Exploring Mechanisms of Cell Shape Control in *Helicobacter pylori*

Kris Blair

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Nina Salama, Chair

Barry Stoddard

Joseph Mougous

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Exploring Mechanisms of Cell Shape Control in *Helicobacter pylori*

Kris Blair

Chair of the Supervisory Committee:
Professor Nina Salama
Molecular and Cellular Biology, University of Washington
Division of Human Biology

Helicobacter pylori chronically infects over half of the global population and is associated with increased risk for gastric cancer, the third leading cause of cancer deaths worldwide. *H. pylori* resides in the gastric mucosa of the human stomach where it utilizes helical cell shape to efficiently colonize the epithelium. However, the underlying molecular mechanisms that *H. pylori* bacteria use to generate and control their helical cell shape are not known. In this dissertation, I took a multidisciplinary approach using genetic, biochemical, and proteomic approaches to discover and characterize protein interactions among known cell shape determining (Csd) proteins to gain mechanistic insight into helical cell shape determination. We

discovered that the non-enzymatic protein Csd5 comprises a multi-protein "Shapeosome" complex that resides in the cytoplasmic membrane where it interacts with ATP synthase, a cytoplasmic peptidoglycan synthase (MurF), a bacterial cytoskeleton protein (CcmA), and the bacterial cell wall. We also performed structure-function studies of Csd5 and the peptidoglycan hydrolase Csd4 to identify functional domains required for protein function. Finally, we show that *H. pylori* isolates obtained from a chronically infected host exhibit significant heterogeneity in cell shape and find that mutations in cell shape determining genes contribute to this cell shape diversity.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

"For the strength of the Pack is the Wolf, and the strength of the Wolf is the Pack."
-Rudyard Kipling, The Jungle Book

This dissertation could not have been completed without the dedication and contributions of many individuals with whom I have had the privilege to know and learn from over the past 15 years as a scientist. An extra special thank you to past and present members of the Salama lab, all of whom I have had a BLAST commiserating with over cake, ice cream, and broken incubators: Jenny Taylor, Desirée Yang, Tina Gall, Ilana Cohen, Jutta Fero, Sarah Talarico, Christina Leverich, Laura Jackson, Kevin Mears, and many others. Former Fred Hutch summer interns: Jesse Domingo, Max Ruben, Daniel Wright, Alex Shang and Andrew Dominguez who all made very meaningful contributions to the science and my enjoyment sharing the successes and failures of science. To all the collaborators and Fred Hutch support staff: John Whitney, Anson Chan, Rick Peek Jr., Phil Gafken, Lisa Jones, Elizabeth Jensen, Luna Yu, Pat Heath and Pam Lindberg, and Cathy Lin. The MCB Program for making graduate school a wonderful experience: Michele Karantsavelos, MaryEllin Robinson, Maia Low, Nomi Odano, and Marci Burden. I also would like to thank my friends and all of my MCB Class of 2011 colleagues; John "J-La" LaMacchia and Andrea Casasola-LaMacchia, Amitabha "Guppy" Gupta, and Oliver Fregoso. I thank my wife and best friend Claire Gonzalez, and my mother Cathy Wood for their sacrifices on this journey. Thank you to my former mentor Dan Kearns who first showed me how to succeed in the lab. But most importantly, I want to thank my advisor Nina Salama for teaching me, letting me fail, and letting me succeed. Without your openness, abundant patience and unmatched curiosity to understand how life works I would not have become the scientist I am today. I am forever grateful to have shared this experience with each and every one of you.

DEDICATION

This dissertation is dedicated to my wife Claire

You Are My Moon and My Shining Star

Without your love and unwavering support

this achievement could not have been possible.

I am forever grateful for you, and our love and companionship.

Thank You

Chapter 1. THE WHY AND HOW OF BACTERIAL CELL SHAPE

1.1 PREFACE

Portions of this chapter are adapted with appropriate permissions of a previously published literature review in collaboration with my colleague Dr. Desiree C. Yang and my Ph.D. advisor Dr. Nina Salama.

> Yang, D.C., Blair, K.M. and Salama, N.R., 2016. Staying in shape: the impact of cell shape on bacterial survival in diverse environments. *Microbiology and Molecular Biology Reviews*, *80*(1), pp.187-203.

> "*The mystery of the beginning of all things is insoluble by us; and I for one must be content to remain an agnostic." – Charles Darwin*

1.2 THE IMPORTANCE OF CELL SHAPE

Helicobacter pylori is a carcinogenic bacterial pathogen that chronically infects more than half of the global human population and is a significant risk factor for the development of stomach cancer, the third leading cause of cancer death worldwide (Ferlay *et al.*, 2015). The helical cell shape of organisms like *H. pylori* has been hypothesized to promote enhanced motility through viscous solutions or gels via a corkscrew or burrowing model of motility (Greenberg and Canale-Parola, 1977). Until recently, there was little experimental evidence to support this theory beyond the early and foundational experiments investigating the motility of spirochetes in high viscosity solutions of polyvinylpyrrolidone or methylcellulose. Greenberg

and Canale-Parola measured the minimum inhibitory viscosity (MIV) of these organisms and found that their translational motility was enhanced under conditions of high viscosity, unlike their motile and flagellated *E. coli* counterparts, in which motility was inhibited in high-viscosity solutions. Furthermore, cell shape mutants of *S. aurantia* with reduced coiling (i.e., reduced helicity) but otherwise normal motility had an MIV of 1/10 that of the normally coiled parental strain, suggesting a direct role for helical shape in promoting motility through a viscous environment (Greenberg and Canale-Parola, 1977).

Helical shapes have arisen multiple times during the evolution of bacteria and are represented by distinct groups of bacterial pathogens that include the spirochetes (e.g., *Treponema*, *Spirochaeta*, *Leptospira*, and *Borrelia* spp.) and the proteobacteria (e.g., gastric and non-gastric *Helicobacter* spp. and the foodborne pathogen *Campylobacter jejuni*), all of which cause a wide variety of diseases in humans and animals. It has been shown for both helical proteobacteria and spirochetes that a loss of motility and/or helical cell shape causes major defects in the colonization potential of these organisms (Goodwin and Shedlarski, 1975; Harman *et al.*, 2012; Frirdich *et al.*, 2012).

More recently, studies of the pathogens *H. pylori* and *C. jejuni* provided further support for the functional importance of helical cell shape in the natural environment of the host. For *H. pylori*, non-helical cell shape mutants exhibited stomach colonization defects in a murine colonization model. In these experiments, mice were challenged with equal amounts of wild-type (helical) and mutant (curved-rod) *H. pylori* bacteria. After a 1-week infection, the helical bacteria outnumbered the non-helical mutants by a magnitude of 1 $\log_{10}$, demonstrating that helical shape is an important pathogenesis factor (Sycuro *et al.*, 2010; Bonis *et al.*, 2010).

1.3 MORPHOLOGICAL VARIATION

Morphological plasticity as a mechanism for pathogens to better adapt to different environments may explain the enormous cell shape heterogeneity seen in clonal isolates of *H. pylori* (Figure 1.1) (Sycuro *et al.*, 2010). Although *H. pylori* is described as a helical rod, cultures grown from clonal isolates display a range of morphologies, including both curved and straight rods of various lengths and helicities (Sycuro *et al.*, 2010). It has been postulated that this heterogeneity of shape allows this pathogen to infect and adapt to different stomach niches depending on the particular individual that is infected. Given the diverse shapes observed for each clonal isolate, a single *H. pylori* strain may be poised for more successful transmission between hosts because, presumably, there will always be cells with different morphologies that are better suited for different gastric environments.

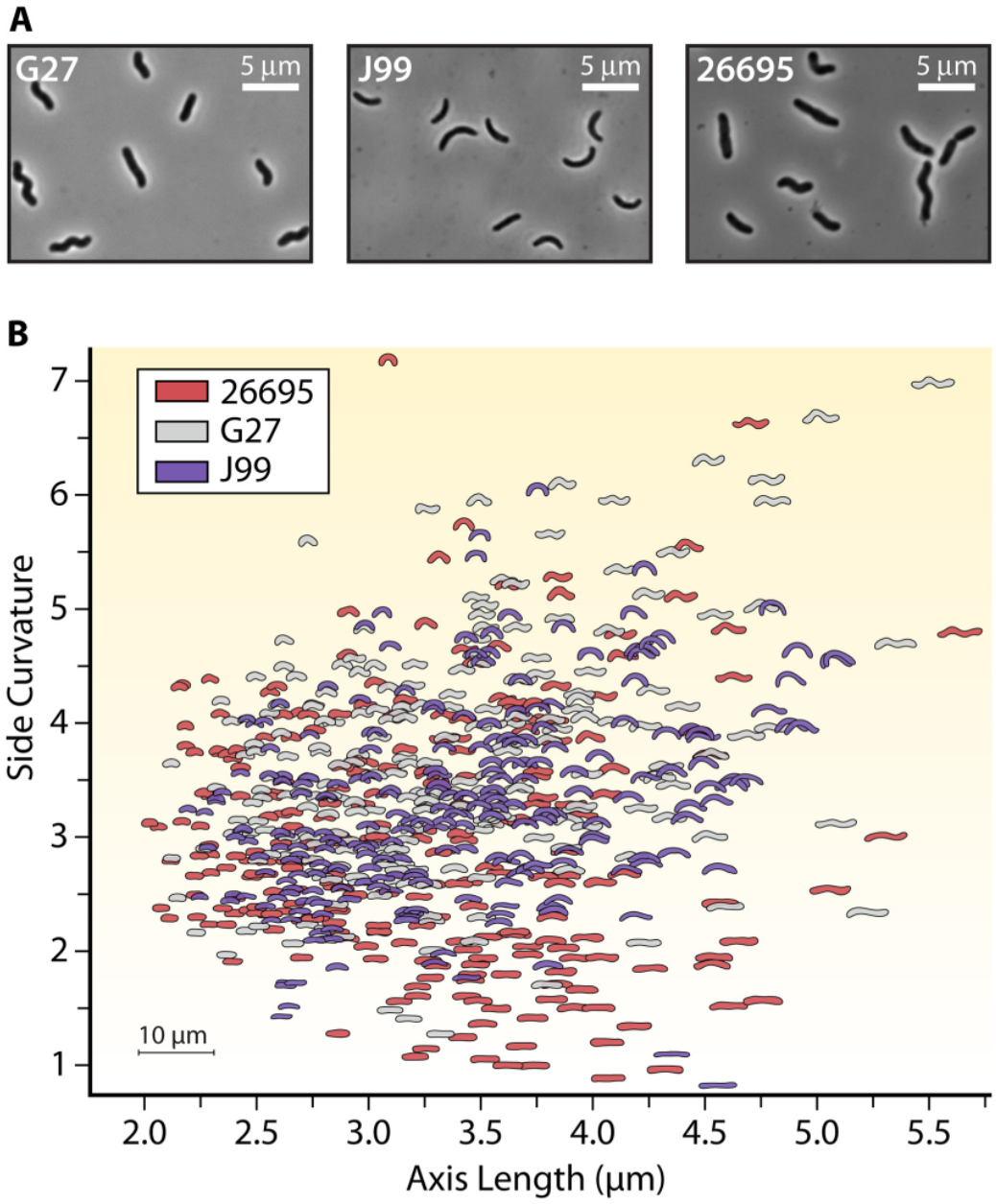

Figure 1.1. *Helicobacter pylori* intraspecies cell shape heterogeneity.

A. Phase-contrast images of three clinical strains of *Helicobacter pylori* (Tomb *et al.*, 1997; Alm *et al.*, 1999; Baltrus *et al.*, 2009). B. Quantitative Celltool scatterplot analysis (Sycuro *et al.*, 2010) showing cell side curvature versus axis length for clonal populations of the indicated strains of *H. pylori* (100 cells/strain).

1.4 THE SHAPE OF BACTERIA

In most bacteria, the cell wall determines the shape of the cell. Peptidoglycan (PG) is the major structural component of the cell wall in both Gram-positive and Gram-negative bacteria. Thus, much research in the bacterial morphology field focuses on defining the mechanisms by which the cell wall is synthesized and modified to produce different cell body shapes. The reader is referred to several excellent reviews on cell wall synthesis and the evolution of bacterial morphogenesis for an in-depth consideration of this topic (Scheffers and Pinho, 2005; Typas *et al.*, 2012; Jiang *et al.*, 2015; Randich and Brun, 2015). The PG polymer consists of a repeating disaccharide with a short peptide stem attached to one of the sugars (Figure 1.2). Transpeptidation between peptide stems can link the sugar strands together, forming a rigid meshwork that encases the cell and prevents lysis due to turgor pressure. The biosynthesis of this essential structure has been an important antibiotic target and an active area of research, which is the subject of several excellent reviews (Typas *et al.*, 2012; Lovering *et al.*, 2012; Desmarais *et al.*, 2013). In Gram-negative bacteria, the PG cell wall resides in the periplasm between the inner and outer membranes. Multiple lines of evidence suggest that the PG layer in Gram-negative bacteria is relatively thin, perhaps a single monolayer, with the glycan strands oriented primarily in a circumferential direction as a disordered lattice (Figure 1.2) (Vollmer and Höltje, 2004; Gan *et al.*, 2008; Typas *et al.*, 2012; Lovering *et al.*, 2012; Desmarais *et al.*, 2013; Gumbart *et al.*, 2014). The outer membrane of Gram-negative bacteria protects the PG from lytic enzymes produced by the host or other microorganisms. PG is thought to be the structural determinant of cell shape because purified PG sacculi retain the shape of the cell from which they are isolated, be it coccoid or a straight, curved, or helical rod (Goodwin and Shedlarski, 1975; de Pedro *et al.*, 1997; Sycuro *et al.*, 2010). The major theoretical model of shape generation for the cell wall

posits asymmetric synthesis driven by cytoskeletal elements that direct the activity of PG synthesis enzymes to particular regions of the cell (Scheffers and Pinho, 2005), though mathematical modeling has suggested that changes in cross-linking could also drive alternate shapes (Figure 1.2) (Huang *et al.*, 2008).

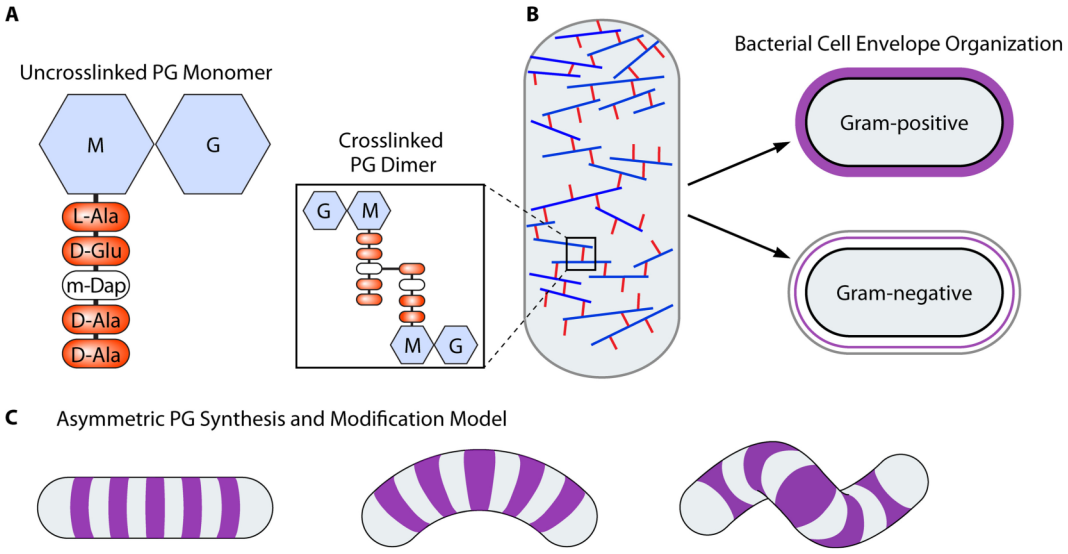

Figure 1.2. Cell wall composition and architecture in different bacterial classes and shapes. A. Monomeric and cross-linked peptidoglycan (PG) structure depicting the *N*-acetylglucosamine (G)–*N*-acetylmuramic acid (M) disaccharide with a cross-link mediating *meso*-diaminopimelic acid (*m*-DAP) in the third position of the pentapeptide stem. B. Glycan strands (repeating PG disaccharide subunits) are generally arranged circumferentially around the cell in both Gram-positive and Gram-negative bacteria. Gram-positive bacteria have a thick layer of PG (shown in purple), while Gram-negative bacteria have a thin layer between the inner (black outline) and outer (gray outline) membranes. C. Model for the generation of curvature and twist in curved and helical bacteria whereby asymmetric wedges or zones of PG synthesis and/or modification may promote increased or decreased PG incorporation on one side or region of the cell.

Interestingly, helical cell shapes appear to be found mainly in Gram-negative bacteria and this restriction of helical morphologies in Gram-negative bacteria is likely a consequence of the

biophysical properties of the PG cell wall that encases the cell. Because Gram-positive bacteria possess a thick, multilayered PG sacculus surrounding the cell, it is likely a more rigid structure than the single layer of PG found in Gram-negative bacteria (Figure 1.2). In order to deviate from a straight-rod morphology, there must be a mechanism of asymmetric incorporation of new PG into the cell wall and/or localized changes in PG cross-linking (Sycuro *et al.*, 2010). Consistent with this hypothesis, the majority of the genes involved in cell shape determination in the proteobacteria *H. pylori* and *C. jejuni* alter the crosslinking or stem-peptide composition of the cell wall when absent or overexpressed (Sycuro *et al.*, 2010; Sycuro *et al.*, 2012; Frirdich *et al.*, 2012; Sycuro *et al.*, 2013; Frirdich *et al.*, 2014; Chan *et al.*, 2015).

1.5 *HELICOBACTER PYLORI* CELL SHAPE PROTEINS

Visual and flow cytometry (FACS) screens of an *H. pylori* mutant library identified genes (*csd4*, *csd5* and *csd6*) that when mutated, gave rise to straight-rod phenotypes and reduced colonization potential (Sycuro *et al.*, 2010; Sycuro *et al.*, 2012). Also identified in these screens were endopeptidases that cleave PG crosslinks (*csd1, csd2*[†]*, csd3*) and a non-enzymatic bactofilin homolog (*ccmA)* that result in curved rod shapes when disrupted (Figure 1.3). There is evidence for genetic and functional conservation of helical shape generating pathways. For example, *pgp1*, a homolog of the *H. pylori csd4* gene, was identified in *C. jejuni* and shown to be important for proper cell shape (Sycuro *et al.*, 2012; Frirdich *et al.*, 2012). Analysis of the muropeptide (PG) content of these mutants demonstrated that Csd6 cleaves monomeric tetrapeptides to produce the tripeptide substrate for Csd4. Csd4 then acts to cleave the terminal mDAP moiety resulting in PG dipeptides that are necessary to form crosslinks. Curiously, the

[†]Csd2 is enzymatically inactive but is required for Csd1 stability (personal communication with Desiree Yang and Nina Salama.

muropeptide composition of *csd5* mutants with the shared straight-rod phenotype of *csd4* and *csd6* mutants, revealed a PG content profile that was identical to wild-type. As Csd5 is predicted to be a non-enzymatic scaffolding protein, we hypothesized that its role in shape determination may be to promote localized modifications to the cell wall by Csd4 and Csd6 enzymes necessary for helical shape determination. The overarching aim of this body of work is to the explore mechanisms of *H. pylori* cell shape determination through discovery and characterization of cell shape protein interactions to gain insight regarding the molecular mechanisms that underlie the helical shape determination of *Helicobacter pylori.*

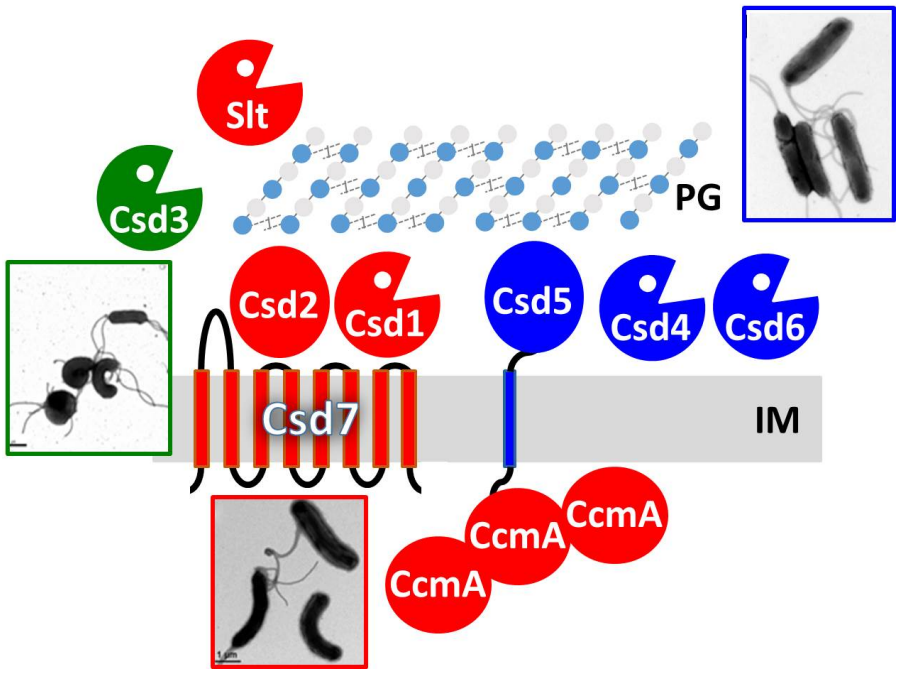

Figure 1.3. Summary of *H. pylori* Csd protein mutant phenotypes, predicted topology and PG hydrolase activity.

Colors indicate shared mutant shape phenotypes (*curved: red, straight: blue, variable: green*). PacMan shapes indicate verified PG hydrolase activity. PG: peptidoglycan, IM: inner membrane. Figure courtesy of Nina Salama.

1.6 SUMMARY

At the beginning of this work our knowledge was limited to a list of known cell shape proteins and a basic understanding of their individual function. In Chapter 2 we describe a structure-function analysis of Csd4 and report the crystal structure of the Csd4 enzyme with bound tripeptide substrate, demonstrate that the two C-terminal domains of unknown function are required for Csd4 protein stability and function, and that an atypical glutamine residue in the active site is required for coordination of a zinc ligand and its enzymatic activity. In Chapter 3 we performed a structure-function analysis of Csd5 to identify interaction partners and find that instead of interaction with periplasmic hydrolases Csd4 and Csd6, Csd5 is part of a multi-protein cell shape promoting complex that spans the inner membrane to bridge cytoplasmic cell shape and PG biosynthetic proteins to the ATP synthase and the peptidoglycan cell wall. These findings along with additional proteomic screens described in Chapter 5 reveal additional candidate protein interactions with other cell shape proteins, molecular chaperones, and proteins of unknown function. These finding suggest the existence of a complex network of protein interactions and complexes with undefined roles in cell shape determination. And finally, in Chapter 4 we present evidence of extensive cell shape variation of single colony isolates obtained from a single human host and uncover genetic variation that may explain the observed cell shapes of these stains.

1.7 BIBLIOGRAPHY

Alm, R.A., Ling, L.-S.L., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., *et al.* (1999) Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* **397**: 176–180.

Baltrus, D.A., Amieva, M.R., Covacci, A., Lowe, T.M., Merrell, D.S., Ottemann, K.M., *et al.* (2009) The complete genome sequence of *Helicobacter pylori* strain G27. *J Bacteriol* **191**: 447–448.

