Subramaniam. Electron microscopic analysis of membrane assemblies formed by the bacterial chemotaxis receptor Tsr. *J. Bacteriol.*, 185(12):3636–43, Jun 2003.

[186] A. C. Ind, S. L. Porter, M. T. Brown, E. D. Byles, J. A. de Beyer, S. A. Godfrey, and J. P. Armitage. Inducible-expression plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*. *Appl. Environ. Microbiol.*, 75(20):6613–5, Oct 2009.

[187] O. Sliusarenko, J. Heinritz, T. Emonet, and C. Jacobs-Wagner. High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Mol. Microbiol.*, 80(3):612–27, May 2011.

[188] J. R. Kremer, D. N. Mastronarde, and J. R. McIntosh. Computer visualization of three-dimensional image data using imod. *J. Struct. Biol.*, 116(1):71–6, 1996.

[189] F. Amat, F. Moussavi, L. R. Comolli, G. Elidan, K. H. Downing, and M. Horowitz. Markov random field based automatic image alignment for electron tomography. *J. Struct. Biol.*, 161(3):260–75, Mar 2008.

[190] S. J. Ludtke, P. R. Baldwin, and W. Chiu. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.*, 128(1):82–97, Dec 1999.

[191] W. Liou, H. J. Geuze, and J. W. Slot. Improving structural integrity of cryosections for immunogold labeling. *Histochem. Cell Biol.*, 106(1):41–58, Jul 1996.

[192] G. Griffiths, K. Simons, G. Warren, and K. T. Tokuyasu. Immunoelectron microscopy using thin, frozen sections: application to studies of the intracellular transport of semliki forest virus spike glycoproteins. *Methods Enzymol.*, 96:466–85, 1983.

[193] Protein expression and purification core facility: protein expression: *E. coli*. *European Molecular Biology Laboratory website*, Accessed: 23 Nov 2015.

[194] S. P. Howard, D. Cavard, and C. Lazdunski. Phospholipase-A-independent damage caused by the colicin A lysis protein during its assembly into the inner and outer membranes of *Escherichia coli*. *J. Gen. Microbiol.*, 137(1):81–9, Jan 1991.

[195] M. Adrian, J. Dubochet, J. Lepault, and A. W. McDowall. Cryo-electron microscopy of viruses. *Nature*, 308(5954):32–6, Mar 1984.

[196] M. le Maire, P. Champeil, and J. V. Moller. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta*, 1508(1-2):86–111, Nov 2000.

[197] T. G. M. Schmidt and A. Skerra. The *Strep*-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protoc.*, 2(6):1528–35, Jun 2007.

[198] D. S. Waugh. Making the most of affinity tags. *Trends Biotechnol.*, 23(6):316–20, Jun 2005.

[199] A. N. Barnakov, L. A. Barnakova, and G. L. Hazelbauer. Efficient adaptational demethylation of chemoreceptors requires the same enzyme-docking site as efficient methylation. *Proc. Natl. Acad. Sci. U.S.A.*, 96(19):10667–72, Sep 1999.

[200] C. Tribet, R. Audebert, and J. L. Popot. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc. Natl. Acad. Sci. U.S.A.*, 93(26):15047–50, Dec 1996.

[201] J. R. Meyerson, P. Rao, J. Kumar, S. Chittori, S. Banerjee, J. Pierson, M. L. Mayer, and S. Subramaniam. Self-assembled monolayers improve protein distribution on holey carbon cryo-EM supports. *Sci. Rep.*, 4:7084, 2014.

[202] H. Stark. GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods Enzymol.*, 481:109–26, 2010.

[203] M. Li, C. M. Khursigara, S. Subramaniam, and G. L. Hazelbauer. Chemotaxis kinase CheA is activated by three neighbouring chemoreceptor dimers as effectively as by receptor clusters. *Mol. Microbiol.*, 79(3):677–85, Feb 2011.