Bonis, M., Ecobichon, C., Guadagnini, S., Prevost, M.C., and Boneca, I.G. (2010) A M23B family metallopeptidase of *Helicobacter pylori* required for cell shape, pole formation and virulence. *Molecular Microbiology* **78**: 809–819.

Chan, A.C.K., Blair, K.M., Liu, Y., Frirdich, E., Gaynor, E.C., Tanner, M.E., *et al.* (2015) Helical shape of *Helicobacter pylori* requires an atypical glutamine as a zinc ligand in the carboxypeptidase Csd4. *J Biol Chem* **290**: 3622–3638.

de Pedro, M.A., Quintela, J.C., Höltje, J.V., and Schwarz, H. (1997) Murein segregation in *Escherichia coli*. *J Bacteriol* **179**: 2823–2834.

Desmarais, S.M., de Pedro, M.A., Cava, F., and Huang, K.C. (2013) Peptidoglycan at its peaks: How chromatographic analyses can reveal bacterial cell wall structure and assembly. *Molecular Microbiology* **89**: 1–13.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., *et al.* (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* **136**: E359–86.

Frirdich, E., Biboy, J., Adams, C., Lee, J., Ellermeier, J., Gielda, L.D., *et al.* (2012) Peptidoglycan-modifying enzyme Pgp1 is required for helical cell shape and pathogenicity traits in *Campylobacter jejuni*. *PLoS Pathog* **8**: e1002602

Frirdich, E., Vermeulen, J., Biboy, J., Soares, F., Taveirne, M.E., Johnson, J.G., *et al.* (2014) Peptidoglycan LD-carboxypeptidase Pgp2 influences *Campylobacter jejuni* helical cell shape and pathogenic properties and provides the substrate for the DL-carboxypeptidase Pgp1. *J Biol Chem* **289**: 8007–8018

Gan, L., Chen, S., and Jensen, G.J. (2008) Molecular organization of Gram-negative peptidoglycan. *Proc Natl Acad Sci USA* **105**: 18953–18957.

Goodwin, S.D., and Shedlarski, J.G. (1975) Purification of cell wall peptidoglycan of the dimorphic bacterium *Caulobacter crescentus*. *Arch Biochem Biophys* **170**: 23–36.

Greenberg, E.P., and Canale-Parola, E. (1977) Relationship between cell coiling and motility of spirochetes in viscous environments. *J Bacteriol* **131**: 960–969.

Gumbart, J.C., Beeby, M., Jensen, G.J., and Roux, B. (2014) *Escherichia coli* peptidoglycan structure and mechanics as predicted by atomic-scale simulations. *PLoS Comput Biol* **10**: e1003475–10.

Harman, M.W., Dunham-Ems, S.M., Caimano, M.J., Belperron, A.A., Bockenstedt, L.K., Fu, H.C., *et al.* (2012) The heterogeneous motility of the Lyme disease spirochete in gelatin mimics dissemination through tissue. *Proc Natl Acad Sci USA* **109**: 3059–3064.

Huang, K.C., Mukhopadhyay, R., Wen, B., Gitai, Z., and Wingreen, N.S. (2008) Cell shape and cell-wall organization in Gram-negative bacteria. *Proc Natl Acad Sci USA* **105**: 19282–19287.

Jiang, C., Caccamo, P.D., and Brun, Y.V. (2015) Mechanisms of bacterial morphogenesis: Evolutionary cell biology approaches provide new insights. *Bioessays* **37**: 413–425.

Lovering, A.L., Safadi, S.S., and Strynadka, N.C.J. (2012) Structural Perspective of Peptidoglycan Biosynthesis and Assembly. *Annu Rev Biochem* **81**: 451–478.

Randich, A.M., and Brun, Y.V. (2015) Molecular mechanisms for the evolution of bacterial morphologies and growth modes. *Front Microbiol* **6**: 938–13.

Scheffers, D.J., and Pinho, M.G. (2005) Bacterial Cell Wall Synthesis: New insights from

localization studies. *Microbiology and Molecular Biology Reviews* **69**: 585–607.

Sycuro, L.K., Pincus, Z., Gutierrez, K.D., Biboy, J., Stern, C.A., Vollmer, W., and Salama, N.R. (2010) Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori's* helical shape and stomach colonization. *Cell* **141**: 822–833.

Sycuro, L.K., Rule, C.S., Petersen, T.W., Wyckoff, T.J., Sessler, T., Nagarkar, D.B., *et al.* (2013) Flow cytometry-based enrichment for cell shape mutants identifies multiple genes that influence *Helicobacter pylori* morphology. *Molecular Microbiology* **90**: 869–883.

Sycuro, L.K., Wyckoff, T.J., Biboy, J., Born, P., Pincus, Z., Vollmer, W., and Salama, N.R. (2012) Multiple peptidoglycan modification networks modulate *Helicobacter pylori's* cell shape, motility, and colonization potential. *PLoS Pathog* **8**: e1002603.

Tomb, J.-F., Tomb, J.F., White, O., White, O., Kerlavage, A.R., Kerlavage, A.R., *et al.* (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**: 539–547.

Typas, A., Banzhaf, M., Gross, C.A., and Vollmer, W. (2012) From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Micro* **10**: 123–136.

Vollmer, W., and Höltje, J.V. (2004) The architecture of the murein (peptidoglycan) in Gram-negative bacteria: Vertical scaffold or horizontal layer(s)? *J. Bacteriol* **186**: 5978–5987.

Chapter 2. HELICAL SHAPE OF *HELICOBACTER PYLORI* REQUIRES AN ATYPICAL GLUTAMINE AS A ZINC LIGAND IN THE CARBOXYPEPTIDASE CSD4

2.1 PREFACE

This Chapter was previously published as:

> Chan, A.C., Blair, K.M., Liu, Y., Frirdich, E., Gaynor, E.C., Tanner, M.E., Salama, N.R. and Murphy, M.E., 2015. Helical shape of *Helicobacter pylori* requires an atypical glutamine as a zinc ligand in the carboxypeptidase Csd4. *Journal of Biological Chemistry*, *290*(6), pp.3622-3638.

As a co-first author on the aforementioned manuscript I am including only the methods for the experiments I performed in producing the manuscript. The Csd4 bioinformatic analysis, crystallography (including protein production), enzymatic assays, and writing of the manuscript were performed by Anson C. Chan. My contribution to this work was the construction of *H. pylori* mutant strains, immunoblotting, morphological analysis, and writing of the manuscript. Please refer to the full publication for additional details pertaining to all procedures, including synthesis of the tripeptide substrate.

2.2 ABSTRACT

Peptidoglycan modifying carboxypeptidases (CPs) are important determinants of bacterial cell shape. Here, we report crystal structures of Csd4, a three-domain protein from the human gastric pathogen *Helicobacter pylori*. The catalytic zinc in Csd4 is coordinated by a rare His-Glu-Gln

configuration that is conserved among most Csd4 homologs, which form a distinct subfamily of CPs. Substitution of the glutamine to histidine, the residue found in prototypical zinc carboxypeptidases, resulted in decreased enzyme activity and inhibition by phosphate. Expression of the histidine variant at the native locus in a *H. pylori csd4* deletion strain did not restore the wild-type helical morphology. Biochemical assays show that Csd4 can cleave a tripeptide peptidoglycan substrate analog to release *m*-DAP. Structures of Csd4 with this substrate analog or product bound at the active site reveal determinants of peptidoglycan specificity and the mechanism to cleave an isopeptide bond to release *m*-DAP. Our data suggest that Csd4 is the archetype of a new CP subfamily with a domain scheme that differs from this large family of peptide-cleaving enzymes.

2.3 INTRODUCTION

The peptidoglycan (PG) sacculus encases the cell and is responsible for maintaining cellular shape and protecting against osmotic stress in most bacteria (Vollmer *et al.*, 2008). PG is a heteropolymer of glycan chains with attached short peptides that are cross-linked to form a rigid, mesh-like structure that retains cell shape when purified (Sycuro *et al.*, 2010). This PG layer undergoes constant remodeling during cell growth, requiring enzymes that cleave (PG hydrolases) and grow (PG synthases) the existing PG structure. The combined effect of PG-modifying enzymes leads to diverse cellular morphologies, such as stars, squares, and spirals (Young, 2006). However, much of our understanding of cell shape determination involves well studied organisms that are cocci, rods, or vibroids, and we are only beginning to identify and understand the components involved in determining other bacterial cell shapes.

The PG glycan chains consist of alternating β-(1-4) linked GlcNAc and *N*-acetylmuramic

acid. The peptide component varies in sequence depending on the species, especially for Gram-negative bacteria (Schleifer and Kandler, 1972). In most Gram negative and some Gram positive species, the PG precursor is composed of the pentapeptide L-Ala-D-Glu-*m-DAP*-D-Ala-D-Ala covalently linked to the *N*-acetylmuramic acid moiety (where *m*-DAP is *meso*-1,6-diaminopimelate) (Vollmer *et al.*, 2008). Once attached to the nascent PG, the pentapeptide can be trimmed sequentially to a tetra-, tri-, and dipeptide by DD-, LD-, and DL-carboxypeptidases, respectively. The *m-DAP* and D-Ala residues from two opposing peptide stems can be directly cross-linked by DD-transpeptidases with the subsequent removal of a terminal D-Ala to form an A1γ type linkage according to the classification scheme of Schleifer and Kandler (Schleifer and Kandler, 1972).

Different classes of enzymes modify the intact peptidoglycan layer (Wyckoff *et al.*, 2012; Frirdich and Gaynor, 2013). Periplasmic DD-carboxypeptidases are well characterized in many organisms and include both the penicillin-binding protein family and unrelated β-lactam-insensitive enzymes (Typas *et al.*, 2012). Much less is known about enzymes that have DL-carboxypeptidase (DL-CP) activity. Recently, we have identified DL-CP genes required for helical shape in the Gram-negative pathogens *Helicobacter pylori* and *Campylobacter jejuni* (Sycuro *et al.*, 2012; Frirdich *et al.*, 2012; Sycuro *et al.*, 2013). *H. pylori* is a specialized Gram negative human pathogen of the class of ε-proteobacteria that colonizes the epithelial surface of the gastric mucous layer and is the causative agent of gastric ulcers and gastric cancer (Salama *et al.*, 2013). Deletion of *H. pylori csd4* led to the loss of helical shape, resulting in slightly curved or straight rods that were impaired in murine stomach colonization assays (Sycuro *et al.*, 2012). A similar study in *C. jejuni*, a closely related human diarrheal pathogen, revealed that deletion of the *csd4* homolog, *pgp1*, led to a similar straight rod phenotype with impaired chick colonization

ability, a model representing the natural reservoir of this pathogen (Frirdich *et al.*, 2012). HPLC analysis of *H. pylori csd4* PG sacculi or purified disaccharide-linked tripeptide digested with Csd4 demonstrated production of a disaccharide dipeptide (*i.e.* DL-carboxypeptidase activity) (Sycuro *et al.*, 2012).

Sequence analysis suggests that *H. pylori* Csd4 contains an N-terminal zinc-containing carboxypeptidase domain of the M14 family and a C-terminal region with unpredicted structural features. Zinc hydrolases have been grouped into over 30 families by Rawlings *et al.* (Rawlings *et al.*, 2012). Family M14 consists of the carboxypeptidases, with the prototypical member being carboxypeptidase A (Lipscomb and Sträter, 1996). One hallmark feature of this family is the zinc-binding motif H*XX*E+H (+H_2O). Zinc coordination lowers the pK_a of the bound catalytic water (Frey and Hegeman, 2007). A second glutamate residue forms a H-bond to the catalytic water molecule and is proposed to abstract a proton forming a hydroxide ion that undergoes a nucleophilic attack on the bound substrate (Gomis-Rüth, 2008).

In this study, we show that Csd4 is a DL-carboxypeptidase with a modified zinc binding site containing a glutamine residue in place of a conserved histidine. To elucidate the hydrolytic mechanism, we have solved crystal structures of Csd4 with bound peptide substrate or product. Crystallographic studies, biochemical assays, and bacterial morphological studies of mutant strains implicate a key role of the glutamine ligand of the active site zinc ion in the formation of helical cell shape. Based on sequence analysis of Csd4 homologs, we propose that Csd4 is the archetype of a new CP subfamily.

2.4 Results and Discussion

Domain Structure of Csd4 —The 1.40 Å resolution Csd4 crystal structure reveals a monomer

with an N-terminal carboxypeptidase domain, followed by two smaller domains of unknown function (domains 2 and 3; Figure 2.1 A and Table 2.1). The CP domain, consisting of residues 1–251 of the native sequence, is globular with a mixed α/β fold containing a nine-stranded antiparallel β-sheet core sandwiched by groups of four and five α-helices. A structural similarity search using DaliLite (Holm and Rosenström, 2010) found significant similarity between the CP domain and the family of funnelin-type carboxypeptidases (*e.g.* human carboxypeptidase B2, PDB code 3LMS, *Z* score 19.1, r.m.s.d. 2.3 Å over 192 aligned residues), a family of metallopeptidases that exhibits CP activity(Gomis-Rüth, 2008). Although the overall funnelin fold was conserved, sequence identities with Csd4 are under 20% over the aligned regions. Domains 2 (residues 252– 343) and 3 (residues 343–438) consist primarily of β-strands with one and two half-turn helices, respectively, in an immunoglobulin-like fold (Figure 2.1 A). Neither domain 2 nor 3 shares significant structural similarity with components found in other carboxypeptidases. Instead, domain 2 shares low structural similarity to a binding domain of human RhoGDI (PDB code 1HH4, *Z* score 3.2, r.m.s.d. 2.9 Å over 65 residues). Domain 3 shares some structural similarity to the periplasmic, non-sugar bound domain of a heparin-sensing two component system (BT4663) of the human gut symbiont *Bacteroides thetaiotaomicron* (PDB code 4A2M, *Z* score 4.9, r.m.s.d. 2.9 Å over 79 residues). Because of weak structural and sequence similarity, domains 2 and 3 are not likely to have similar functions as their top Dali hits. Instead, we predict that these domains participate in protein-protein or protein-PG interactions.

Table 2.1. Data collection and refinement statistics for Csd4

	Apo-Csd4	Zn-Csd4	TriZn-Csd4	Q46H-Csd4	Csd4-initial
Data collection					
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions					
a, *b*, *c* (Å)	53.28, 66.78, 145.53	53.02, 66.77, 144.56	53.23, 66.83, 145.05	53.60 66.65 145.40	53.11 66.92 146.02
Resolution (Å)	42.99-1.40 (1.45-1.40)	42.76-1.85 (1.92-1.85)	49.15-1.75 (1.81-1.75)	49.13-1.75 (1.81-1.75)	50.00-2.10 (2.14-2.10)
R_{merge}	0.070 (0.644)	0.062 (0.434)	0.058 (0.539)	0.078 (0.463)	0.148(0.469)
$I/\sigma I$	17.5 (3.6)	21.4 (4.6)	21.2 (2.8)	15.4 (3.0)	21.2 (2.9)
Completeness (%)	100.0 (99.9)	99.8 (99.3)	99.6 (96.3)	99.8 (97.9)	99.9 (98.3)
Redundancy	7.1	7.1	7.0	7.0	27.4
Refinement					
No. unique reflections	102944	44613	52911	53300	31191
R_{work} / R_{free}	0.14/0.16	0.17/0.20	0.18/0.21	0.18/0.21	0.17/0.22
No. atoms					
Protein	7002	6803	6803	6823	3428
Substrate/Product	25	25	55	25	13
Water	466	401	327	451	384
Average *B*-factors (Å^2)					
Protein	23.1	28.6	30.0	25.0	17.6
Substrate/Product	19.9	21.1	44.9	35.1	14.5
Water	35.1	35.2	37.2	34.5	26.0
R.m.s. deviations					
Bond lengths (Å)	0.013	0.007	0.011	0.012	0.010
Bond angles (°)	1.49	1.09	1.33	1.36	1.41
PDB Accession Code	4WCK	4WCL	4WCN	4WCM	

Domains 2 and 3 are required for stable Csd4 expression because replacement of the endogenous *csd4* gene with a 3x-FLAG-tagged allele truncated at the end of either domain 1 or 2 at the native locus resulted in no detectable expression within whole cell lysates (using an anti-FLAG antibody) and showed a straight rod phenotype similar to the null allele (Figure 2.2). In contrast, replacement with full-length Csd4 containing the same 3x-FLAG epitope revealed the robust expression of the epitope-tagged full-length allele and showed normal morphology. From inspection of the crystal structures, the two domain interfaces (945 Å^2 between domains 1 and 2; 820 Å^2 between domains 2 and 3) are largely hydrophobic, suggesting that both domains are

required to stabilize the protein within cells.

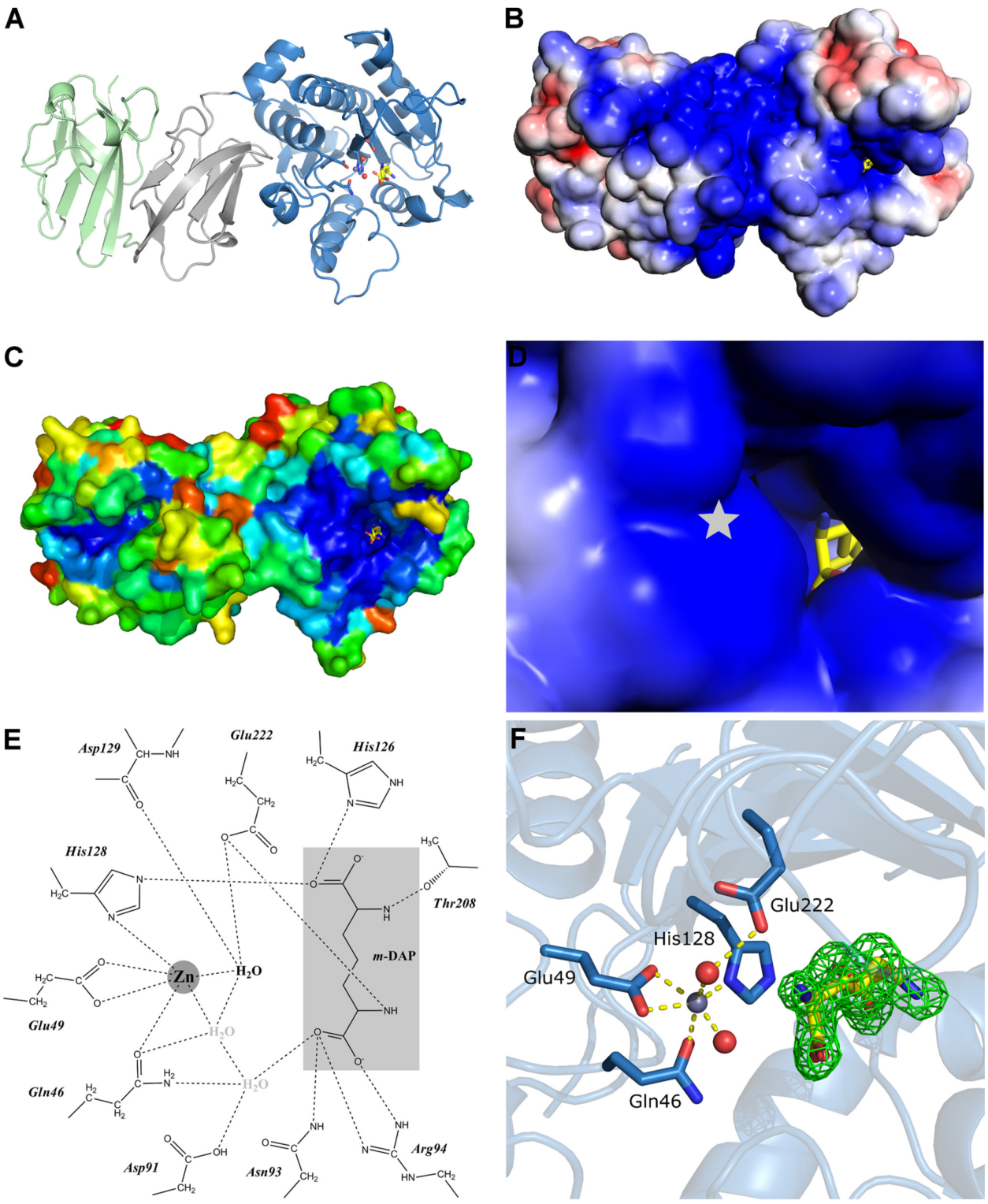

Figure 2.1. The crystal structure of Csd4 (PDB code 4WCL)

A. The overall monomeric structure of Csd4 with Zn^{2+} and *m*-DAP bound. The catalytic domain and domains 2 and 3 are colored *blue*, *grey* and *green*, respectively. B and D. Overall (B) and active-site (D) magnified electrostatic surface potential of Csd4 contoured at ±3 k_bT/e_c. Electropositive regions are colored *blue*; electronegative regions are colored *red*; position of buried Zn^{2+} indicated with a *star*. C. Distribution of conserved residues mapped onto the surface of Csd4. Most conserved regions are colored *blue*; least conserved is colored *red*. E. 2D interaction map between Csd4 and the product *m*-DAP (*light gray*). The predicted catalytic water highlighted in *bold type*. F. Corresponding Zn-Csd4 active site with key ligands and an omit F_o-F_c difference density map for the density of the bound *m*-DAP product contoured at 3 σ.

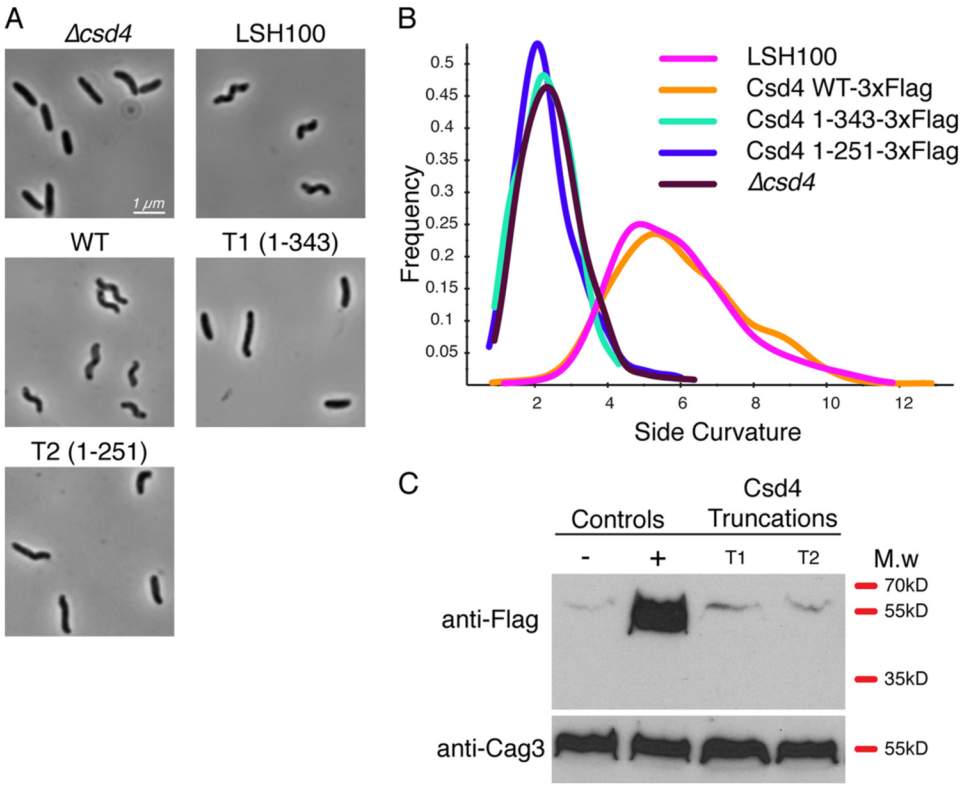

Figure 2.2. Complementation and western blot analysis of Csd4-3xFLAG control, and Csd4 C-terminal truncations.