[204] N. L. Bartelli and G. L. Hazelbauer. Differential backbone dynamics of companion helices in the extended helical coiled-coil domain of a bacterial chemoreceptor. *Protein Sci.*, 24(11):1764–76, Nov 2015.

[205] P. Ames and J. S. Parkinson. Constitutively signaling fragments of Tsr, the *Escherichia coli* serine chemoreceptor. *J. Bacteriol.*, 176(20):6340–8, Oct 1994.

[206] A. Briegel, P. Ames, J. C. Gumbart, C. M. Oikonomou, J. S. Parkinson, and G. J. Jensen. The mobility of two kinase domains in the *Escherichia coli* chemoreceptor array varies with signalling state. *Mol. Microbiol.*, 89(5):831–41, Sep 2013.

[207] Q. Wang, C. P. Mercogliano, and J. Löwe. A ferritin-based label for cellular electron cryotomography. *Structure*, 19(2):147–54, Feb 2011.

[208] G. Furness. The transfer of motility and tyrosine requirement to *Escherichia coli* strain B by recombination with *E. coli* strain K12. *J. Gen. Microbiol.*, 18(3):782–6, Jun 1958.

[209] E. M. Witkin. Genetics of resistance to radiation in *Escherichia coli*. *Genetics*, 32(3):221–48, May 1947.

[210] J. S. Parkinson. Complementation analysis and deletion mapping of *Escherichia coli* mutants defective in chemotaxis. *J. Bacteriol.*, 135(1):45–53, Jul 1978.

[211] M. C. Konopka, I. A. Shkel, S. Cayley, M. T. Record, and J. C. Weisshaar. Crowding and confinement effects on protein diffusion *in vivo*. *J. Bacteriol.*, 188(17):6115–23, Sep 2006.

[212] K. A. Sochacki, I. A. Shkel, M. T. Record, and J. C. Weisshaar. Protein diffusion in the periplasm of *E. coli* under osmotic stress. *Biophys. J.*, 100(1):22–31, Jan 2011.

[213] P. J. Walser, N. Ariotti, M. Howes, C. Ferguson, R. Webb, D. Schwudke, N. Leneva, K.-J. Cho, L. Cooper, J. Rae, M. Floetenmeyer, V. M. J. Oorschot, U. Skoglund, K. Simons, J. F. Hancock, and R. G. Parton. Constitutive formation of caveolae in a bacterium. *Cell*, 150(4):752–63, Aug 2012.

[214] J. Shin, Y.-H. Jung, D.-H. Cho, M. Park, K. E. Lee, Y. Yang, C. Jeong, B. H. Sung, J.-H. Sohn, J.-B. Park, and D.-H. Kweon. Display of membrane proteins on the heterologous caveolae carved by caveolin-1 in the *Escherichia coli* cytoplasm. *Enzyme Microb. Technol.*, 79-80:55–62, Nov 2015.

[215] X. Fu, B. A. Himes, D. Ke, W. J. Rice, J. Ning, and P. Zhang. Controlled bacterial lysis for electron tomography of native cell membranes. *Structure*, 22(12):1875–82, Dec 2014.

[216] S. S. Lam, J. D. Martell, K. J. Kamer, T. J. Deerinck, M. H. Ellisman, V. K. Mootha, and A. Y. Ting. Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods*, 12(1):51–4, Jan 2015.

[217] K. L. McDonald. A review of high-pressure freezing preparation techniques for correlative light and electron microscopy of the same cells and tissues. *J. Microsc.*, 235(3):273–81, Sep 2009.

[218] J. Harapin, M. Börmel, K. T. Sapra, D. Brunner, A. Kaech, and O. Medalia. Structural analysis of multicellular organisms with cryo-electron tomography. *Nat. Methods*, 12(7):634–6, Jul 2015.