The strains used were LSH18 (Δ*csd4*), LSH100 WT (no 3x-FLAG tag), KBH19 (WT-3x-FLAG), KBH35 (T1–3x-FLAG), and KBH37 (T2–3x-FLAG). A. 1000X phase contrast images

of wild-type and mutant *H. pylori*. B. Smooth histograms displaying population cell curvature (*x* axis) as a density function (*y* axis). C. Western blot analysis of Csd4 (detected by anti-FLAG M2 antibody); LSH100 WT (-); KBH19 (+) with a predicted molecular mass of 48 kDa; KBH35 (T1) with a predicted molecular mass of 39 kDa; KBH37 (T2) with a predicted molecular mass of 28 kDa; and Cag3 with a predicted molecular mass of 55 kDa. Equivalent amounts of cell extract based on optical density of the culture were loaded for each strain.

PG Peptide Binding to the Active Site of Csd4 —In contrast to all other characterized members of the CP family, the catalytic zinc site of Csd4 consists of Glu, His, and an atypical Gln residue. Zn^2 was soaked into Csd4 crystals to obtain a structure with zinc bound at 90% occupancy (Zn-Csd4). The zinc site is situated along the inner edge of a shallow, positively charged substrate-binding pocket located on the surface of the CP domain (Figure 2.1 B and D). The metal ion is coordinated by His- 128 Nδ1 (2.0 Å), the side chain carboxylate of Glu-49 in a symmetric, bidentate manner (2.3/2.3 Å), and Gln-46 through the carboxamide Oε (2.2 Å) in Zn-Csd4 (Figure 2.1, E and F). Two additional density peaks that are best modeled as either water or hydroxide ions (2.2 and 2.3 Å) are also observed (Table 2.1), together forming an infrequently observed six-coordinate zinc site (Sousa *et al.*, 2009).

The initial crystal structure of Csd4 (Csd4-initial) did not contain bound metal in the active site, as observed for other carboxypeptidases (Hoyland *et al.*, 2014), suggesting weak affinity or insufficient metal availability during protein purification. Zinc binding results in structural changes to the active site ligands. His-128 exhibits an 18° rotation about Cβ to bring the imidazole ring closer to the zinc ion. In the absence of a bound metal, Gln-46 has elevated *B*-factors, and the carboxamide side chain is primarily directed away from the active site. Upon zinc binding, Gln-46 reorients toward the zinc ion, strongly supporting formation of a Gln-46-

Zn^2 ligand interaction. No significant displacement of Glu-49 is observed.

In the substrate-binding pocket of Csd4-initial, apo-Csd4, and Zn-Csd4, positive electron density was observed matching the proposed tripeptide cleavage product, *m*-DAP, based on comparisons between wild-type and *csd4* PG from *H. pylori* (Sycuro *et al.*, 2012). Because no exogenous *m*-DAP was included during the purification or crystallization of Csd4, affinity for *m*-DAP was sufficiently high for co-purification from *E. coli* lysate. The *m*-DAP forms direct interactions with Asn-93, Arg-94, His- 126, the Zn^2 ligand His-128, Thr-208, and Glu-222 (Figure 2.1, E and F). Asp-91 directly forms an H-bond to the Zn^2 ligand Gln-46 and to a water molecule that is H-bonded to both Gln-46 and the outer carboxylate oxygen of *m*-DAP. Three hydrophobic residues also line one side of the substrate-binding pocket (Trp-148, Ile-153, and Met-203) to interact with the alkyl portion of *m*-DAP. Additionally, a buried network of water molecules is present that form interactions between *m-DAP* and both Csd4 backbone and side chain residues.

To examine the interactions of Csd4 with substrate, Csd4 crystal soaking experiments were performed with Zn^2 and a tripeptide representing a portion of the PG substrate (Ac-L- Ala-D-Glu-*m*-DAP; refer to full manuscript for additional details (Chan *et al.*, 2015). The substrate was synthesized as a mixture of two stereoisomers: one with the (*R*)-stereocenter of *m*-DAP attached to γ-D-Glu and one with the (*S*)-stereocenter of *m*-DAP attached to γ-D-Glu. The crystal structure of tripeptide-Zn^2 -Csd4 complex (TriZn-Csd4) was solved to 1.75 Å (Table 2.1 and Table 2.2). No significant difference in the overall fold was observed when compared with Zn-Csd4 (r.m.s.d. 0.2 Å over all Cα atoms). Electron density at the metal binding site was weaker, and a zinc ion was modeled at 50% occupancy. As in Zn-Csd4, Gln-46, His-128, and Glu-49 are

coordinated to the zinc ion; additionally, a single solvent molecule with well defined electron density is observed coordinated to the zinc and is modeled as a water at full occupancy (Table 2.2 and Figure 2.4 A). This water is H-bonded to the backbone carbonyl group of Asp- 129 (2.8 Å) and the side chain of Glu-222 (2.9 and 3.1 Å, bidentate) and is situated 3.3 Å from C7 of the tripeptide scissile bond. Therefore, it is poised to be the catalytically essential water (Gomis-Rüth, 2008), although conformational changes at the active site may be required during the catalytic process. The second zinc coordinated water molecule in Zn-Csd4 appears to have been displaced by the bound tripeptide.

Table 2.2. Zinc ligand bond lengths in the crystal structures of Csd4

	Zn-Csd4	TriZn-Csd4	Q46H-Csd4
Gln46/His46	2.2	2.0	2.1
Glu49	2.3/2.3	2.4/2.6	2.1/2.5
His128	2.0	2.5	2.1
H_2O (catalytic)	2.3	2.9	-
H_2O (other)	2.2	-	-
Phosphate	-	-	1.9

The tripeptide substrate analog was modeled at full occupancy in TriZn-Csd4 (Figure 2.4). Under the crystal soaking conditions, tripeptide cleavage is sufficiently impaired that the substrate and not the product was observed. Only the naturally occurring stereoisomer (the (*S*)-stereocenter of *m*-DAP attached to γ-D-Glu) was observed in the active site, indicating that the enzyme selectively bound the preferred isomer from solution. The *m*-DAP moiety of the tripeptide overlays with the product structure, and the interactions with Csd4 are conserved (Figure 2.4 B). The tripeptide substrate extends outward past the binding pocket, and only the carbonyl O of D-Glu provides an additional significant interaction with Arg-86 (Figure 2.4 A

and Figure 2.5) Accordingly, the B-factors of the buried *m*-DAP moiety are low (20 $Å^2$) and increase toward the surface-exposed end of the tripeptide (Figure 2.4 *C*).

Plotting the degree of amino acid conservation among the homologs of Csd4 (see below) on the surface of the structure reveals highest conservation at the substrate-binding site (Figure 2.1 C). Although the majority of these conserved residues are part of the extensive zinc and substrate-binding network, additional conserved residues are present at the CP domain surface surrounding the active site pocket. This surface is composed primarily of loops, the largest of which consists of residues Tyr- 133–Trp-148 that are poised to interact with the polysaccharide backbone of the PG substrate.

***Gln-46 Is Required for Full Csd4 Activity*—**To examine the role of Gln-46 in the catalytic activity of Csd4, three active site variants were constructed (Q46H, Q46A, and Q46E) and assayed for *m*-DAP release from the tripeptide substrate. Wild-type Csd4 demonstrated highest activity in either sodium/potassium phosphate or Bis-tris buffer (Figure 2.3 A).

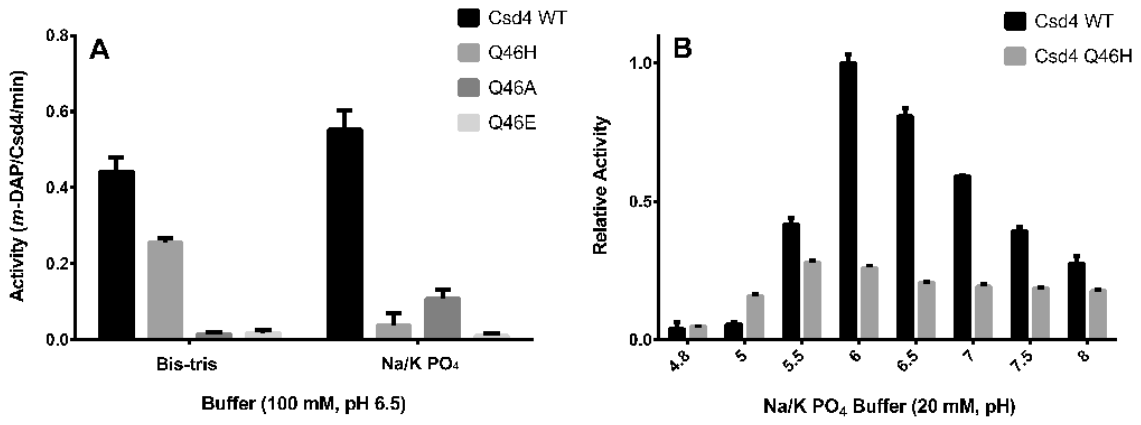

Figure 2.3. Wild-type Csd4 exhibits higher catalytic activity on the tripeptide substrate than its active site variants.

A. Csd4 activity was continuously monitored via the activity of *meso*-diaminopimelate dehydrogenase (*m*-DAP), which consumes the Csd4 product *m*-DAP to produce NADPH.

Buffer-based activity rate differences of Csd4 and its variants are shown. B. The pH-based activity differences between wild-type Csd4 and the Q46H variant was examined by examining the amount of product produced after 20 min. Mean values are shown with error bars representing the standard deviation based on at least three experimental replicates. The P-value for all variants is < 0.0005 as compared to wild-type in their respective buffers utilizing the t test in panel A. P-values between wild-type and Q46H are < 0.0006 for all pH values except pH 4.8 in panel B.

The peptidase activity of the Q46H variant was half that of wild-type in the Bis-tris-buffered system but was 10-fold reduced in the presence of phosphate. The complete loss of a zinc ligand (Q46A variant) retained 20% of wild-type activity in phosphate but no significant activity in Bis-tris buffer. Substitution of Gln with acidic Glu resulted in no significant activity and appeared to disrupt the active site. Wild-type Csd4 had highest activity at pH 6, whereas the Q46H optimal was at pH 5.5 but was much less pH-sensitive (Figure 2.3 B).

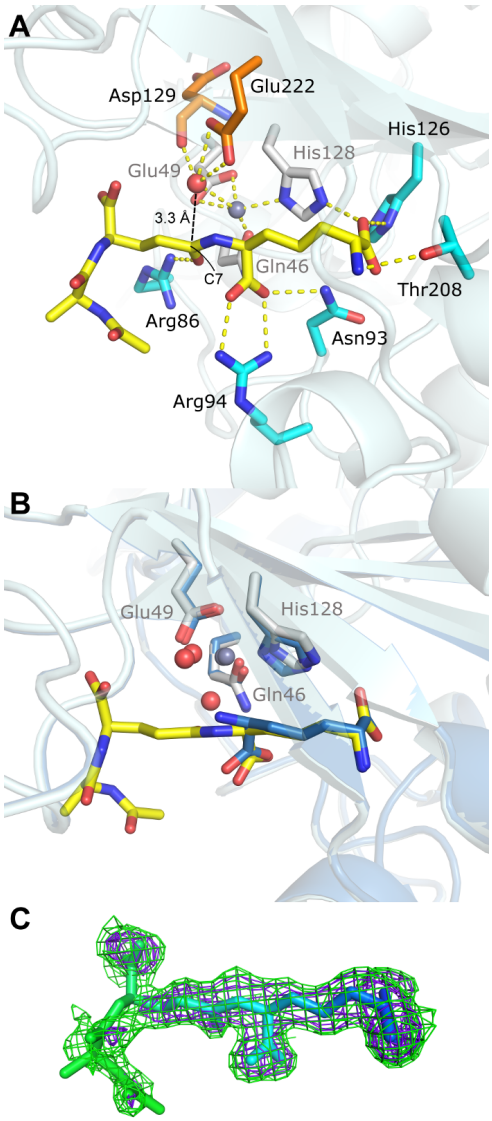

Figure 2.4. Tripeptide substrate binding site (PDB code 4WCN).

A. Key zinc and substrate interactions with Csd4 are shown. Zinc ligands are colored *white*; tripeptide ligands are colored *cyan*; the predicted catalytic water is *red* and its ligands are *orange*; zinc is *grey*; key interactions are shown as *dotted lines*. Hydrophobic residues and other waters are not shown. B. Structural alignment between substrate- and product-bound Csd4. View is of the active site in panel A with a 30° rotation about the x-axis. C. Tripeptide omit F_o-F_c difference map contoured at 2 σ (*green*) and 3 σ (*purple*) showing well-defined omit density for the deeply buried portions of the tripeptide substrate. Increased tripeptide flexibility in relation to the degree of surface exposure are indicated by a B-factor based coloring scheme (*blue*=~20 Å2 and *green*=~60 Å2).

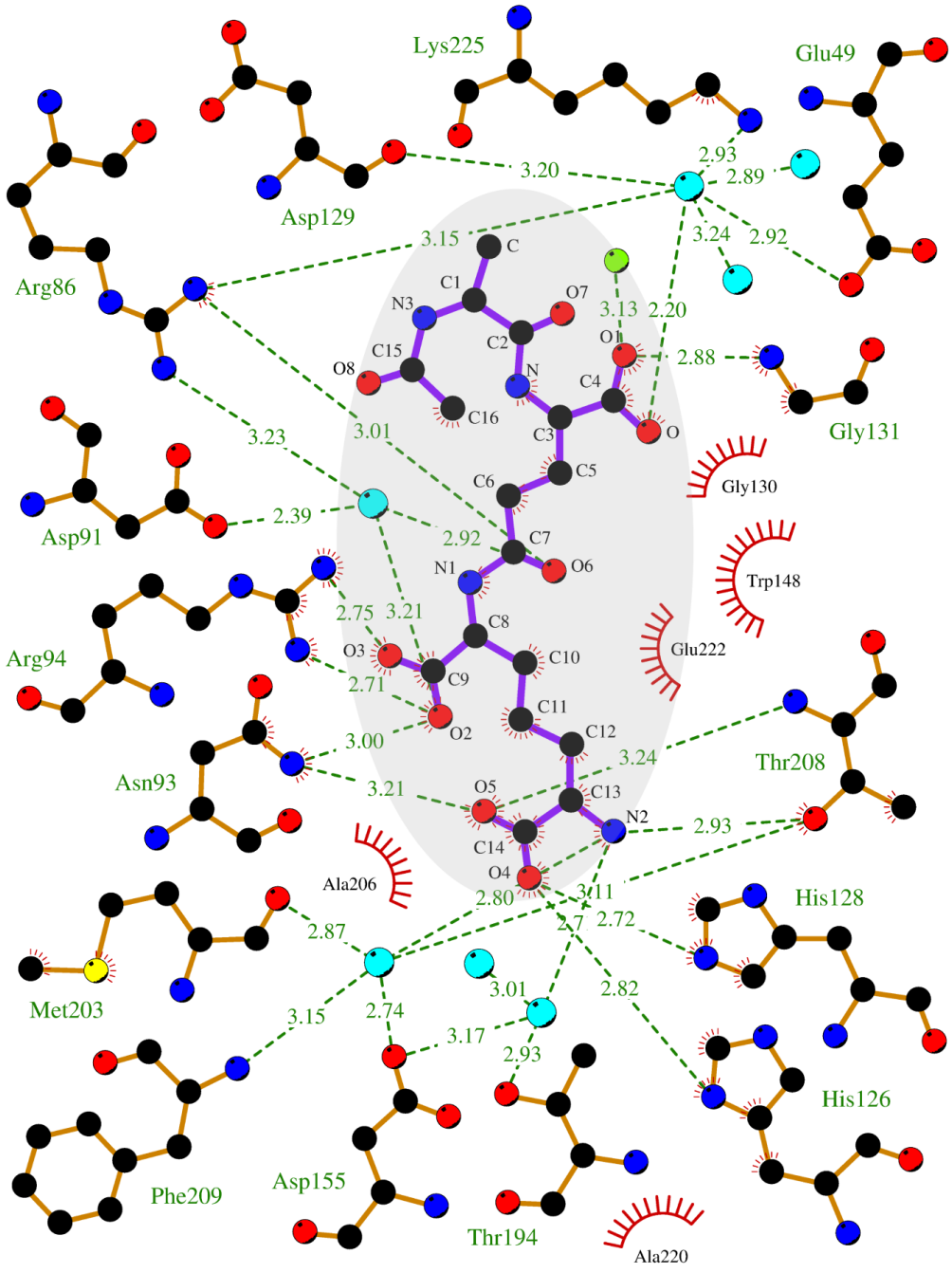

Figure 2.5. 2D tripeptide-Csd4 interaction map.

Tripeptide substrate is highlighted in *grey*; waters are colored *cyan*; iodide in *green*; *red bristled arcs* depict nearby hydrophobic interactions; zinc cofactor and its ligands are not shown. Drawn using LigPlot+ (Laskowski and Swindells, 2011).

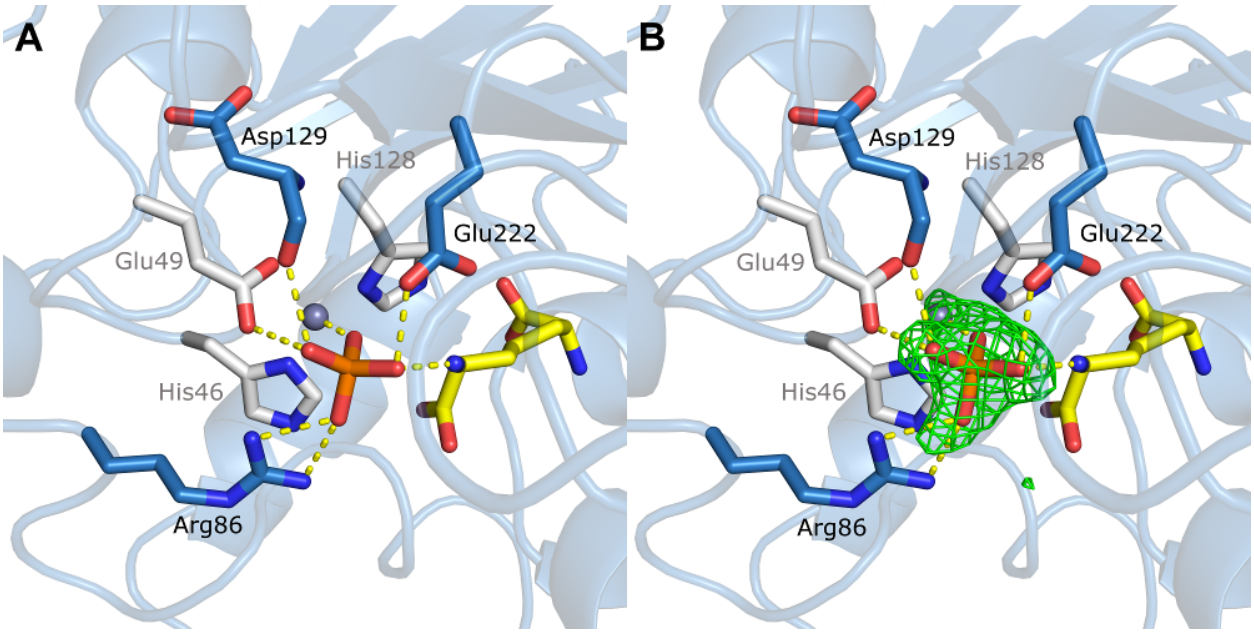

Figure 2.6. Active site of the Zn-bound Q46H variant.

A.Zinc ligands are colored white; *m*-DAP is in yellow; phosphate is in orange; other residues interacting with the phosphate are colored blue; phosphate interactions are shown as dotted lines. B. Omit Fo-Fc difference density map contoured at 3 σ showing density for a bound phosphate.

To explore the structural basis for altered activity by the Q46H variant, the crystal structure of zinc-bound Q46H was solved to 1.75 Å resolution (Figure 2.6 A). As expected, the His-His- Glu metal site bound Zn^2 with modest changes in ligand geometry (Table 2.2). A tetragonal-shaped density was observed near the zinc ion in a F_o - F_c map (Figure 2.6 B). Based on the presence of phosphate in the protein purification buffer, the density was modeled as a phosphate molecule at 90% occupancy. Structural alignment between TriZn-Csd4 and the Q46H variant revealed that the phosphate molecule occupies the space of the zinc-coordinated solvent molecule. Arg-86 and Glu-222 have rotated to form H-bonds with the phosphate group. Density for the product, *m*-DAP, is observed in the structure of Q46H and is situated in the same location as in the wild-type Csd4 structures.

To determine whether Gln-46 is required for normal helical cell shape, we generated strains of *H. pylori* expressing Q46H and Q46A variants fused to a C-terminal 3x-FLAG epitope integrated at the native locus. Both mutants exhibited non-helical cell morphology consistent with a *csd4* null phenotype of slightly curved or straight rods with occasional kinks or bends (Figure 2.7, A and B) (Sycuro *et al.*, 2012). Western blot analysis with anti-FLAG monoclonal antibodies indicated no detectable differences in protein expression between the wild-type and mutant variants of Csd4 (Figure 2.7 D).

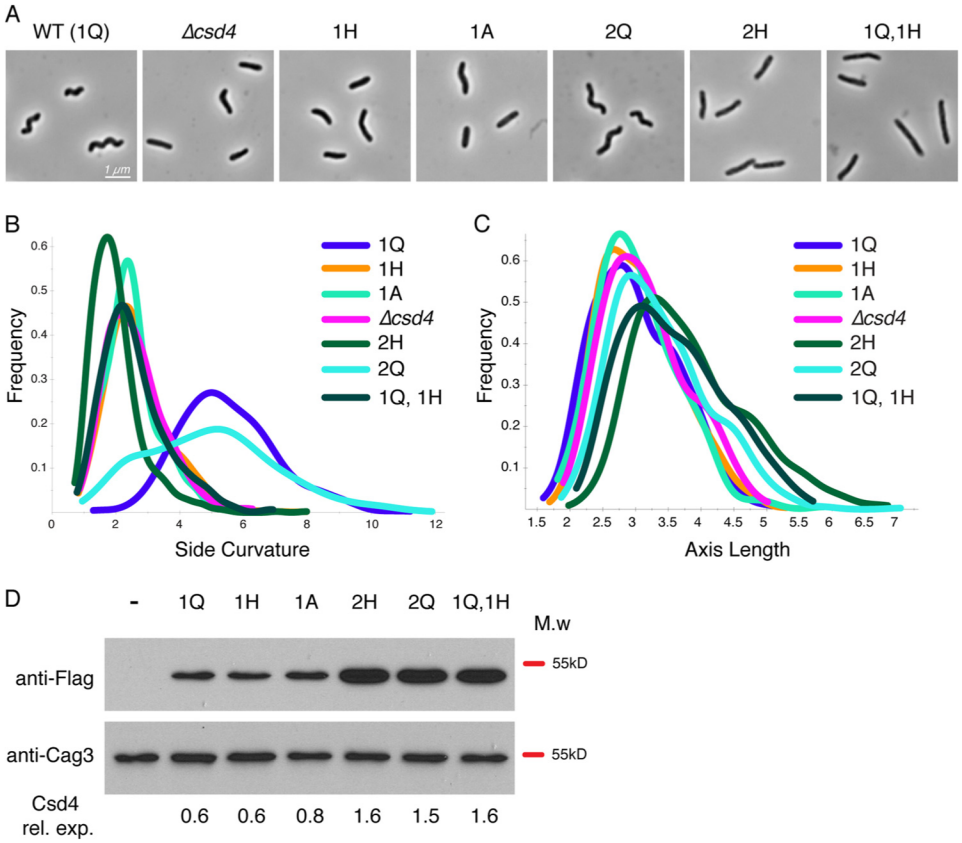

Figure 2.7. Complementation and overexpression analysis of Csd4 active site variants.