[219] C. D. J. Parmenter, M. W. Fay, C. Hartfield, and H. M. Eltaher. Making the practically impossible "merely difficult"-cryogenic fib lift-out for "damage free" soft matter imaging. *Microsc. Res. Tech.*, 79(4):298–303, Apr 2016.

[220] D. Hanahan and R. A. Weinberg. Hallmarks of cancer: the next generation. *Cell*, 144(5):646–74, Mar 2011.

[221] B. DeLaBarre, J. Hurov, G. Cianchetta, S. Murray, and L. Dang. Action at a distance: allostery and the development of drugs to target cancer cell metabolism. *Chem. Biol.*, 21(9):1143–61, Sep 2014.

[222] R. A. Cairns, I. S. Harris, and T. W. Mak. Regulation of cancer cell metabolism. *Nat. Rev. Cancer*, 11(2):85–95, Feb 2011.

[223] T. Soga. Cancer metabolism: key players in metabolic reprogramming. *Cancer Sci.*, 104(3):275–81, Mar 2013.

[224] S. Sundaram, A. R. Johnson, and L. Makowski. Obesity, metabolism and the microenvironment: Links to cancer. *J. Carcinog.*, 12:19, 2013.

[225] J. Yun, C. Rago, I. Cheong, R. Pagliarini, P. Angenendt, H. Rajagopalan, K. Schmidt, J. K. V. Willson, S. Markowitz, S. Zhou, L. A. Diaz, Jr, V. E. Velculescu, C. Lengauer, K. W. Kinzler, B. Vogelstein, and N. Papadopoulos. Glucose deprivation contributes to the development of KRAS pathway mutations in tumor cells. *Science*, 325(5947):1555–9, Sep 2009.

[226] O. Warburg. The chemical constitution of respiration ferment. *Science*, 68(1767):437–43, Nov 1928.

[227] O. Warburg, F. Wind, and E. Negelein. The metabolism of tumors in the body. *J. Gen. Physiol.*, 8(6):519–30, Mar 1927.

[228] W. H. Koppenol, P. L. Bounds, and C. V. Dang. Otto Warburg's contributions to current concepts of cancer metabolism. *Nat. Rev. Cancer*, 11(5):325–37, May 2011.

[229] L. Stryer. *Biochemistry*, chapter Glycolysis, pages 483–508. W.H. Freeman and Company, New York, 4th edition, 1995.

[230] E. Meléndez-Hevia, T. G. Waddell, and M. Cascante. The puzzle of the krebs citric acid cycle: assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. *J. Mol. Evol.*, 43(3):293–303, Sep 1996.

[231] L. L. Spriet, R. A. Howlett, and G. J. Heigenhauser. An enzymatic approach to lactate production in human skeletal muscle during exercise. *Med. Sci. Sports Exerc.*, 32(4):756–63, Apr 2000.

[232] O. Warburg. On the origin of cancer cells. *Science*, 123(3191):309–14, Feb 1956.

[233] V. R. Fantin, J. St-Pierre, and P. Leder. Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology, and tumor maintenance. *Cancer Cell*, 9(6):425–34, Jun 2006.

[234] R. Moreno-Sánchez, S. Rodríguez-Enríquez, A. Marín-Hernández, and E. Saavedra. Energy metabolism in tumor cells. *FEBS J.*, 274(6):1393–418, Mar 2007.

[235] X. Tong, F. Zhao, and C. B. Thompson. The molecular determinants of de novo nucleotide biosynthesis in cancer cells. *Curr. Opin. Genet. Dev.*, 19(1):32–7, Feb 2009.

[236] J. W. Locasale, A. R. Grassian, T. Melman, C. A. Lyssiotis, K. R. Mattaini, A. J. Bass, G. Heffron, C. M. Metallo, T. Muranen, H. Sharfi, A. T. Sasaki, D. Anastasiou, E. Mullarky, N. I. Vokes, M. Sasaki, R. Beroukhim, G. Stephanopoulos, A. H. Ligon, M. Meyerson, A. L. Richardson, L. Chin, G. Wagner, J. M. Asara, J. S. Brugge, L. C. Cantley, and M. G. Vander Heiden. Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nat. Genet.*, 43(9):869–74, Sep 2011.