Strain labels indicate the copy number (1 or 2) and amino acid residue at position 46 (*Q*, *A*, or *H*) of *csd4 –3x-FLAG*. Strains with two copies have one copy at the native locus and the second copy at the *rdxA* locus. The strains used were (-) LSH100 WT (no 3x-FLAG tag), KBH54 (Δ*csd4*), KBH19 (*1Q*), KBH33 (*1H*), KBH42 (*1A*), KBH60 (*2H*), KBH65 (*2Q*), and KBH66 (*1Q*, *1H*). A-C. 1000X phase contrast images of wild-type (A) and *csd4* mutant *H. pylori*

(B and C) smooth histograms displaying population cell curvature (*x*-axis, *B*) and population axis length (*x*-axis, C) as a density function (*y*-axis). D, Western blot analysis of Csd4 (detected by anti-FLAG-M2 antibody, predicted molecular mass of 48 kDa) and Cag3 (periplasmic protein loading control, predicted molecular mass is 55 kDa). Equivalent amounts of cell extract based on optical density of the culture were loaded for each strain. ImageJ software was used for densitometry analysis of Csd4 variant expression relative to Cag3 and is indicated below each lane. *M.w*, molecular mass; *rel. exp.*, relative expression.

Because the Q46H mutant retained some *m*-DAP cleavage activity *in vitro*, we generated a merodiploid strain expressing a second copy of *csd4-Q46H-3x-FLAG* (KBH60) at the *rdxA* locus (used for *csd4* complementation in previous studies) (Sycuro *et al.*, 2012) to explore whether overexpression of the Q46H variant might rescue helical shape. Although we observed a 2.4-fold higher protein expression in the strain expressing two copies of *csd4Q46H* (Fig. 8*D*), we saw no restoration of helical morphology (Figure 2.7, A and B). In contrast, a strain expressing two copies of the Q allele (KBH65) supports normal morphology in most cells (Figure 2.7 A) but has an increased population of straight cells that have side curvature values less than 4 compared with the strain with a single copy of the Q allele (Figure 2.7 B), as was reported previously (Sycuro *et al.*, 2013). The perturbation of cell shape during overexpression was suggested to result from a requirement of precise asymmetric localization of Csd4 activity to generate proper helical curvature. In this model, loss of expression prevents the induction of curvature whereas extra protein expression at additional sites may break asymmetry and again lead to loss of curvature. We then created a strain expressing the wild-type *csd4* allele at the native locus and the *csd4Q46H* allele at *rdxA* (KBH66) and observed a dominant-negative effect of the *csd4Q46H* allele with a complete loss of helical cell morphology (Figure 2.7, A and B). This unexpected result may suggest that in cells Csd4 acts cooperatively in a complex. Previously we observed an

increase in cell length in a wild-type *csd4* merodiploid strain (Sycuro *et al.*, 2013). Interestingly, all strains containing two copies of *csd4* show increased cell length regardless of whether one or both copies contain Q46H (Figure 2.7 D). Taken together with the *in vitro* activity data, these results suggest that the Q46H variant is unable to generate helical cell morphology caused by a perturbation of enzyme activity, but enzymatic activity is not required for the effects of Csd4 on cell length (Figure 2.7 C).

***Csd4 as an Archetype for a New Family of CPs*—**A sequence analysis of Csd4 homologs was performed to identify conserved features. Bacterial homologs containing all three domains of Csd4 were identified based on a sequence similarity search; the sequences were aligned, and a phylogenetic tree was constructed (Figure 2.8 and Figure 2.9). Homologs were identified primarily within the δ and ε-Proteobacteria, Deferribacteres, and Aquificae (Frirdich *et al.*, 2012). Although primarily found in helical or curved rod-shaped organisms, homologs of Csd4 were identified in a number of rod-shaped bacteria, including those isolated from deep sea hydrothermal vents and coastal sediments. Homologs of Csd4 generally cluster based on bacterial class and by the presence of the Gln zinc ligand *versus* an equivalent His. However, *Helicobacter* homologs fall into two distinct branches. The first branch includes *H. pylori* and other ε-Proteobacteria (group Epsilon-Q*). All members of this branch contain a Gln at position equivalent to Gln-46 except for *Helicobacter hepaticus* and *Helicobacter cinaedi*, which contain a histidine but are still helical. Additionally, five of these *Helicobacter* species, including *H. hepaticus* and *H. cinaedi*, contain an extended C-terminal region of 200-300 residues as compared with Csd4. The second branch contains ε-Proteobacteria and have the His variation at position 46 (group Epsilon-H).

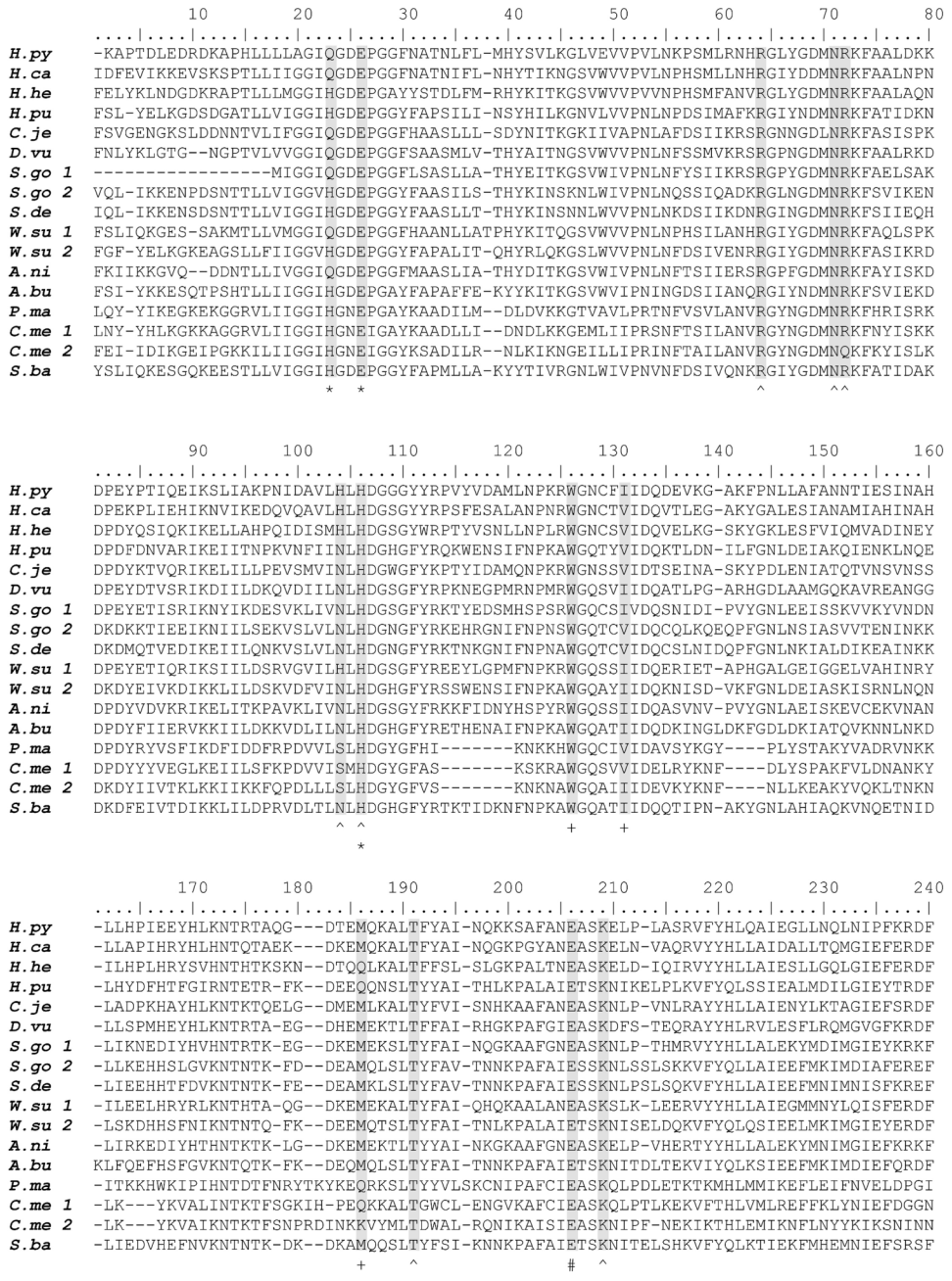

Figure 2.8. Conservation of zinc and substrate binding residues for select homologs of Csd4. *Helicobacter pylori* G27 (H.py) YP_002265985.1, *Helicobacter canis* (H.ca) WP_023929364.1, *Helicobacter hepaticus* ATCC 51449 (H.he) NP_860063.1, *Helicobacter pullorum* (H.pu) WP_005022665.1, *Campylobacter jejuni* 81-176 (C.je) YP_001001002.1, *Desulfovibrio vulgaris* str Miyazaki F (D.vu) YP_002436677.1, *Sulfurimonas gotlandica* GD1 (S.go 1: SMGD1_0946, WP_008338748.1; S.go 2: SMGD1_2299, WP_008339471.1), *Sulfurimonas denitrificans* DSM 1251 (S.de) YP_393892.1, *Wolinella succinogenes* DSM 1740

(W.su 1: WS0230, NP_906487.1; W.su 2: WS0783, NP_906997.1), *Arcobacter nitrofigilis* DSM 7299 (A.ni) YP_003657196.1, *Arcobacter butzleri* RM4018 (A.bu) YP_001489842.1, *Persephonella marina* EX-H1 (P.ma) YP_002731141.1, *Caminibacter mediatlanticus* TB-2 (C.me 1: CMTB2_05747, WP_007475257.1; C.me 2: CMTB2_06881, WP_007473913.1), *Sulfurospirillum barnesii* SES-3 (S.ba) YP_006404626.1, *Escherichia coli* str K-12 substr MG1655 MpaA (E.co) AAC74408.2, *Vibrio campbellii* ATCC BAA-1116 MpaA (V.ca) ABU74930.1. Only the region containing the Csd4 carboxypeptidase domain is shown. Highlighted are key amino acid ligands in the Csd4 crystal structures: zinc-binding ligands (*), substrate/product binding residues (^), hydrophobic binding pocket residues (+), catalytically important glutamate (#).

Whether the Csd4 homologs in those organisms play an active role in determining their shape is not currently known. The zinc ligands Glu-49 and His-128 are absolutely conserved among the identified Csd4 homologs. Although Gln-46 is required for full activity and helical shape of *H. pylori*, approximately one-quarter of the species identified have a histidine in the equivalent position. All of the bacteria isolated from deep-sea vents, whether they belong to the phylum Aquificae or are ε-proteobacteria, have such a histidine. Although His- 46-containing organisms include spiral-shaped species, the majority of deep-sea vent isolates appear to be rod-like; however, whether these species have classical rod morphology is unknown.

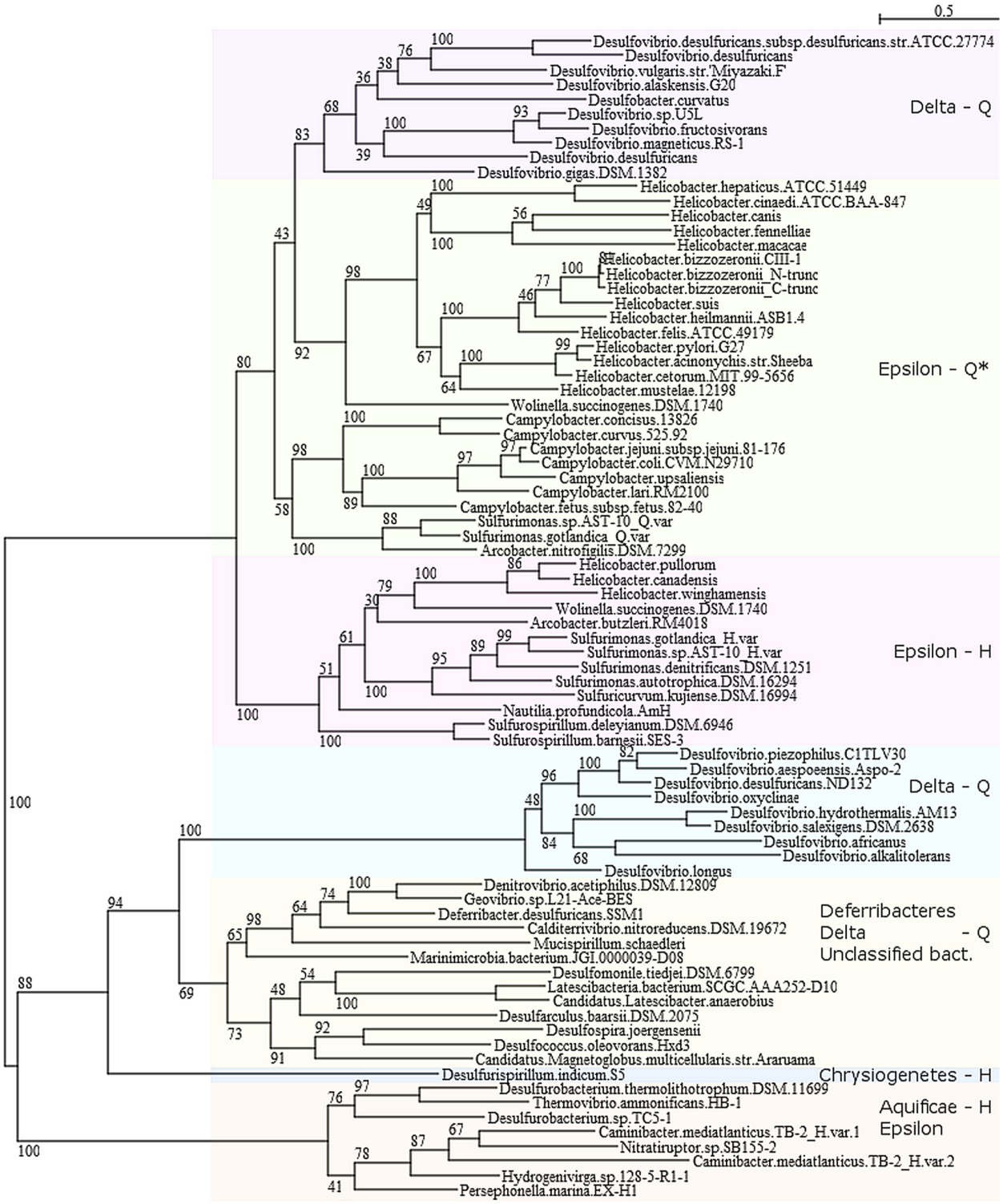

Figure 2.9. A bootstrapped tree of the Csd4 family of CPs.

Homologs of Csd4 were identified from the non-redundant database at the National Center for Biotechnology Information utilizing BLASTP and an E-value cutoff of 1×10^{-7}. Identical

protein sequences derived from different strains of the same species and proteins with short alignment coverage were removed. The sequences were aligned with Clustal Omega (Sievers *et al.*, 2011) and a bootstrapped tree (with 500 replicates, subtree pruning and regrafting and five random starts) was generated using PhyML (Guindon *et al.*, 2010) within Seaview (Gouy *et al.*, 2010). * *Helicobacter hepaticus* and *Helicobacter cinaedi* form their own sub-branch and contain a His at the equivalent position of Gln46.

2.5 DISCUSSION

Our product and substrate-bound crystal structures with an active site zinc gives further insight into cell wall substrate recognition and hydrolysis by enzymatically active Csd4 and expands our knowledge on the larger family of related funnelin CPs (Gomis-Rüth, 2008). In the Csd4 resting state, as represented by the Zn- Csd4 crystal structure, two water molecules are bound to the active site zinc. Substrate binding (*i.e.* TriZn-Csd4) displaces one of these two water molecules, leaving the other water positioned for nucleophilic attack of C7 of the substrate scissile amide bond. Glu-222 is the conserved glutamate positioned to act as the essential general base that abstracts a proton from the catalytic water to enhance nucleophilic attack on the amide carbonyl carbon (C7). Both zinc and Arg-86 are positioned to stabilize a negatively charged *gem*-diolate intermediate that is proposed in the generally accepted mechanism of CPs (Gomis-Rüth, 2008). In contrast to other funnelins, no Tyr residue is present to H-bond to the amide nitrogen. Protonation of the amide nitrogen by Glu-222 is associated with cleavage of the peptide bond. The PG dipeptide moiety appears to easily dissociate; however, the *m*-DAP remains bound and is likely displaced by the next PG tripeptide to repeat the cycle. Typically, funnelin family CPs hydrolyze the C-terminal residues of folded proteins in contrast to the isopeptide bond of the PG tripeptide. The unique structural features in Zn-Csd4 and TriZn-Csd4 may be a consequence

of interacting with PG as the substrate. Recently, the crystal structures of muramyltripeptide and *m*-DAP bound Csd4 have been solved with calcium in the active site (Kim *et al.*, 2014). As expected, the sugar moiety was not observable in the electron density (PDB code 4Q6N).

Most residues that bind the catalytic zinc or that interact with the substrate are highly conserved among the homologs of Csd4. Glu-222 is absolutely conserved in accordance with its proposed direct role in the Csd4 reaction mechanism. Moreover, mutation of Glu-222 in Csd4 alters *H. pylori* cell shape (Sycuro *et al.*, 2012). Further insight into the Csd4 reaction may be gained by monitoring kinetics of residue variants, substrate analogs, and inhibitors such as phosphate. The two notable exceptions to the strict conservation of active site residues are the zinc ligand Gln-46 and His-126. His-126 interacts with *m*-DAP and is often replaced by polar uncharged residues, which may fulfill the same role. In some Csd4 homologs, Gln-46 is substituted by His, as observed in most funnelin CPs. Although the substitution of Gln to His requires a single point mutation, this residue is required for function and is largely conserved within branches of the Csd4 family. Additionally, the specific spatial positioning of a glutamine at position 46 appears to be required, because glutamine is not found in place of the other histidine zinc ligand, His-128.

The binding of phosphate at the zinc site in the structure of the Q46H variant may be due to structural similarity to the tetrahedral *gem*-diolate intermediate of the catalytic cycle. The rearrangement of Arg-86 and Glu-222 to accommodate this phosphate molecule may reflect the role of these residues in stabilization of the intermediate formed during normal catalysis. The observed inhibition of the activity of the Q46H variant by phosphate together with phosphate bound in the variant crystal structure suggests that coordination of the zinc by Gln in Csd4 may serve to prevent phosphate inhibition. Notably, phosphate was observed bound to *Bacillus*

subtilis LdcB (Hoyland *et al.*, 2014) and *Streptococcus pneumoniae* DacB (Abdullah *et al.*, 2014), two LD-carboxypeptidases that remove the fourth amino acid from PG peptides and have a His-His-Asp zinc-binding motif. Although LD-CPs are also found in *H. pylori* (Csd6) (Sycuro *et al.*, 2013) and *C. jejuni* (Pgp2) (Frirdich *et al.*, 2014), neither are related to the Gram (-) LD-CPs, nor to DL-CPs such as Csd4.

Gln is required for Csd4 function yet is an uncommon zinc ligand. Zinc ligands play a key role in modulating the pK_a and nucleophilicity of the bound catalytic water and therefore the activity of the enzyme. Gln is a polar amino acid that has a similar size and chemical properties as histidine, yet substitution of Gln-46 by His results in the loss of both CP activity and *in vivo* function. Zinc binding sites in proteins are typically formed by His, Cys, Glu, and Asp residues (Sousa *et al.*, 2009). A previous analysis in the PDB database found 6200 zinc-containing sites (Andreini and Bertini, 2012). From this list (C. Andreini, personal communication), we have found 12 wild-type entries, representing 6 proteins that are observed with Gln as a zinc ligand, none of which are peptidoglycan peptidases (Table 2.3). One well characterized example is human glyoxalase I, an unrelated zinc enzyme with a zinc site composed of a His, two Glu, and a Gln, whereas glyoxalase I from other organisms have a second His ligand instead of Gln. A Q33E variant of human glyoxalase I showed significantly reduced activity *in vitro* (1.3% of wild-type) (Ridderström *et al.*, 1998). However, a Q33H variant of human glyoxalase I was not examined, and therefore whether there is a similar deleterious effect as the Q46H substitution in Csd4 is not clear.

Table 2.3. Characterized Gln-containing Zn proteins identified from the Protein Data Bank.

Protein	Organism	PDB-ID	Zn ligands	Class
Phosphomannose isomerase	*Candida albicans*	1pmi	Q-H--E-H-H_2O	Isomerase
Glyoxalase I	*human, mouse*	1bh5, 1fro, 1qip, 1qin, 2za0	H-E-Q-E-H_2O	Lyase
ING4	*human*	2k1j	Q-H (and 2 weak C interactions)	Gene Regulation/ Zinc-finger
L-Histidinol Dehydrogenase	*E. coli*	1kae	Q-H-D-H	Oxidoreductase
Glucose Dehydrogenase I	*Sulfolobus solfataricus (archaea)*	2cd9	C-H-E-Q	Oxidoreductase
5'-Nucleotidase	*E. coli*	1hp1, 1ush, 2ush	D-H-D-Q	Hydrolase

Typical to the members of the M14 family, the tripeptide binding pocket of Csd4 forms a characteristic cul-de-sac that determines substrate length and specificity. At the bottom of this pocket in Csd4 is Thr-208, an absolutely conserved residue that forms a H-bond to the terminal amine of the *m*-DAP moiety through the side chain O_γ (3.0 Å). The equivalent position is responsible for substrate specificity in other CPs. An example is subfamily M14A, which is further subdivided into A-type (*e.g.* carboxypeptidase A) and B-type (*e.g.* carboxypeptidase B) enzymes (Gomis-Rüth, 2008). In A-type CPs, Thr-208 is replaced by small hydrophobic residues to preferentially interact with aromatic or small aliphatic side chains, whereas B-type CPs, which prefer substrates terminating in basic amino acids, have a negatively charged residue at the position equivalent to 208. The presence of Thr-208 rather than an acidic residue like in B-type CPs in part explains the lack of detectable activity on hippuryl-Lys, a chromogenic substrate commonly used to assay the activity of B-type carboxypeptidases (data not shown). The Csd4 substrate binding pocket is designed to accommodate the stem tripeptide portion of PG but does

not have sufficient space for a fourth amino acid, explaining the lack of activity on a disaccharide tetrapeptide substrate analog reported previously (Sycuro *et al.*, 2012).

Csd4 shares similar substrate preferences to members of subfamily M14C, characterized by two bacterial CPs involved in the cleavage of murein-derived substrates: endopeptidase I from *Lysinibacillus sphaericus* and the *E. coli* murein peptide amidase (MpaA). Although both enzymes are CPs and have a canonical zinc-binding motif, neither has detectable sequence similarity to Csd4 by BLAST analyses (*E* values 0.1). Endopeptidase I cleaves *m*-DAP-D-Ala from the PG tetrapeptide (DL- endopeptidase activity) and also *m*-DAP from the tripeptide (DL-carboxypeptidase activity) (Garnier *et al.*, 1985). The CP domain of endopeptidase I is preceded by two tandem copies of a LysM domain predicted to bind PG. Conversely, MpaA is a single domain cytoplasmic CP involved in the catabolic usage of murein-derived peptides as a nutrient source. Unlike Csd4, MpaA cleaves murein tripeptides but has little to no activity on *N*-acetylmuramic acid tripeptides and tetrapeptides (Maqbool *et al.*, 2012). The structure of the *Vibrio harveyi* MpaA homolog (solved only in the apo form) shows overall homology to other CPs including Csd4 but contains an extra loop region over the substrate-binding site that is proposed to determine the preference for small, sugarless substrates. The absence of such a loop in Csd4 is consistent with its ability to interact with the large intact PG sacculus.

Csd4 is a structurally unique CP that hydrolyzes tripeptides to modify PG. We have shown that Csd4 contains an M14-type domain and has the requisite DL-carboxypeptidase activity to release *m*-DAP using a zinc site with a Gln ligand. Interestingly, the MEROPS database of CPs (Rawlings *et al.*, 2012) currently lists *Helicobacter* as a genus lacking any members of the M14 family of CPs and a BLAST search of the Csd4 sequence against this

database yields CP sequences with only weak significance scores (*E* values 1 10 ^{3}). In addition, Csd4 and its close homologs contain two additional C-terminal domains that are unlike those found in the other characterized M14 subfamilies. Therefore, we propose that Csd4 be defined as the prototypical member of a new subfamily tentatively named M14E.

Open questions include how Csd4 functions within the context of the bacterial cell. The C-terminal domains play a role in Csd4 stability within the bacterial cell but may have additional functions in coordinating Csd4 localization through interactions with other shape determining proteins or the PG itself. Although overexpression of the Q46H variant did not support helical cell shape generation, it did promote the increased cell length previously observed during overexpression of wild-type Csd4 (Sycuro *et al.*, 2013), suggesting additional functions for the protein that do not require enzymatic activity. *H. pylori* Csd3, an unrelated M23B family metallopeptidase, was shown to have DD-carboxypeptidase activity and can produce tetrapeptides from un-cross-linked pentapeptides (Bonis *et al.*, 2010). In turn, the tetrapeptide cleaving LD-carboxypeptidase Csd6 was recently shown to provide the tripeptide substrate for Csd4. Overexpression (as well as loss) of either Csd4 or Csd6 perturbs helical morphology; thus peptide hydrolysis by Csd4 to achieve a helical bacterial shape is likely to require precise spatial coordination within the PG layer (Sycuro *et al.*, 2013).

2.6 Methods and Supporting Information

***Bacterial Strains and Growth Conditions*—**Strains used in this work, as well as primers and plasmids used in strain construction are described in Table 2.4. *H. pylori* were grown in Brucella broth with 10% fetal bovine serum (Invitrogen) without antimicrobials or on horse blood agar plates with antimicrobials as described (Sycuro *et al.*, 2010). Bacteria were incubated at 37 °C

under micro-aerobic conditions in a trigas incubator (10% CO_2, 10% O_2, and 80% air). For resistance marker selection, horse blood plates were supplemented with 15 μg ml^{-1} chloramphenicol, 25 μg ml^{-1} kanamycin, or 60 mg ml^{-1} sucrose. For plasmid selection and maintenance in *Escherichia coli*, cultures were routinely grown in LB broth or agar at 37 °C supplemented with 100 μg ml^{-1} ampicillin.

Table 2.4. Plasmids, strains, and primers used in this study.

Plasmids	**Relevant or Descriptive Genotype**	**Reference or Source**
pLKS2	Csd4 in pBluescript II SK+	Sycuro 2012
pLC292	pRdxA with ampicillin resistance marker	Sycuro 2012
pKB20	Csd4 WT:3xFlag (in pBluescript II SK+)	This work
pKB27	Csd4 1-251:3xFlag (in pBluescript II SK+)	This work
pKB29	Csd4 1-343:3xFlag (in pBluescript II SK+)	This work
pKB37	Csd4 Q46A:3xFlag (in pBluescript II SK+)	This work
pKB38	Csd4 Q46H:3xFlag (in pBluescript II SK+)	This work
pKB44	*rdxA*::Csd4 WT:3xFlag (in pLC292)	This work
pKB46	*rdxA*::Csd4 Q46H:3xFlag (in pLC292)	This work
pAC-C4H	Csd4 Q46H (in pET15)	This work
pAC-C4E	Csd4 Q46E (in pET15)	This work
pAC-C4A	Csd4 Q46A (in pET15)	This work
Strains	**Relevant or Descriptive Genotype**	**Reference or Source**
LSH100	Wild-type *H. pylori* (G27 background)	Sycuro 2012
LSH18	*csd4::catsacB*	Sycuro 2012
LSH108	*rdxA::kansacB*	Sycuro 2010
LSH122	*csd4::cat*	Sycuro 2012
TRH1	Csd4 4xFlag (at native locus)	This work
KBH19	Csd4 WT:3x-FLAG (at native locus)	This work
KBH33	Csd4 Q46H:3x-FLAG (at native locus)	This work
KBH35	Csd4 1-343:3x-FLAG (at native locus)	This work

KBH37	Csd4 1-251:3x-FLAG (at native locus)	This work
KBH47	Csd4 Q46H:3x-FLAG, *rdxA::kansacB*	This work
KBH54	*csd4::cat, rdxA::kansacB*	This work
KBH60	Csd4 Q46H:3xFlag, *rdxA*::Csd4 Q46H:3xFLAG	This work
KBH64	Csd4 WT:3xFlag, *rdxA::kansacB*	This work
KBH65	Csd4 WT:3xFlag, *rdxA*::Csd4 WT:3xFLAG	This work
KBH66	Csd4 WT:3xFlag, *rdxA*::Csd4 Q46H:3xFLAG	This work

Primers	**Description**	**Sequence**
53	Csd4FLAG_UP_F	cgggccccccctcgaatgatagaagcttgcaaagcgtctt
54	Csd4FLAG_UP_R	acttttcaatctatatcttacttgtcg
55	Csd4FLAG_DN_F	tatagattgaaaagtgaatgggtgaatcttgccaaaagat
56	Csd4FLAG_DN_R	cgggctgcaggaattaacatgcatagctccatcaggtttca
94	Csd4-Q46A-SDM	tgttgcttttagcagggattgcaggcgatgagcctgggggt
95	Csd4-Q46H-SDM	ttgcttttagcagggattcatggcgatgagcctgggggtt
85	Csd4-3xF_UP_KpnI_F	gaattgggtaccgggccccc
86	Csd4-1-251_3xF_DNR	ataatccttatcgtcatcgtctttataatcaaaagggatattgagttgattgagta a
88	Csd4-1-343_3xF_DNR	ataatccttatcgtcatcgtctttataatcaggatcaaactctatgtaaaaaggc
89	Csd4-3xF_DN_F	gattataaagacgatgacgataaggattat
90	Csd4-3xFDNSacI_R	aagctggagctccaccgcggt
120	Csd4-3xFl-BamHI	aggagggatccattggttttaaaatcgcttgtttcaac
121	Csd4-3xFl-EcoRI	ctcctgaattcttctaatgtatggcatgactatcctt
A1	Csd4-Q46H – For	gttgcttttagcagggattcatggcgatgagcc
A2	Csd4-Q46H – Rev	ggctcatcgccatgaatccctgctaaaagcaac
A3	Csd4-Q46E – For	ttgttgcttttagcagggattgaaggcgatgagc
A4	Csd4-Q46E – Rev	gctcatcgccttcaatccctgctaaaagcaacaa
A5	Csd4-Q46A – For	aggctcatcgcctgcaatccctgctaaaagcaacaaatgg
A6	Csd4-Q46A – Rev	ccatttgttgcttttagcagggattgcaggcgatgagcct

Construction of H. pylori Strains Expressing Tagged and Mutant Csd4—Plasmids containing wild-type and mutant derivatives of *csd4* fused to the 3x-FLAG epitope were used for

complementation and creation of merodiploid strains and are described in Table 2.4. To generate pKB20, the *csd4* gene and the *3x-FLAG* epitope were individually amplified by PCR (Phusion, NEB) using TRH1 (*csd4:4xFLAG*) genomic DNA as template, with oligonucleotides 53/54 (csd4) and 55/56 (3x-FLAG), which contained XhoI and EcoRI sequences appended. Strain TRH1 was generated using PCR SOEing (Horton, 1995) to add the 3x-FLAG sequences and a stop codon right before the endogenous *csd4* stop codon, which was then used to replace *csd4*::*catsacB* in LSH18 using counterselection (Copass *et al.*, 1997; Pinto-Santini and Salama, 2009; Sycuro *et al.*, 2012).The vector pLKS2 (Sycuro *et al.*, 2012) was digested with XhoI and EcoRI (NEB) and gel purified (Qiagen). The vector backbone and purified PCR fragments were ligated together using the HD In-Fusion cloning kit (Clontech). Point mutant derivatives of *csd4* were generated using a modified quick change procedure (Stratagene) with the pKB20 master vector as template and oligonucleotides 94 (Q46A) and 95(Q46H). To generate pKB27 and pKB29, the *csd4* truncation alleles and the 3x-FLAG epitope were individually amplified using pKB20 as template with primers 85/86 ($Csd4^{1–251}$), 85/87 ($Csd4^{1–343}$), and 89/90 (3x-FLAG). The *csd4* alleles and *3x-FLAG* were gel-purified and stitched together by PCR SOEing (Horton, 1995) using primers 85/90, which contained KpnI and SacI sequences appended. The vector pKB20 and the stitched PCR products were digested with KpnI and SacI and gel-purified. The vector backbone and PCR fragments were ligated together using standard T4 ligase-based methodology and verified by Sanger sequencing of candidate clones (Fred Hutchinson Cancer Research Center Genomics Shared Resource). To generate pKB44 and pKB46 for integration of mutant alleles at the *rdxA* locus, the *csd4–3x-FLAG* alleles were PCR-amplified from KBH33 (Csd4 Q46H:3x-FLAG) and KBH42 (Csd4 Q46A:3x-FLAG) genomic DNA using primers 120/121 with BamHI and EcoRI sequences appended, digested with BamHI and EcoRI, and gel

purified. The vector pLC292 (Terry *et al.*, 2005) was digested with BamHI and EcoRI (NEB) and gel purified (Qiagen). The vector backbone and purified PCR fragments were ligated together using standard T4 ligase-based methodology and verified by Sanger sequencing of candidate clones (Fred Hutchinson Cancer Research Center Genomics Shared Resource). Recipient *H. pylori* bacteria containing a *csd4*::*catsacB* (LSH18) or *rdxA*::*kansacB* (LSH108) cassette (LSH18) were transformed with 2 - 4 g of the appropriate plasmid DNA (Qiagen) containing DNA sequences homologous to regions flanking the *rdxA*::*kansacB* or *csd4*::*catsacB* cassette. Genomic DNA was isolated from candidate clones and verified by Sanger sequencing of PCR amplified *csd4* (Fred Hutchinson Cancer Research Center Genomics Shared Resource).

***H. pylori Morphological Analyses*—**Phase contrast micros- copy was performed as described (Sycuro *et al.*, 2010), and resulting images were thresholded using the ImageJ software package. Quantitative analysis of threshold images of bacteria (300-350 cells/strain) to measure side curvature and central axis length was performed with the CellTool software package as described. Side curvature is the reciprocal of the radius of a circle tangent to a curve at any point; as such, a straight line has zero curvature, whereas a point on a bent line has a curvature proportional to the tightness of the bend. Cell length was estimated using the central axis length calculated by CellTool. Statistical comparison of cell shape distributions were performed using a CellTool module that calculates a bootstrap distribution of Kolmogorov-Smirnov statistics, as described (Sycuro *et al.*, 2012).

Immunoblotting*—H. pylori* whole cell extracts were prepared by harvesting log phase grown bacteria by centrifugation for 2 min at maximum speed in a microcentrifuge and resuspending in 2x SDS-PAGE sample buffer at 10.0 optical density (600 nm) per ml and boiled for 10 min. Proteins were separated by 12.5% SDS-PAGE and transferred onto PVDF membranes using a

semidry transfer system (I-blot; Invitrogen) according to the manufacturer's instructions. Membranes were blocked for 1 h at room temperature or overnight at 4 °C with 5% nonfat milk-Tris-buffered saline with Tween 20 (TBS-T; 0.5 M Tris, 1.5 M NaCl, pH 7.6, plus 0.05% Tween 20), followed by incubation overnight at 4 °C with primary antibody at 1:2500 dilution for anti-FLAG M2 (Sigma) or 1:20,000 dilution for anti-Cag3, in TBS-T (Sycuro *et al.*, 2013). Four 10-min washes with TBS-T were followed by a 1-h incubation at room temperature with appropriate horse- radish peroxidase-conjugated anti-immunoglobulin G (Santa Cruz Biotechnology) antibody at 1:20,000 dilution in TBS-T (anti-rabbit for Cag3 blots, anti-mouse for 3x-FLAG blots). After four more TBS-T washes, antibody detection was performed with ECL Plus immunoblotting (Cag3) or Millipore Immobilon (3x-FLAG) detection kits, following the manufacturer's protocol (GE Healthcare).

2.7 BIBLIOGRAPHY

Abdullah, M.R., Gutiérrez-Fernández, J., Pribyl, T., Gisch, N., Saleh, M., Rohde, M., *et al.* (2014) Structure of the pneumococcal l,d-carboxypeptidase DacB and pathophysiological effects of disabled cell wall hydrolases DacA and DacB. *Molecular Microbiology* **93**: 1183–1206.

Andreini, C., and Bertini, I. (2012) A bioinformatics view of zinc enzymes. *Journal of Inorganic Biochemistry* **111**: 150–156.

Bonis, M., Ecobichon, C., Guadagnini, S., Prevost, M.C., and Boneca, I.G. (2010) A M23B family metallopeptidase of *Helicobacter pylori* required for cell shape, pole formation and virulence. *Molecular Microbiology* **78**: 809–819.

Chan, A.C.K., Blair, K.M., Liu, Y., Frirdich, E., Gaynor, E.C., Tanner, M.E., *et al.* (2015) Helical shape of *Helicobacter pylori* requires an atypical glutamine as a zinc ligand in the carboxypeptidase Csd4. *J Biol Chem* **290**: 3622–3638.

Copass, M., Grandi, G., and Rappuoli, R. (1997) Introduction of unmarked mutations in the

Helicobacter pylori vacA gene with a sucrose sensitivity marker. *Infect Immun* **65**: 1949–1952.

Frey, P.A., and Hegeman, A.D. (2007) *Enzymatic Reaction Mechanisms*. Oxford University Press, USA.

Frirdich, E., and Gaynor, E.C. (2013) Peptidoglycan hydrolases, bacterial shape, and pathogenesis. *Current Opinion in Microbiology* **16**: 767–778.

Frirdich, E., Biboy, J., Adams, C., Lee, J., Ellermeier, J., Gielda, L.D., *et al.* (2012) Peptidoglycan-modifying enzyme Pgp1 is required for helical cell shape and pathogenicity traits in *Campylobacter jejuni*. *PLoS Pathog* **8**: e1002602.

Frirdich, E., Vermeulen, J., Biboy, J., Soares, F., Taveirne, M.E., Johnson, J.G., *et al.* (2014) Peptidoglycan LD-carboxypeptidase Pgp2 influences *Campylobacter jejuni* helical cell shape and pathogenic properties and provides the substrate for the DL-carboxypeptidase Pgp1. *J Biol Chem* **289**: 8007–8018.

Garnier, M., Vacheron, M.J., Guinand, M., and Michel, G. (1985) Purification and partial characterization of the extracellular gamma-D-glutamyl-(L)*meso*-diaminopimelate endopeptidase I, from *Bacillus sphaericus* NCTC 9602. *Eur J Biochem* **148**: 539–543.

Gomis-Rüth, F.X. (2008) Structure and mechanism of metallocarboxypeptidases. *Crit Rev Biochem Mol Biol* **43**: 319–345.

Gouy, M., Guindon, S., and Gascuel, O. (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**: 221–224.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**: 307–321.

Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Research* **38**: W545–9.

Horton, R.M. (1995) PCR-mediated recombination and mutagenesis. SOEing together tailor-made genes. *Mol Biotechnol* **3**: 93–99.

Hoyland, C.N., Aldridge, C., Cleverley, R.M., Duchêne, M.-C., Minasov, G., Onopriyenko, O., *et al.* (2014) Structure of the LdcB LD-carboxypeptidase reveals the molecular basis of peptidoglycan recognition. *Structure* **22**: 949–960.

Kim, H.S., Kim, J., Im, H.N., An, D.R., Lee, M., Hesek, D., *et al.* (2014) Structural basis for the recognition of muramyltripeptide by *Helicobacter pylori* Csd4, a D,L-carboxypeptidase controlling the helical cell shape. *Acta Crystallogr Sect D Biol Crystallogr* **70**: 2800–2812.

Laskowski, R.A., and Swindells, M.B. (2011) LigPlot+: Multiple ligand–protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**: 2778–2786.

Lipscomb, W.N., and Sträter, N. (1996) Recent advances in zinc enzymology. *Chem Rev* **96**: 2375–2434.

Maqbool, A., Hervé, M., Mengin-Lecreulx, D., Wilkinson, A.J., and Thomas, G.H. (2012) MpaA is a murein-tripeptide-specific zinc carboxypeptidase that functions as part of a catabolic pathway for peptidoglycan-derived peptides in γ-proteobacteria. *Biochem J* **448**: 329–341.

Pinto-Santini, D.M., and Salama, N.R. (2009) Cag3 is a novel essential component of the *Helicobacter pylori* Cag type IV secretion system outer membrane subcomplex. *J Bacteriol* **191**: 7343–7352.

Rawlings, N.D., Barrett, A.J., and Bateman, A. (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* **40**: D343–50.

Ridderström, M., Cameron, A.D., Jones, T.A., and Mannervik, B. (1998) Involvement of an active-site Zn2+ ligand in the catalytic mechanism of human glyoxalase I. *Journal of Biological Chemistry* **273**: 21623–21628.

Salama, N.R., Hartung, M.L., and Müller, A. (2013) Life in the human stomach: Persistence strategies of the bacterial pathogen *Helicobacter pylori*. *Nat Rev Micro* **11**: 385–399.

Schleifer, K.H., and Kandler, O. (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* **36**: 407–477.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539.

Sousa, S.F., Lopes, A.B., Fernandes, P.A., and Ramos, M.J. (2009) The Zinc proteome: A tale of stability and functionality. *Dalton Trans* **2**: 7946–7956.

Sycuro, L.K., Pincus, Z., Gutierrez, K.D., Biboy, J., Stern, C.A., Vollmer, W., and Salama, N.R. (2010) Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori*'s helical shape and stomach colonization. *Cell* **141**: 822–833.

Sycuro, L.K., Rule, C.S., Petersen, T.W., Wyckoff, T.J., Sessler, T., Nagarkar, D.B., *et al.* (2013) Flow cytometry-based enrichment for cell shape mutants identifies multiple genes that influence *Helicobacter pylori* morphology. *Molecular Microbiology* **90**: 869–883.

Sycuro, L.K., Wyckoff, T.J., Biboy, J., Born, P., Pincus, Z., Vollmer, W., and Salama, N.R. (2012) Multiple peptidoglycan modification networks modulate *Helicobacter pylori's* cell shape, motility, and colonization potential. *PLoS Pathog* **8**: e1002603.

Terry, K., Williams, S.M., Connolly, L., and Ottemann, K.M. (2005) Chemotaxis plays multiple roles during *Helicobacter pylori* animal infection. *Infect Immun* **73**: 803–811.

Typas, A., Banzhaf, M., Gross, C.A., and Vollmer, W. (2012) From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Micro* **10**: 123–136.

Vollmer, W., Blanot, D., and de Pedro, M.A. (2008) Peptidoglycan structure and architecture. *FEMS Microbiology Reviews* **32**: 149–167.

Wyckoff, T.J., Taylor, J.A., and Salama, N.R. (2012) Beyond growth: Novel functions for bacterial cell wall hydrolases. *Trends in Microbiology* 1–8.

Young, K.D. (2006) The selective value of bacterial shape. *Microbiol Mol Biol Rev* **70**: 660–703.

Chapter 3. *HELICOBACTER PYLORI* CSD5 LINKS A CELL SHAPE PROMOTING PROTEIN COMPLEX TO THE CELL WALL AND ATP SYNTHASE TO PROMOTE HELICAL SHAPE

3.1 PREFACE

This chapter is presented as a modified version of a manuscript currently in revision with Molecular Microbiology:

> Blair, KM., Mears KS., Taylor JA., Fero J., Jones LA., Gafken PR., Whitney JC., and Salama NR. 2018 "*Helicobacter pylori* Csd5 links a cell shape promoting complex to the cell wall and ATP synthase." *In Revision*

3.2 ABSTRACT

Chronic infection with *Helicobacter pylori* can lead to the development of gastric ulcers and stomach cancers. The helical cell shape of *H. pylori* promotes stomach colonization. Screens for loss of helical shape have identified several periplasmic peptidoglycan (PG) hydrolases and non-enzymatic putative scaffolding proteins, including Csd5. Both over and under expression of the PG hydrolases perturb helical shape, but the mechanism used to coordinate and localize these activities is not known. Using immunoprecipitation and mass spectrometry we identified Csd5 interactions with cytosolic proteins CcmA, a bactofilin required for helical shape, and MurF, a PG precursor synthase, as well as the inner membrane spanning ATP synthase. A combination of Csd5 domain deletions and point mutations revealed that N-terminal and transmembrane

domains promote MurF, CcmA, and ATP synthase interactions while the C-terminal SH3 domain mediates PG binding. We conclude that Csd5 promotes helical shape as part of a membrane associated, multi-protein shape promoting complex that links the cell wall to PG precursor synthesis enzymes, ATP synthase, and the bacterial cytoskeleton.

3.3 INTRODUCTION

H. pylori infection accounts for a significant global cancer burden as the primary cause of stomach cancer, the third leading cause of cancer deaths worldwide (Ferlay *et al.*, 2015). The helical cell shape of *H. pylori* is required for efficient stomach colonization (Sycuro *et al.*, 2010). Thus, understanding the mechanisms by which *H. pylori* achieves helical shape may inform novel therapeutic strategies.

The shape of most bacteria is determined by the peptidoglycan (PG) cell wall, a single macromolecule encasing the cell that determines cell shape and provides protection from osmolysis (Höltje, 1998; Typas *et al.*, 2012). The PG cell wall consists of a meshwork of glycan chains composed of alternating N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) cross-linked by short peptides attached to NAM (Vollmer and Bertsche, 2008). For bacterial cells to grow without compromising the structural integrity of the cell wall, they must coordinate the action of PG hydrolases that cleave PG with the action of PG synthases that insert newly synthesized PG precursors into the growing cell wall. PG biosynthesis begins in the cytoplasm where precursors consisting of undecaprenol (UDP) linked NAG-NAM-pentapeptides are assembled onto a lipid carrier through a series of cytoplasmic and membrane associated steps. The final lipid-linked precursor (lipidII) is then flipped into the periplasm where PG synthases assimilate the precursors into the existing wall by transglycosylation of the disaccharide (Sham

et al., 2014; Scheffers and Tol, 2015). Nascent PG strands are then cross-linked to existing PG by transpeptidation to uncross-linked PG peptides.