[237] D. Anastasiou, G. Poulogiannis, J. M. Asara, M. B. Boxer, J. K. Jiang, M. Shen, G. Bellinger, A. T. Sasaki, J. W. Locasale, D. S. Auld, C. J. Thomas, M. G. Vander Heiden, and L. C. Cantley. Inhibition of pyruvate kinase M2 by reactive oxygen species contributes to cellular antioxidant responses. *Science*, 334(6060):1278–83, Dec 2011.

[238] R. B. Hamanaka and N. S. Chandel. Cell biology. Warburg effect and redox balance. *Science*, 334(6060):1219–20, Dec 2011.

[239] X. Chen, Y. Qian, and S. Wu. The Warburg effect: evolving interpretations of an established concept. *Free Radic. Biol. Med.*, 79:253–63, Feb 2015.

[240] T. Fiaschi and P. Chiarugi. Oxidative stress, tumor microenvironment, and metabolic reprogramming: a diabolic liaison. *Int. J. Cell. Biol.*, 2012:762825, 2012.

[241] N. Maulik. Redox signaling of angiogenesis. *Antioxid. Redox Signal.*, 4(5):805–15, Oct 2002.

[242] P. Storz. Reactive oxygen species in tumor progression. *Front. Biosci.*, 10:1881–96, 2005.

[243] O. Blokhina, E. Virolainen, and K. V. Fagerstedt. Antioxidants, oxidative damage and oxygen deprivation stress: a review. *Ann. Bot.*, 91 Spec No:179–94, Jan 2003.

[244] Z. Darzynkiewicz, L. Staiano-Coico, and M. R. Melamed. Increased mitochondrial uptake of rhodamine 123 during lymphocyte stimulation. *Proc. Natl. Acad. Sci. U.S.A.*, 78(4):2383–7, Apr 1981.

[245] C. J. Hedeskov. Early effects of phytohaemagglutinin on glucose metabolism of normal human lymphocytes. *Biochem. J.*, 110(2):373–80, Nov 1968.

[246] T. Wang, C. Marquardt, and J. Foker. Aerobic glycolysis during lymphocyte proliferation. *Nature*, 261(5562):702–5, Jun 1976.

[247] S. Christen and U. Sauer. Intracellular characterization of aerobic glucose metabolism in seven yeast species by 13C flux analysis and metabolomics. *FEMS Yeast Res*, 11(3):263–72, May 2011.

[248] S. Y. Lunt and M. G. Vander Heiden. Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. *Annu. Rev. Cell Dev. Biol.*, 27:441–64, 2011.

[249] H. R. Christofk, M. G. Vander Heiden, N. Wu, J. M. Asara, and L. C. Cantley. Pyruvate kinase M2 is a phosphotyrosine-binding protein. *Nature*, 452(7184):181–6, Mar 2008.

[250] T. Noguchi, H. Inoue, and T. Tanaka. The M1- and M2-type isozymes of rat pyruvate kinase are produced from the same gene by alternative RNA splicing. *J. Biol. Chem.*, 261(29):13807–12, Oct 1986.

[251] H. R. Christofk, M. G. Vander Heiden, M. H. Harris, A. Ramanathan, R. E. Gerszten, R. Wei, M. D. Fleming, S. L. Schreiber, and L. C. Cantley. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature*, 452(7184):230–3, Mar 2008.

[252] T. Hitosugi, S. Kang, M. G. Vander Heiden, T.-W. Chung, S. Elf, K. Lythgoe, S. Dong, S. Lonial, X. Wang, G. Z. Chen, J. Xie, T.-L. Gu, R. D. Polakiewicz, J. L. Roesel, T. J. Boggon, F. R. Khuri, D. G. Gilliland, L. C. Cantley, J. Kaufman, and J. Chen. Tyrosine phosphorylation inhibits PKM2 to promote the Warburg effect and tumor growth. *Sci. Signal.*, 2(97):ra73, 2009.