Genetic screens have identified several proteins that when deleted give rise to distinctly non-helical cell shapes. The majority of these proteins function as PG hydrolases that either break crosslinks (Csd1, Csd3) (Sycuro *et al.*, 2010; Bonis *et al.*, 2010) or trim uncross-linked peptide stems (Csd6, Csd4) (Sycuro *et al.*, 2012; Sycuro *et al.*, 2013), giving rise to either curved-rod or straight-rod shaped cells, respectively. Two non-enzymatic proteins, CcmA and Csd5, were also identified in these screens. *H. pylori* CcmA has homology to bactofilins, bacterial cytoskeletal elements capable of self-assembly without a cofactor that are widely distributed in eubacteria (Kühn *et al.*, 2010). In *Caulobacter crescentus*, the bactofilins BacA and BacB form membrane associated cytoplasmic filaments that associate with the inner membrane to direct stalk PG synthesis (Kühn *et al.*, 2010). *ccmA* mutants have altered global PG profiles similar to *csd1* and *csd3* suggesting it may regulate one or both of these enzymes. Csd5 is unique to *H. pylori* and *H. acinonychous* (Eppinger *et al.*, 2006). Deletion of *csd5* gives rise to straight cells yet global PG content analysis indicates no significant alterations in PG composition compared to wild-type (Sycuro *et al.*, 2012). Thus, the mechanism by which Csd5 promotes helical shape is unclear.

Here we performed a structure-function analysis of *H. pylori* Csd5 to gain a mechanistic understanding of how this protein promotes helical cell shape. We show that N-terminal cytoplasmic (NT) and transmembrane (TM) domains, plus a C-terminal SH3 domain are each required for helical shape. Csd5 interacts directly with peptidoglycan via its C-terminal SH3 domain and the N-terminal domains promote interactions with CcmA, the PG precursor synthesis enzyme MurF, and F_1F_OATP synthase. This work expands our understanding of the

biochemical activities required to support helical shape generation and provides new insights as to how these activities may be spatially organized.

3.4 RESULTS

Csd5 has conserved N and C terminal domains that each contribute to helical shape.

Iterative searches using Jackhmmer (Finn *et al.*, 2015) with HPG27_1195 (strain G27) and jhp1171 (strain J99) protein sequences were used to identify and predict Csd5 functional domains, sequence motifs, and protein topology (Figure 3.1 A). We mapped predicted secondary structure features derived from these queries onto a multiple sequence alignment of diverse Csd5 variants (Figure 3.2) from a hand-curated list of fully sequenced *H. pylori* genomes downloaded from the PATRIC database (Wattam *et al.*, 2017). This analysis revealed distinct regions of high sequence conservation in the N and C-terminal regions of Csd5. The less conserved middle domain is predicted to be disordered, and contains within it a repetitive, glutamine rich coiled-coil (CC) motif with variable length across diverse strains of *H. pylori* (Figure 3.2). The N-terminal region has a predicted transmembrane helix that may anchor Csd5 to the inner membrane while the C-terminus contains a bacterial SH3 domain (Kamitori and Yoshida, 2015). Bacterial SH3 domains can mediate protein-protein (Leon *et al.*, 2010; Xu *et al.*, 2015), protein-PG (Baba and Schneewind, 1996; Lu *et al.*, 2006; Xu *et al.*, 2009; Rolain *et al.*, 2013), protein-metal (Pohl *et al.*, 1999), or protein-nucleic acid (Wilkinson *et al.*, 2016) interactions.

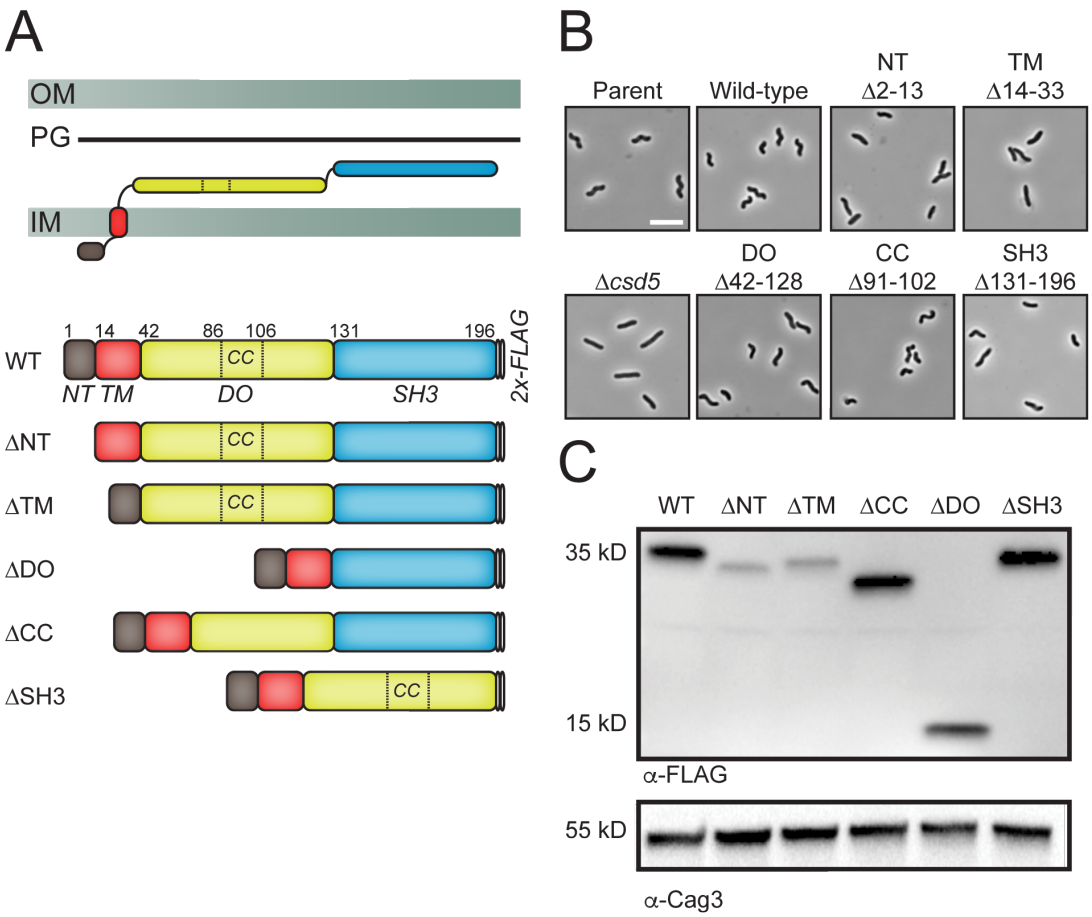

Figure 3.1. The N and C terminal domains of Csd5 are required for helical cell shape.

A. Schematic of the Gram-negative cell envelope and Csd5 predicted topology (top) and Csd5-2xFLAG deletion variants expressed at the native locus (bottom). Numbers indicate amino acid residues of predicted domain boundaries for Csd5. B. Phase contrast micrographs of parent (LSH100), Δ*csd5* (LSH36), Csd5-2xFLAG (WT, KBH127), ΔNT (KBH128), ΔTM (KBH129), ΔCC (KBH131), ΔDO (KBH132), ΔSH3 (KBH133). Scale bar = 5 µm. C. Western blotting of whole cell extracts probed with α-FLAG M2 to detect Csd5 variants and a polyclonal antibody to Cag3 as a periplasmic loading control. WT (wild-type), IM (Inner membrane), PG (Peptidoglycan), OM (Outer membrane), CC (Coiled-coil), NT (N-terminal) TM (Transmembrane), DO (Disordered).

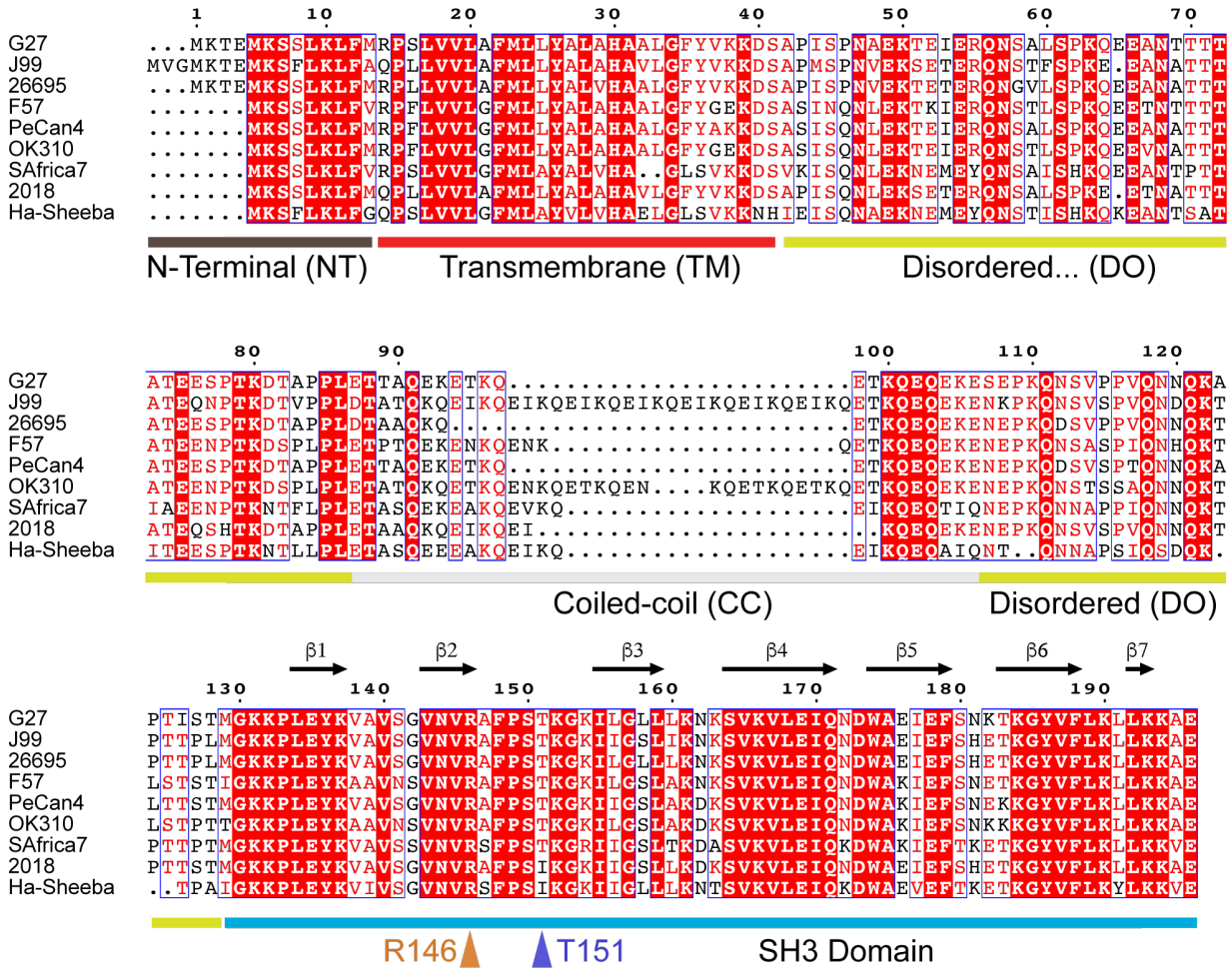

Figure 3.2. The N and C-terminal domains of Csd5 are conserved among geographically diverse strains of *H. pylori.*

H. pylori Csd5 sequences were aligned using Clustal omega (Sievers *et al.*, 2011) and ESPRIT 3.0 (Robert and Gouet, 2014) to display amino acid conservation using the default amino acid equivalency global cutoff of 0.7. Red boxes and white letters, strict identity; red letters, similar within amino acid groups; blue frame, similar across all amino acid groups. Colored bars below sequences indicate Csd5 domain boundaries (Figure 3.1). Arrows (beta sheets) indicate secondary structure of the SH3 domain based on our homology model.

To assess the contribution of each of these domains or motifs to the overall function of Csd5, we generated a series of deletion variants, each containing a C-terminal 2xFLAG tag integrated in single-copy at the native locus (Figure 3.1). Deletion of the N-terminal tail (ΔNT),

the transmembrane helix (ΔTM), or the SH3 domain resulted in rod shaped bacteria, similar to *csd5* null bacteria (Figure 3.1 B and Figure 3.3).

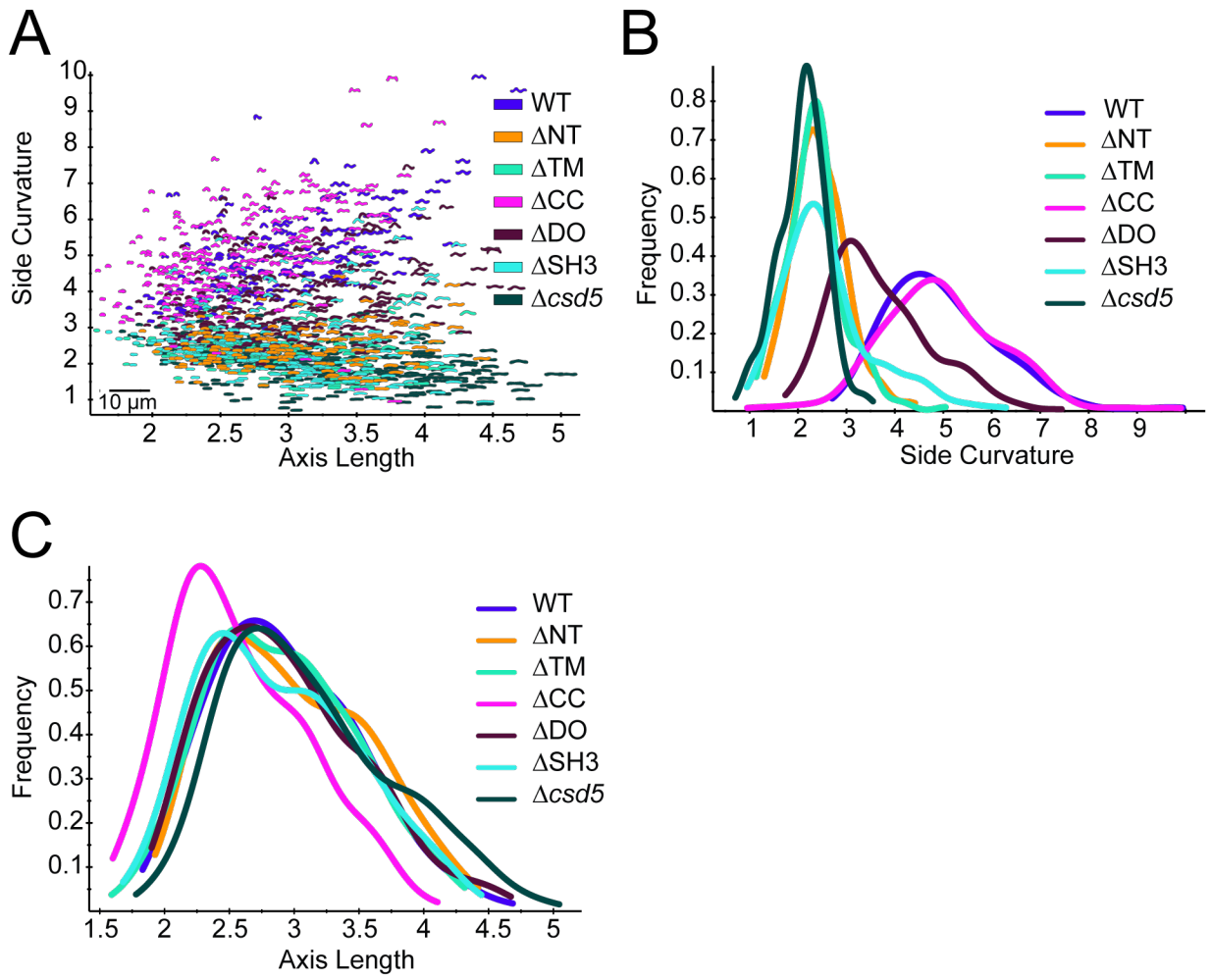

Figure 3.3. Quantitative cell shape analysis of Csd5-2xFLAG domain deletion variants. A. CellTool (Sycuro *et al.*, 2010) scatter plot showing cell population axis length vs. side curvature of strains WT (KBH127) , ΔNT (KBH128), ΔTM (KBH129), ΔQL (KBH131), ΔDO (KBH132), ΔSH3 (KBH133), and Δ*csd5* (LSH36). B. Smooth histogram showing cell population side curvature (x-axis) as a density function. C. Smooth histogram showing cell population axis length (x-axis) as a density function.

To ensure that our Csd5-2xFLAG variants were expressed we probed for Csd5 protein by immunoblotting using α-FLAG M2 antibody (Figure 3.1 C). The Csd5 ΔNT and ΔTM deletion

variants exhibited somewhat reduced protein expression while the SH3 deletion variant was expressed at similar levels to wild-type protein. Deletion of either the disordered (DO) domains or predicted coiled-coil (CC) motif in the glutamine rich linker resulted in helical cells with no (CC) or only minor (DO) changes to the overall helical cell shape parameters (Figure 3.1 B and Figure 3.3) and had wild-type levels of protein expression. These data indicate a functional role for the SH3 domain in cell shape generation. Furthermore, the N-terminal domains may be required for proper expression and/or localization of the SH3 domain.

Csd5 interacts with CcmA, MurF, and ATP synthase-

We investigated Csd5 protein interaction partners using immunoprecipitation (IP) and mass spectrometry with two independent Csd5 protein fusions, Csd5-VSV-G (Table 3.5, Table 3.6, Table 3.9, and Figure 3.5) and Csd5-2xFLAG (Figure 3.6. and Figure 3.8.). We first performed IP and mass spectrometry using Csd5-VSV-G alongside control IPs of functional (helical cell shape) VSV-G fusions to MinD (cytoplasmic) (Vecchiarelli *et al.*, 2016), Slt (periplasmic) (Chaput *et al.*, 2006), and a wild-type strain bearing no epitope fusion for controls. The top scoring proteins (Csd5-VSV-G IP) are known or predicted to function in the bacterial cytoplasm (Table 1, Table S3) and cytoplasmic membrane. The top unique hit in our proteomic screen was MurF, a cytoplasmic PG precursor synthesis enzyme that catalyzes the addition of D-Ala-D-Ala dipeptide to the terminal *meso*-diaminopimelic acid (*m*-DAP) residue of the UDP-linked NAM-tripeptide during cytoplasmic PG precursor biosynthesis (Smith, 2006). We also identified interactions with CcmA, a known *H. pylori* cell shape protein and bactofilin homolog that when mutated gives rise to curved-rod cells (Sycuro *et al.*, 2010). Finally, our data showed robust

enrichment for multiple components of the essential F_1F_O ATP synthase suggesting that Csd5 may be associated with this molecular machine (McGowan *et al.*, 1997).

Table 3.5. Top 10 proteins identified from Csd5-VSV-G immunoprecipitation and mass spectrometry

Rank[a]	Protein Name	*H. pylori G*27 Locus Tag	% Coverage (± SD)[b]	AVG PSM (± SD)[b,c]	Mass (kD)	Description
1	MurF	HPG27_696	64 ± 1	122 ± 7	56	Cytoplasmic PG precursor synthase
2	AtpA	HPG27_1079	62 ± 1	85 ± 9	55	α subunit: F_1ATP Synthase
3	AtpD	HPG27_1077	76 ± 5	78 ± 3	51	β subunit: F_1ATP Synthase
4	Csd5	HPG27_1195	52 ± 3	60 ± 6	22	Cell shape protein
5	AtpG	HPG27_1078	52 ± 3	28 ± 2	34	λ subunit: F_OATP Synthase
6	CcmA	HPG27_1480	67 ± 9	27 ± 1	15	Cell shape protein; Putative bactofilin
7	AtpH	HPG27_1080	57 ± 5	19 ± 3	20	δ subunit: F_1ATP Synthase
8	AtpF	HPG27_1081	34 ± 1	14 ± 1	20	b subunit: F_OATP Synthase
9	AtpC	HPG27_1076	53 ± 5	14 ± 2	13	ε subunit: F_1ATP Synthase
10	AtpX	HPG27_1082	49 ± 0	12 ± 1	16	b' subunit: F_OATP Synthase

Table 3.6. Top 20 scoring proteins uniquely present or five-fold enriched in Csd5-VSV-G vs. LSH100 (no bait control) IPs based on Positive Spectral Matches (PSMs)

*H. pylori G*27 Locus Tag	Protein Name	Description	LSH100[a]		Csd5-VSV-G		Fold Δ[b]
			PSM AVG	(± SD)	PSM AVG	(± SD)	
Proteins uniquely present in Csd5-VSV-G IP vs. Control							
HPG27_696	MurF	UDP-MurNac-pentapeptide presynthetase	0.00	0.00	122.00	7.00	na
HPG27_1079	AtpA	ATP synthase F1. subunit alpha	0.00	0.00	85.00	8.72	na
HPG27_1195	Csd5	Csd5	0.00	0.00	60.00	6.00	na
HPG27_1078	AtpG	ATP synthase F1, subunit gamma	0.00	0.00	28.00	2.00	na
HPG27_1480	ccmA	CcmA	0.00	0.00	26.67	1.15	na
HPG27_1080	AtpH	ATP synthase F1, subunit delta	0.00	0.00	19.00	3.00	na
HPG27_1081	AtpF	ATP synthase F0. subunit b	0.00	0.00	14.00	1.00	na
HPG27_1076	AtpC	ATP synthase F1, subunit epsilon	0.00	0.00	14.00	2.00	na
HPG27_1082	AtpX	ATP synthase F0. subunit b'	0.00	0.00	12.33	0.58	na
HPG27_1525	YaaW	hypothetical proteinDUF3944	0.00	0.00	2.33	1.15	na
HPG27_297		hypothetical protein	0.00	0.00	1.67	1.53	Na

HPG27_1099	MurG	UDP-N-acetylglucosamine lipid transferase	0.00	0.00	1.00	1.00	na
HPG27_896		hypothetical protein HPG27_896	0.00	0.00	1.00	0.00	na
HPG27_1327	Mod	putative type III restriction enzyme M protein	0.00	0.00	1.00	0.00	na
HPG27_980		208434933 ref|YP_002266599.1| adenine specific DNA methyltransferase	0.00	0.00	0.67	1.15	na
Proteins Enriched Greater than 5-fold vs. Control							
HPG27_1077	AtpD	ATP synthase F1, subunit beta	1.67	1.53	77.67	3.06	46.60
HPG27_257	GppA	guanosine pentaphosphate phosphohydrolase	1.00	1.00	7.33	0.58	7.33
HPG27_1225	TrpC	N-(5'-phospho-ribosyl) anthranilate isomerase	6.33	2.52	35.33	7.57	5.58
HPG27_228	HflC	hypothetical proteinSPFH Domain	3.33	1.53	17.33	2.52	5.20
HPG27_189	Mpr	ATP-binding protein	0.33	0.58	1.67	1.15	5.00

[a] similar results obtained from Slt-VSV-G and MinD-VSV-G IPs. [b] vs. LSH100 no bait control.

To validate these candidate interactions we performed reciprocal IPs using tagged strains followed by western blotting with polyclonal (CcmA, Csd5, AtpD) and monoclonal (FLAG, VSV-G) antibodies. We generated individual strains bearing CcmA-2xFLAG, MurF-3xVSV-G, or AtpF-3xFLAG expressed at the native locus and showed these fusions are functional (Figure 3.4).

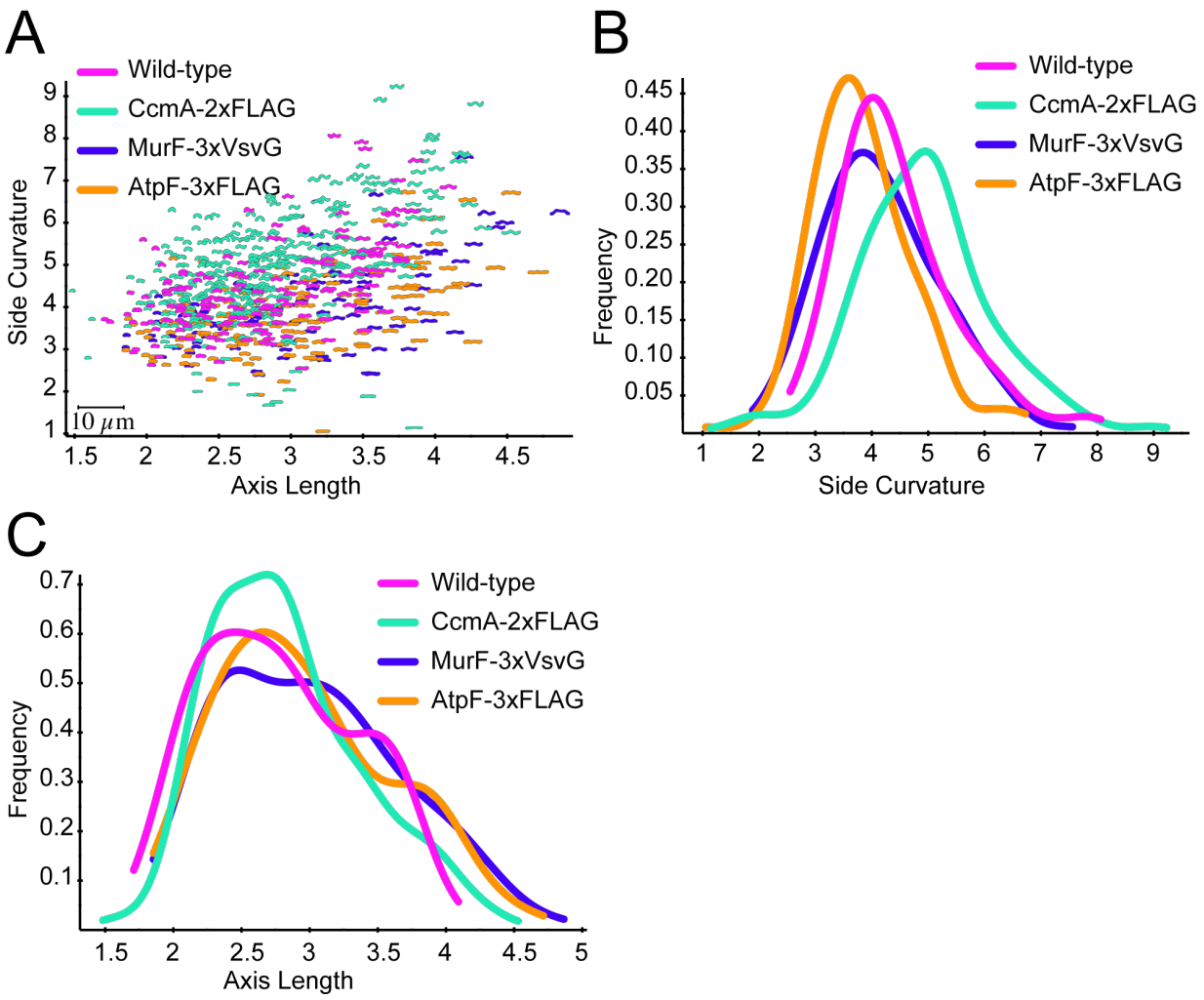

Figure 3.4. Quantitative cell shape analysis of CcmA, MurF, and AtpF fusion proteins. A. CellTool (Sycuro *et al.*, 2010) scatter plot showing cell population axis length vs. side curvature of strains wild-type (LSH100, no tag), CcmA-2FLAG (JTCF), MurF-3xVSV-G (KBH157), and AtpF-3xFLAG (KBH159). B. Smooth histogram showing cell population side curvature (x-axis) as a density function. C. Smooth histogram showing cell population axis length (x-axis) as a density function

CcmA-2xFLAG cells have normal helical shape and MurF-3xVSV-G and AtpF-3xFLAG cells are both viable (*murF* and *atpF* are essential) and helical. AtpF (b-subunit*)* in the F_1F_O ATP

synthase is part of the peripheral stalk that connects the membrane embedded F_O complex to the cytoplasmic F_1 ATPase (Weber, 2006; Rühle and Leister, 2015) (Figure 3.8. A).

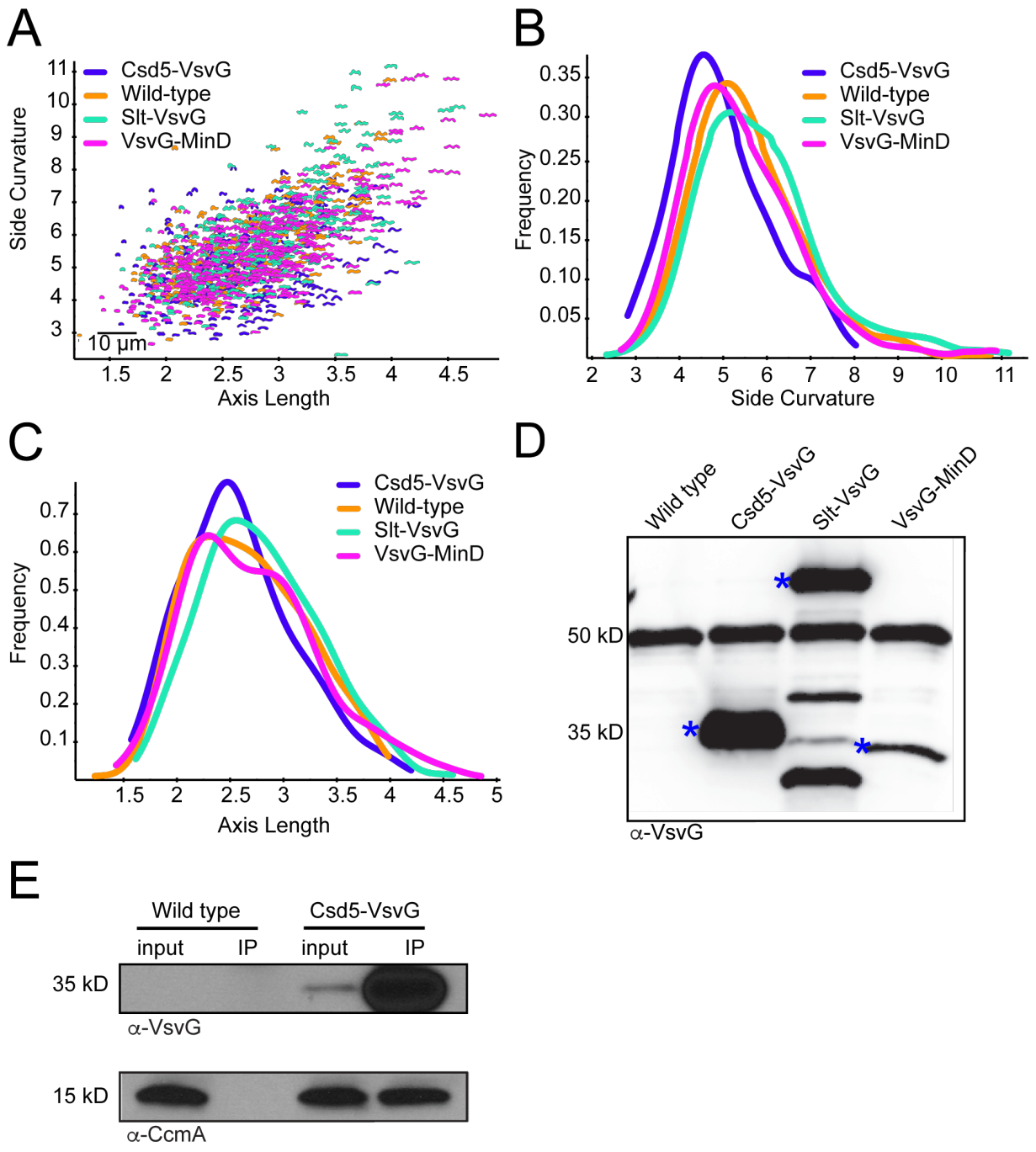

Figure 3.5. Validation of VSV-G protein fusions and IPs used to identify Csd5 interactions.

For CcmA, MurF, and AtpF we demonstrate reciprocal Co-IP of Csd5 or Csd5-2xFLAG (Figure 3.6. A). In addition, we show that with MurF, we also Co-IP CcmA and AtpD (b-subunit) of ATP synthase (Figure 3.6. B). Taken together, these results suggest that Csd5, MurF, CcmA, and ATP synthase interact as part of one or more multi-protein complexes.

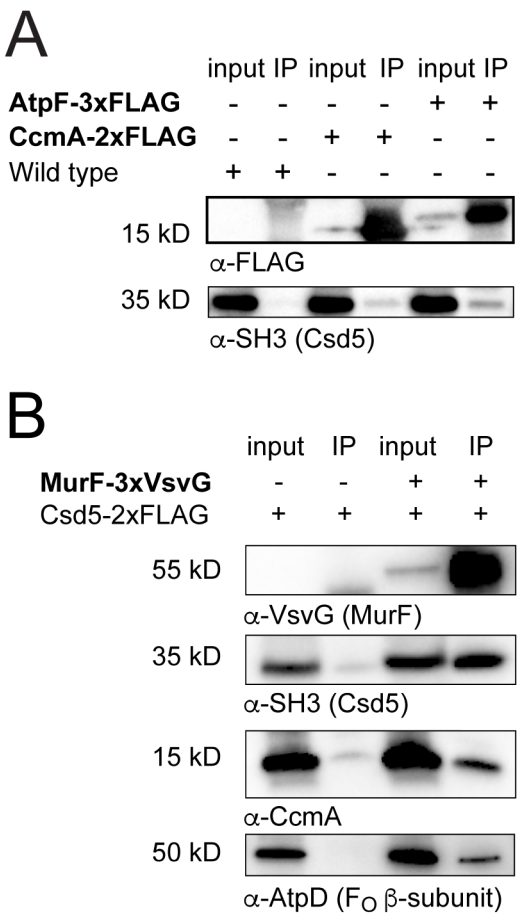

Figure 3.6. Csd5 interacts with MurF, CcmA, and ATP synthase.

H. pylori strains bearing the indicated protein fusions were subject to immunoprecipitation (IP) using detergent solubilized whole cell extracts and α-FLAG (A) or α-VSV-G (B) antibodies to pull-down CcmA-2xFLAG (JTH3), AtpF-3xFLAG (KBH159), MurF-3xVSV-G (KBH157) or a no tag-control (A, LSH100; B, KBH127). A. CcmA (15 kD) and AtpF (20 kD) were detected in the input and IP fractions with α-FLAG while Csd5 (35 kD) was detected with α-SH3 antibody. B. Csd5, CcmA and the b-subunit of ATP synthase (AtpD) were detected in the input

and IP fractions with α-SH3, α-CcmA, and α-AtpD antibodies. Representative experiments shown from a minimum of 2 biological replicates.

The Csd5 N-terminal domains are required for interactions with MurF and CcmA-

We next used the domain deletions to begin to localize Csd5 interactions with MurF, CcmA, and ATP synthase. We could uniformly IP each of the variant bait proteins (Figure 3.7. A) with comparable yields despite lower levels of protein expression for the ΔNT and ΔTM domain deletion strains (Figure 3.1 C). Deletion of either the Csd5 DO domain or C-terminal SH3 domain had no effect on our ability to Co-IP MurF, CcmA, or AtpD (Figure 3.7. B and E). However, loss of the TM domain resulted in a complete loss of MurF and CcmA positive spectra (Figure 3.7. B) and loss of CcmA and AtpD protein by western blotting (Figure 3.7. E). In addition, we observed an overall decrease in the abundance of each component of the ATP synthase (Figure 3.8. B). The NT domain also appears to be important for these interactions as we observed a significant decrease in MurF spectra and a complete loss of CcmA. These data suggest that the NT and TM domain of Csd5 are required for interactions with CcmA, MurF, and ATP synthase.

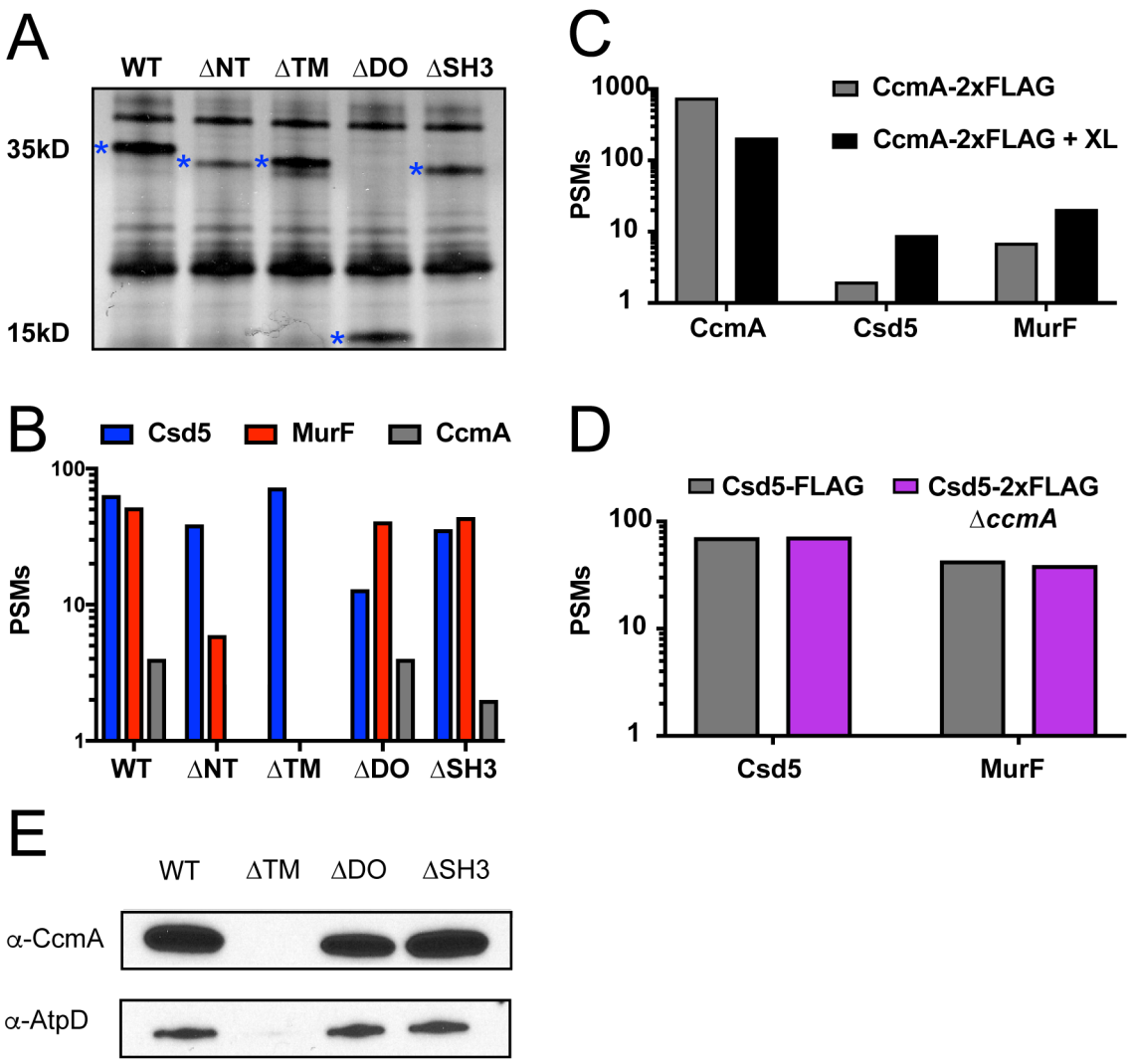

Figure 3.7. Csd5 N-terminal and transmembrane domains promote protein-protein interactions with MurF, CcmA, and ATP Synthase.

A. Silver stain of Csd5 domain deletion immunoprecipitation (IP) fractions. Bait proteins indicated with a blue asterisk. B. Mass spectrometry positive spectral match (PSM) counts for Csd5, CcmA and MurF from Csd5 domain deletion IPs; WT (KBH127), ΔNT (KBH128), ΔTM (KBH129), ΔDO (KBH132) and ΔSH3 (KBH133). C. PSM counts for CcmA, Csd5, and MurF from CcmA-2xFLAG IPs with and without formaldehyde crosslinking. D. PSM counts for Csd5 and MurF from Csd5-2xFLAG (KBH127) and Csd5-2xFLAG Δ*ccmA* (KBH126) IPs. E. Immunoblot detection of CcmA and AtpD in IPs of indicated Csd5 deletion variants using α-CcmA and α-AtpD antibodies respectively.

To probe whether Csd5, MurF, CcmA, and ATP synthase exist in a single complex, we performed IP and mass spectrometry using CcmA-2xFLAG as bait in the presence or absence of formaldehyde crosslinking. We observed a crosslinking specific increase in spectral counts for both MurF and Csd5 (Figure 3.7. C) and also for components of the ATP synthase (Table 3.10). We next tested whether Csd5 could pull down MurF in the absence of CcmA using a *ccmA* deletion strain. Deletion of *ccmA* had no effect on immunoprecipitation of MurF by Csd5 (Figure 3.7. D) suggesting that the Csd5 and MurF interaction is direct. Taken together we conclude that Csd5, MurF, and CcmA interact individually and/or together with ATP synthase to promote maintenance of cell shape.

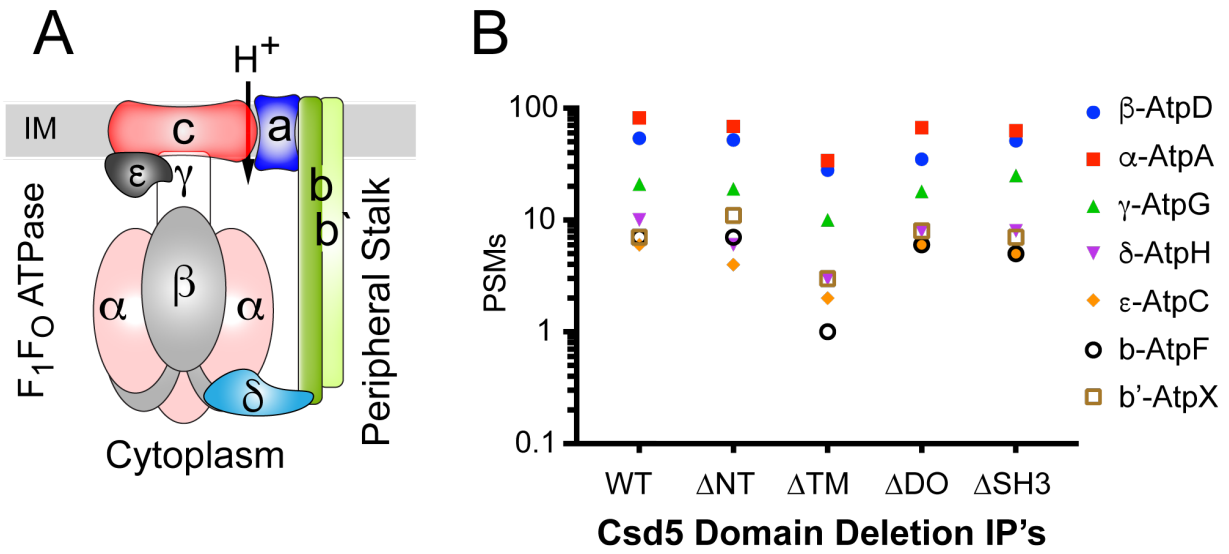

Figure 3.8. Csd5 transmembrane domain promotes interaction with F_1F_O ATP synthase. A. Schematic and topology of the *H. pylori* F_1F_O ATP synthase. IM (Inner Membrane). B. Positive spectral matches (PSMs) for each identified subunit of the F_1F_O ATP synthase identified from Csd5 domain deletion immunoprecipitations (IPs); WT (KBH127), ΔNT (KBH128), ΔTM (KBH129), ΔDO (KBH132), and ΔSH3 (KBH133) in Figure 3.1.

The SH3 domain RT-loop is required for SH3 domain stability and interaction with PG-

None of the observed Csd5 protein-protein interactions required the SH3 domain. To further probe the function of the SH3 domain we investigated a putative interaction with the PG cell wall. Bacterial SH3 domains have been implicated as cell wall targeting domains that bind directly to PG (Baba and Schneewind, 1996; Lu *et al.*, 2006). We identified an amino acid Csd5^{R146} in the RT-loop motif that is strictly conserved in *H. pylori* and also conserved with the cell wall targeting SH3 domain of Lystostaphin (ALE-1, PDB ID 1R77) from *Staphylococcus simulans* (Figure 3.2 and Figure 3.9.) (Lu *et al.*, 2006). We identified a second amino acid, Csd5^{T151} that is slightly less conserved in both Csd5 and ALE-1 and is positioned at the tip of the surface-accessible RT-loop in our homology model (Figure 3.9. A). We generated strains expressing single copy variants of *csd5-R146A* or *csd5-T151A* at the native locus. Mutation of the conserved arginine resulted in undetected Csd5 protein *in vivo* and a loss of helical cell shape apparently from protein instability (Figure 3.9. B). In contrast, the Csd5-^{T151A} mutant resulted in no visible change in protein production but cell shape was affected, yielding cells with reduced side curvature indicating this loop region is important for maintaining helical cell shape (Figure 3.9. B).