[253] D. K. Stammers and H. Muirhead. Three-dimensional structure of cat muscle pyruvate kinase at 3.1Å resolution. *J. Mol. Biol.*, 112(2):309–16, May 1977.

[254] H. Muirhead, D. A. Clayden, D. Barford, C. G. Lorimer, L. A. Fothergill-Gilmore, E. Schiltz, and W. Schmitt. The structure of cat muscle pyruvate kinase. *EMBO J.*, 5(3):475–81, Mar 1986.

[255] T. M. Larsen, M. M. Benning, I. Rayment, and G. H. Reed. Structure of the bis(Mg2+)-ATP-oxalate complex of the rabbit muscle pyruvate kinase at 2.1 Å resolution: ATP binding over a barrel. *Biochemistry*, 37(18):6247–55, May 1998.

[256] J. D. Dombrauckas, B. D. Santarsiero, and A. D. Mesecar. Structural basis for tumor pyruvate kinase M2 allosteric regulation and catalysis. *Biochemistry*, 44(27):9417–29, Jul 2005.

[257] M. S. Jurica, A. Mesecar, P. J. Heath, W. Shi, T. Nowak, and B. L. Stoddard. The allosteric regulation of pyruvate kinase by fructose-1,6-bisphosphate. *Structure*, 6(2):195–210, Feb 1998.

[258] W. J. Israelsen and M. G. Vander Heiden. Pyruvate kinase: Function, regulation and role in cancer. *Semin. Cell Dev. Biol.*, 43:43–51, Jul 2015.

[259] T. J. Bollenbach, A. D. Mesecar, and T. Nowak. Role of lysine 240 in the mechanism of yeast pyruvate kinase catalysis. *Biochemistry*, 38(28):9137–45, Jul 1999.

[260] W. Luo, H. Hu, R. Chang, J. Zhong, M. Knabel, R. O'Meally, R. N. Cole, A. Pandey, and G. L. Semenza. Pyruvate kinase M2 is a PHD3-stimulated coactivator for hypoxia-inducible factor 1. *Cell*, 145(5):732–44, May 2011.

[261] K. Akhtar, V. Gupta, A. Koul, N. Alam, R. Bhat, and R. N. K. Bamezai. Differential behavior of missense mutations in the intersubunit contact domain of the human pyruvate kinase M2 isozyme. *J. Biol. Chem.*, 284(18):11971–81, May 2009.

[262] H. P. Morgan, F. J. O'Reilly, M. A. Wear, J. R. O'Neill, L. A. Fothergill-Gilmore, T. Hupp, and M. D. Walkinshaw. M2 pyruvate kinase provides a mechanism for nutrient sensing and regulation of cell proliferation. *Proc. Natl. Acad. Sci. U.S.A.*, 110(15):5881–6, Apr 2013.

[263] K. H. Ibsen, R. H. Chiu, H. R. Park, D. A. Sanders, S. Roy, K. N. Garratt, and M. K. Mueller. Purification and properties of mouse pyruvate kinases K and M and of a modified K subunit. *Biochemistry*, 20(6):1497–506, Mar 1981.

[264] R. D. Koler and P. Vanbellinghen. The mechanism of precursor modulation of human pyruvate kinase i by fructose diphosphate. *Adv. Enzyme Regul.*, 6:127–42, 1968.

[265] K. Ashizawa, P. McPhie, K. H. Lin, and S. Y. Cheng. An *in vitro* novel mechanism of regulating the activity of pyruvate kinase M2 by thyroid hormone and fructose 1, 6-bisphosphate. *Biochemistry*, 30(29):7105–11, Jul 1991.

[266] J. Schulz, G. Sparmann, and E. Hofmann. Alanine-mediated reversible inactivation of tumour pyruvate kinase caused by a tetramer-dimer transition. *FEBS Lett.*, 50(3):346–50, Feb 1975.

[267] R. Williams, T. Holyoak, G. McDonald, C. Gui, and A. W. Fenton. Differentiating a ligand's chemical requirements for allosteric interactions from those for protein binding. Phenylalanine inhibition of pyruvate kinase. *Biochemistry*, 45(17):5421–9, May 2006.

[268] G. Sparmann, J. Schulz, and E. Hofmann. Effects of L-alanine and fructose (1,6-diphosphate) on pyruvate kinase from ehrlich ascites tumour cells. *FEBS Lett.*, 36(3):305–8, Nov 1973.

[269] K. H. Ibsen and S. W. Marles. Inhibition of chicken pyruvate kinases by amino acids. *Biochemistry*, 15(5):1073–9, Mar 1976.

[270] K. E. Keller, I. S. Tan, and Y.-S. Lee. SAICAR stimulates pyruvate kinase isoform M2 and promotes cancer cell survival in glucose-limited conditions. *Science*, 338(6110):1069–72, Nov 2012.

[271] K. E. Keller, Z. M. Doctor, Z. W. Dwyer, and Y.-S. Lee. SAICAR induces protein kinase activity of PKM2 that is necessary for sustained proliferative signaling of cancer cells. *Mol. Cell.*, 53(5):700–9, Mar 2014.

[272] W. Yang, Y. Zheng, Y. Xia, H. Ji, X. Chen, F. Guo, C. A. Lyssiotis, K. Aldape, L. C. Cantley, and Z/ Lu. ERK1/2-dependent phosphorylation and nuclear translocation of PKM2 promotes the Warburg effect. *Nat. Cell. Biol.*, 14(12):1295–304, Dec 2012.

[273] C. Kung, J. Hixon, S. Choe, K. Marks, S. Gross, E. Murphy, B. DeLaBarre, G. Cianchetta, S. Sethumadhavan, X. Wang, S. Yan, Y. Gao, C. Fang, W. Wei, F. Jiang, S. Wang, K. Qian, J. Saunders, E. Driggers, H. K. Woo, K. Kunii, S. Murray, H. Yang, K. Yen, W. Liu, L. C. Cantley, M. G. Vander Heiden, S. M. Su, S. Jin, F. G. Salituro, and L. Dang. Small molecule activation of PKM2 in cancer cells induces serine auxotrophy. *Chem. Biol.*, 19(9):1187–98, Sep 2012.

[274] K. M. Parnell, J. M. Foulks, R. N. Nix, A. Clifford, J. Bullough, B. Luo, A. Senina, D. Vollmer, J. Liu, V. McCarthy, Y. Xu, M. Saunders, X.-H. Liu, S. Pearce, K. Wright, M. O'Reilly, M. V. McCullar, K.-K. Ho, and S. B. Kanner. Pharmacologic activation of PKM2 slows lung tumor xenograft growth. *Mol. Cancer. Ther.*, 12(8):1453–60, Aug 2013.

[275] G. H. Reed and S. D. Morgan. Kinetic and magnetic resonance studies of the interaction of oxalate with pyruvate kinase. *Biochemistry*, 13(17):3537–41, Aug 1974.

[276] L. Lv, D. Li, D. Zhao, R. Lin, Y. Chu, H. Zhang, Z. Zha, Y. Liu, Z. Li, Y. Xu, G. Wang, Y. Huang, Y. Xiong, K.-L. Guan, and Q.-Y. Lei. Acetylation targets the M2 isoform of pyruvate kinase for degradation through chaperone-mediated autophagy and promotes tumor growth. *Mol. Cell.*, 42(6):719–30, Jun 2011.

[277] Y.-P. Lv, L.and Xu, D. Zhao, F.-L. Li, W. Wang, N. Sasaki, Y. Jiang, X. Zhou, T.-T. Li, K.-L. Guan, Q.-Y. Lei, and Y. Xiong. Mitogenic and oncogenic stimulation of K433 acetylation promotes pkm2 protein kinase activity and nuclear localization. *Mol. Cell.*, 52(3):340–52, Nov 2013.

[278] M. Anitha, G. Kaur, N. Z. Baquer, and R. Bamezai. Dominant negative effect of novel mutations in pyruvate kinase-M2. *DNA Cell Biol.*, 23(7):442–9, Jul 2004.

[279] V. Gupta, P. Kalaiarasan, M. Faheem, N. Singh, M. A. Iqbal, and R. N. K. Bamezai. Dominant negative mutations affect oligomerization of human pyruvate kinase M2 isozyme and promote cellular growth and polyploidy. *J. Biol. Chem.*, 285(22):16864–73, May 2010.

[280] M. A. Iqbal, F. A. Siddiqui, N. Chaman, V. Gupta, B. Kumar, P. Gopinath, and R. N. K. Bamezai. Missense mutations in pyruvate kinase M2 promote

cancer metabolism, oxidative endurance, anchorage independence, and tumor growth in a dominant negative manner. *J. Biol. Chem.*, 289(12):8098–105, Mar 2014.

[281] P. Wang, C. Sun, T. Zhu, and Y. Xu. Structural insight into mechanisms for dynamic regulation of PKM2. *Protein Cell*, 6(4):275–87, Apr 2015.

[282] A. Bartesaghi, D. Matthies, S. Banerjee, A. Merk, and S. Subramaniam. Structure of β-galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.*, 111(32):11709–14, Aug 2014.

[283] S. H. W. Scheres. A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.*, 415(2):406–18, Jan 2012.

[284] G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, and S. J. Ludtke. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.*, 157(1):38–46, Jan 2007.

[285] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.*, 25(13):1605–12, Oct 2004.

[286] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, 66(Pt 4):486–501, Apr 2010.

[287] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, Ralf W. G.-K., A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystall.*, 66(Pt 2):213–21, Feb 2010.

[288] A. Kucukelbir, F. J. Sigworth, and H. D. Tagare. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods*, 11(1):63–5, Jan 2014.

[289] B. Chaneton, P. Hillmann, L. Zheng, A. C. L. Martin, O. D. K. Maddocks, A. Chokkathukalam, J. E. Coyle, A. Jankevics, F. P. Holding, K. H. Vousden, C. Frezza, M. O'Reilly, and E. Gottlieb. Serine is a natural ligand and allosteric activator of pyruvate kinase M2. *Nature*, 491(7424):458–62, Nov 2012.

[290] S. Khrapunov. Circular dichroism spectroscopy has intrinsic limitations for protein secondary structure analysis. *Anal. Biochem.*, 389(2):174–6, Jun 2009.

[291] M. Liao, E. Cao, D. Julius, and Y. Cheng. Single particle electron cryomicroscopy of a mammalian ion channel. *Curr. Opin. Struct. Biol.*, 27:1–7, Aug 2014.

[292] S. Mazurek. Pyruvate kinase type M2: a key regulator within the tumour metabolome and a tool for metabolic profiling of tumours. *Ernst Schering Found. Symp. Proc.*, 2007(4):99–124, Jun 2008.

[293] X.-C. Bai, C. Yan, G. Yang, P. Lu, D. Ma, L. Sun, R. Zhou, S. H. W. Scheres, and Y. Shi. An atomic structure of human γ-secretase. *Nature*, 525(7568):212–7, Sep 2015.

[294] X.-C. Bai, E. Rajendra, G. Yang, Y. Shi, and S. H. W. Scheres. Sampling the conformational space of the catalytic subunit of human γ-secretase. *eLife*, 4, 2015.

[295] A. Shamirian, A. Ghai, and P. T. Snee. QD-based FRET probes at a glance. *Sensors*, 15(6):13028–51, 2015.