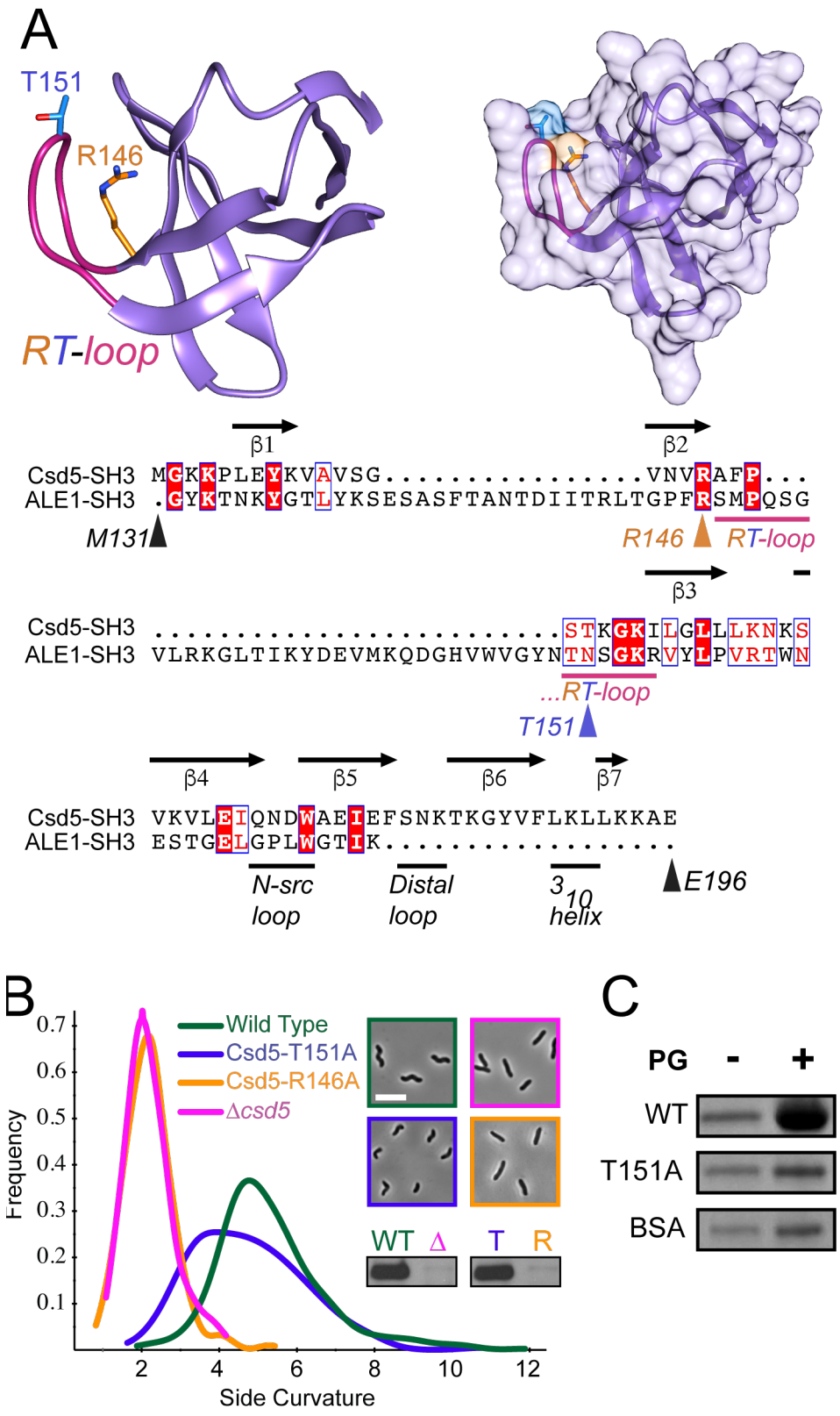

Figure 3.9. The Csd5-SH3 domain interacts with peptidoglycan.

A. Phyre2 (Kelley *et al.*, 2015) homology model (>99% confidence) of the Csd5 SH3 domain highlighting the RT-loop (pink) and mutated sites R146 (Orange) and T151 (Blue) as sticks (Left). Surface representation of SH3 domain model (Right). Graphics generated using Chimera (Pettersen *et al.*, 2004). Sequence alignment of *H. pylori* Csd5 SH3 domain with *Staphyloccocus simulans* ALE-1 cell wall binding domain (PDB ID 1R77) and Csd5 secondary structure and loop domain assignments (RT-loop, N-src loop, distal loop, and 3_{10} helix) based on established SH3 domain loop nomenclature (Noble *et al.*, 1993; Kamitori and Yoshida, 2015). Colored arrows or bars indicate mutated sites and loops, respectively (Bottom). Alignment graphic generated using Clustal Omega (Sievers *et al.*, 2011) and ESPRIT 3.0 (Robert and Gouet, 2014). B. Phase-contrast microscopy and quantitative cell shape analysis using CellTool software (Sycuro *et al.*, 2010) of wild-type (LSH100), Δ*csd5* (LSH36) and complemented full length Csd5-SH3 variants R146A (KMH1) and T151A (KMH2). Smooth histogram showing cell population side curvature (x-axis) as a density function (188-382 cells analyzed per strain). Western blots performed using whole cell extracts with an α-Csd5-SH3 domain antibody. C. SDS-PAGE of PG sedimentation assay in the presence (+) or absence (-) of PG cell walls with purified GST-SH3 fusions to wild-type, mutant (T151A) or BSA as a negative control, stained with Coomasie Blue. Representative of 2 independent experiments.

To determine if the SH3 domain can interact with PG directly, we purified wild-type and mutant (T151A) SH3 domains as recombinant fusions to GST Figure 3.12. Each protein (WT and T151A variant or controls) was individually mixed and incubated in the presence or absence of purified PG from wild-type *H. pylori* bacteria. We show that the wild-type but not the T151A variant is pulled down in a PG dependent manner (Figure 3.9. C and Figure 3.10. A). Cleavage and separation of the GST domain revealed that the WT SH3 domain in isolation binds PG sacculi (Figure 3.10. B) while GST alone does not (Figure 3.10. C). These combined data establish the RT-loop as a functional determinant for direct SH3 domain interaction with the bacterial cell wall.

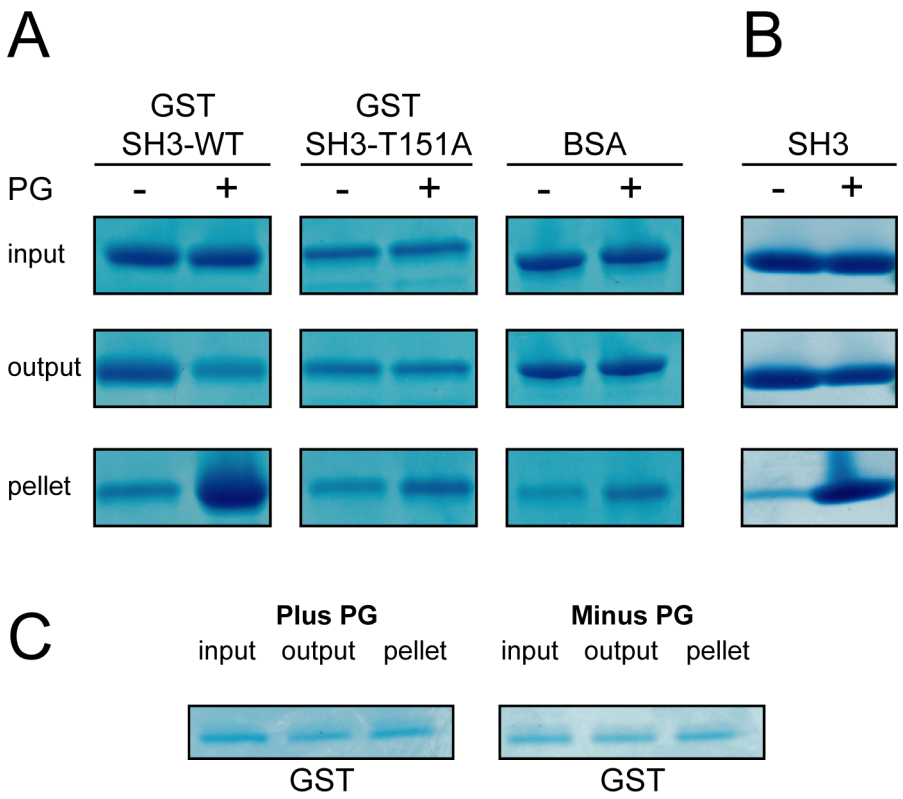

Figure 3.10. The Csd5-SH3 domain interacts with peptidoglycan with and without GST domain.

A. SDS-PAGE of peptidoglycan (PG) sedimentation assay using purified PG cell walls with purified GST-SH3, GST-SH3-T151A or BSA as a negative control, stained with Coomasie Blue; input and output fractions from Fig. 4C. B. PG sedimentation assay using cleaved GST-SH3. C. PG sedimentation assay using GST (Sigma) alone.

3.5 Discussion

Curvature generation in curved-rod organisms such as *Caulobacter crescentus* and *Vibrio cholera* is thought to result from biased PG insertion along the outer curve of the cell, opposite the inner curvature localized cytoskeletal or periskeletal cell spanning filaments (Cabeen *et al.*, 2009; Bartlett *et al.*, 2017). In contrast, the helical curvature of *H. pylori* has been suggested to result from modulation of PG crosslinking (Sycuro *et al.*, 2010; Bonis *et al.*, 2010). Shared

phenotypes (straight-rod cell shape) between Δ*csd5* and the periplasmic PG carboxypeptidases Δ*csd6* and Δ*csd4* led to a hypothesis that Csd5 may localize these enzymes to either promote or inhibit the activity of periplasmic PG hydrolases that successively trim uncross-linked PG-tetrapeptide to PG-dipeptide that can no longer participate in crosslinking (Sycuro *et al.*, 2012; Sycuro *et al.*, 2013; Kim *et al.*, 2014; Chan *et al.*, 2015). Despite significant efforts and a report that recombinant Csd5 and Csd4 interact directly (Kim *et al.*, 2014), we have not been able to reproduce this finding with our *in vivo* IP's and formaldehyde crosslinking experiments. Instead, in this work we have identified and localized Csd5 interactions with a cell shape associated putative bactofilin (CcmA), a PG precursor synthase (MurF) and to the PG cell wall. In addition, we have identified an association between these factors and ATP synthase. These results suggest the existence of a shape promoting protein complex that spans the cytoplasmic membrane to connect the cell wall in the periplasm with the cytoplasmic cytoskeleton (Figure 3.11). Future studies to investigate both the role of the ATP synthase as well as the membrane association and filament formation activities of CcmA would lend further support to this model.

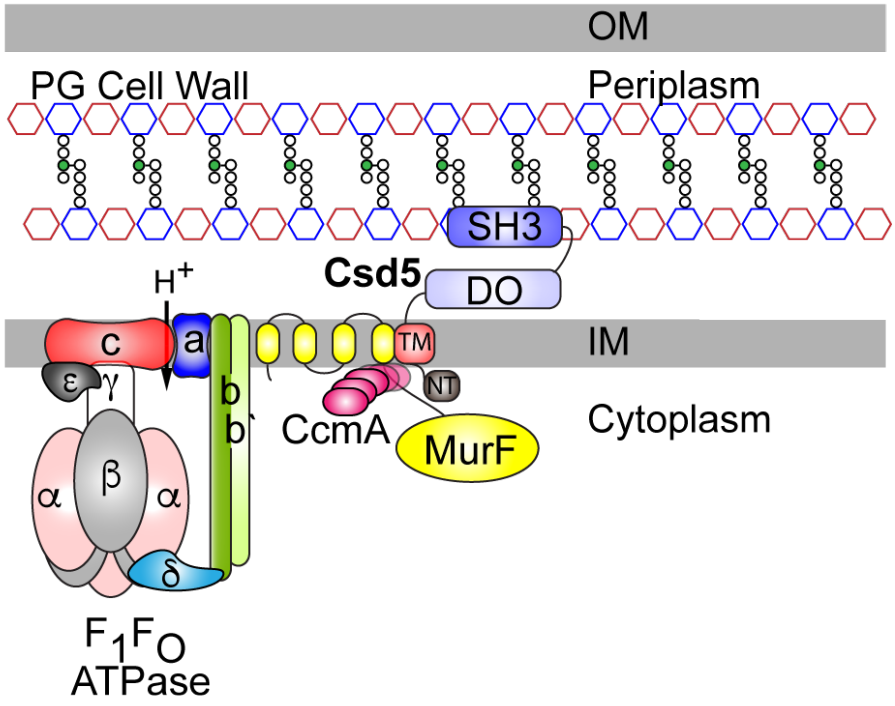

Figure 3.11. Model of the Csd5 *H. pylori* cell shape promoting protein complex.

Csd5 protein is anchored in the inner membrane where it mediates interactions between a putative cytoskeletal bactofilin CcmA, the PG precursor synthase MurF (D-Ala-D-Ala ligase), and the rotary F_1F_O ATP synthase. The C-terminal SH3 domain interacts with peptidoglycan in the periplasm and links Csd5 to cytoplasmic cell shape and cell elongation factors essential for helical shape (CcmA) and growth (MurF). OM (outer membrane), IM (inner membrane).

The SH3 domain of Csd5 could serve as a sensor for recognition of specific PG ligands for proper positioning of the protein complex in the membrane. The similarity of the *H. pylori* Csd5 SH3 domain to NlpC/P60 SH3b2 domain (Xu *et al.*, 2015) and to the Gram (+) ALE-1 cell wall targeting domain (Lu *et al.*, 2006) suggests it may recognize either free stem peptides (e.g. di-, tri-, tetra-, or penta-) or specific crosslink species present in the cell wall sacculus (Gründling and Schneewind, 2006). Crystal structures of ALE-1 with (PDB ID 5LEO) and without (PDB ID

1R77) a penta-glycine PG fragment are deposited in the RCSB (PDB; http://www.rcsb.org/pdb) protein databank (Berman *et al.*, 2000). Our finding that mutations in conserved residues of the RT-loop can affect cell shape and disrupt binding to PG *in vitro*, indicates a role for the SH3 RT-loop in PG substrate recognition.

The ATP dependent synthesis of PG precursors is carried out in the cytoplasm and is divided into cytosolic (MurA-F) and membrane associated (MurG, MraY, MurJ) steps (Laddomada *et al.*, 2016). In *H. pylori* it appears that MurF may deviate from this model as iterative searching using Jackhmmer seeded with the *H. pylori* MurF sequence identified 4 predicted N-terminal transmembrane domains preceding the enzymatic domain of this highly conserved and essential protein. A PSI-BLAST search (excluding the epsilonproteobacteria) seeded with the *H. pylori* MurF sequence was limited to orthologs with greater than 80% query coverage. This analysis revealed widespread prevalence of N-terminal transmembrane extensions in MurF from Gram-negative and, Gram-positive, organisms with minimal or reduced genomes (Nelson and Stegen, 2015; Brown *et al.*, 2015) (data not shown). That our Csd5 ΔTM deletion mutants fail to IP MurF, CcmA, and ATP synthase may suggest intra-membrane protein interactions that maintain these proteins within a single complex. MurF has been shown to associate directly with the elongasome and cytoskeleton through interactions with MurG (Favini-Stabile *et al.*, 2013), MraY (White *et al.*, 2010), and MreB (Mohammadi *et al.*, 2007). Consistent with this, we observe a low level of MurG spectral hits in our Csd5-VSV-G IP (Table 3.6). Csd5 may thus be localized to sites of new PG synthesis through its association with MurF to modulate cell wall growth.

The *H. pylori* ATP synthase subunit stoichiometry varies from the prototypical bacterial synthase, where instead of a single dimeric b-subunit, there exist two distinct b (AtpF) and b'

(AtpX) subunits that form the peripheral stalk dimer (Figure 3.8. and Figure 3.11) (Weber, 2006). In both mitochondria and chloroplasts, the formation of ATP synthase dimers and even larger superassemblies have been associated with regions of high membrane curvature (Rexroth *et al.*, 2004; Seelert and Dencher, 2011; Hahn *et al.*, 2016). In mitochondria, the dimer rows of ATP synthase are essential for the formation of highly curved lamellar cristae (Paumard *et al.*, 2002; Davies *et al.*, 2012). Specialized subunits promote dimerization at the interface between the peripheral stalks of adjacent ATP synthases (Arnold *et al.*, 1998). To date, ATP synthase dimers have not been observed in Eubacteria and future work will explore this possibility. Alternatively, association between MurF and ATP synthase could be involved in modulating local rates of PG synthesis or ATP turnover (Lee *et al.*, 2013).

The identification of Csd5 protein interactions with CcmA, a known cell shape protein and putative cytoskeletal bactofilin, and with MurF, a known cell elongation factor, were unexpected but not surprising given the connections between PG synthesis, cell shape and intermediate filament proteins in curved organisms (Ausmees *et al.*, 2003; Koch *et al.*, 2011). We have discovered the first example of a Gram-negative SH3 domain interaction with peptidoglycan and provide evidence of a cell shape promoting protein complex in *H. pylori*. Future work characterizing these interactions in more detail will expand our understanding of the diversity of mechanisms cells use to regulate and control cell shape.

3.6 METHODS AND SUPPORTING INFORMATION

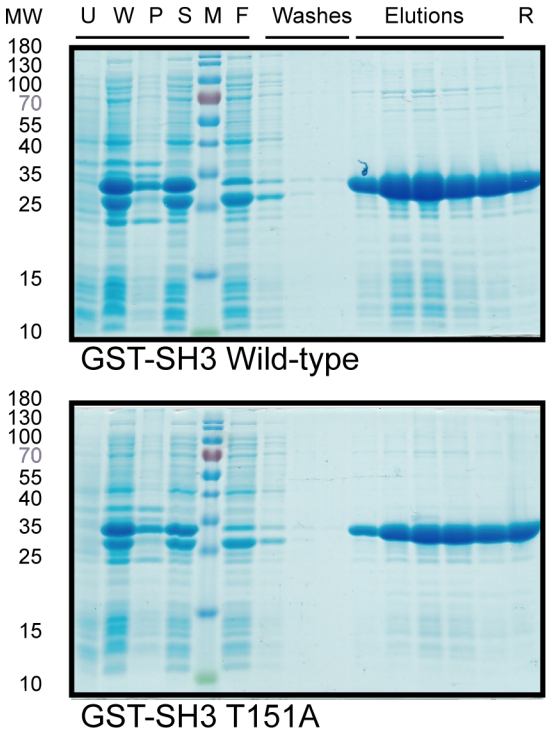

Figure 3.12. Purification of wild-type and mutant GST-SH3 domains.

SDS-PAGE gels of GST-SH3 glutathione affinity resin purifications stained with Coomasie Blue. GST-SH3 wild-type (KBE7) and GST-SH3-T151A (KBE58). Uninduced (U), Whole-cells (W), Pellet (P), Soluble (S), M (Marker, PAGE-RULER), Flow-thru (F), R (Resin).

Table 3.7. Strains and plasmids used in this study

Strains and Plasmids	Relevant or Descriptive Genotype	Reference or Source
H. pylori Strains		
LSH100	Wild-type	Sycuro et al., 2010
LSH36	*csd5::catsacB*	Sycuro et al., 2012
LSH108	*rdxA::kansacB*	Sycuro et al., 2010
KBH17	*csd5-WT*	This work
KMH1	*csd5-R146A*	This work
KMH2	*csd5-T151A*	This work
TRH4	*csd5-VSV-G*	This work
MHH9	*slt::cat*	This work
KBH89	*slt::cat rdxA::slt-VSV-G*	This work
KBH91	*minD::kansacB*	This work

KBH103	*VSV-G-minD*	This work
JTH1	*csd5-2xFLAG::cat*	This work
JTH2	*ccmA-2xFLAG::cat*	This work
JTH3	*ccmA-2xFLAG::kan*	This work
KBH126	*csd5-2xFLAG::cat, ccmA::kan*	This work
KBH127	*csd5-WT-2xFLAG*	This work
KBH128	*csd5-ΔNT-2xFLAG*	This work
KBH129	*csd5-ΔTM-2xFLAG*	This work
KBH131	*csd5-ΔQL-2xFLAG*	This work
KBH132	*csd5-ΔDO-2xFLAG*	This work
KBH133	*csd5-ΔSH3-2xFLAG*	This work
KBH157	*csd5-2xFLAG, murF-3xVSV-G::cat*	This work
KBH159	*atpF-3xFLAG::cat*	This work
E. coli Strains		
KBE7	BL21-DE3 pKB7	This work
KBE58	BL21-DE3 pKM5	This work
KBE57	BL21-DE3 pKM4	This work
Plasmids		
pCR2.1 TA	*pTOPO* TA	Life Technologies
pCR4 Blunt-TOPO®	*pTOPO* BLUNT	Life Technologies
pLC292	*pRdxA*	Terry et al., 2005
pGEX2T	*GST* Expression vector	Life Technologies
pJF-1480	*GST-ccmA*	This work
pGEX4T3	*GST* Expression vector	Life Technologies
pKB7	*GST-csd5 (wild-type)*	This work
pKM5	*GST-csd5-T151A*	This work
pKB21	*pTOPO::csd5-WT*	This work
pKM1	*pTOPO::csd5-R146A*	This work
pKM2	*pTOPO::csd5-T151A*	This work
pKB57	*pTOPO::slt-VSV-G*	This work
pKB59	*rdxA::slt-VSV-G*	This work
pKB61	*pTOPO::minD::kansacB*	This work
pKB66	*pTOPO::VSV-G-minD*	This work
pKB80	*pTOPO::csd5-2xFLAG::cat*	This work
pKB100	*pTOPO::csd5-2xFLAG*	This work
pKB101	*pTOPO::csd5-ΔNT-2xFLAG*	This work
pKB102	*pTOPO::csd5-ΔTM-2xFLAG*	This work
pKB104	*pTOPO::csd5-ΔCC-2xFLAG*	This work
pKB105	*pTOPO::csd5-ΔDO-2xFLAG*	This work
pKB106	*pTOPO::csd5-ΔSH3-2xFLAG*	This work
pKB113	*pTOPO::murF-3xVSV-G*	This work
pKB116	*pTOPO::atpF-3xFLAG*	This work

Table 3.8. Primers used in this study

Primer ID	Description	Sequence 5`3`
For Plasmids and Strain Construction		
JTH1		
JTP1	G27 1195 F1 P1	CAAGAAACTAAACAAGAGCAA
JTP2	G27 1195 FLAG P3	TCACTTATCGTCATCGTCTTTATAATCCTTATCGTC ATCGTCTTTATAATCTTCAGCCTTTTTTAAAAGTTT
JTP3	G27 1195 CATDN P4	CCCAGTTTGTCGCACTGATAAAAGGCTGAATGAA AGAATAATGAAATT
JTP4	G27 1195 DNSTM P2	CACAAGCTCATCATCTTCCAA
JTP5	G27 464 FLAG CATUP P5	CTTATCGTCATCGTCTTTATAATCTAGGATATAGA TTGAAAAGTGGAT
JTH2 and JTH3		
JTP6	G27 1480 F1 P1	TTGCATGTAGATGGCGAATTA
JTP7	G27 1480 FLAG P3	CTACTTATCGTCATCGTCTTTATAATCCTTATCGTC ATCGTCTTTATAATCTTTATTTTCAATTTTCTTTTC
JTP8	G27 1480 CATDN P4	CCCAGTTTGTCGCACTGATAAATTGAAAATAAATA GGGAATGATCCAATC
JTP9	G27 1480 DNSTRM P2	AAGGGTGCAATAACGCGCTAA
TRH4		
THP1	csd5-VSV-G UP F	GCGCCGCCTTTAGACACAGCC
THP2	csd5-VSV-G UP R	ATCCACTTTTCAATCTATATCTCATCATTTGCCTAA TCTATTCATTTCAATATCAGTATATCATTTGCCTAA TCTATTCATTTCAATATCAGTATATTCAGCCTTTTT TAAAAGTTT
THP3	csd5-VSV-G DN F	CCCAGTTTGTCGCACTGATAAAAGAATAATGAAAT TAAAATCT
THP4	csd5-VSV-G DN R	TCATCATTTTCTAAAACAAGC
MHH9		
MLP1	slt::cat UP F	GCAAAGTAAGCATGTGAG
MLP2	slt::cat UP R	ATCCACTTTTCAATCTATATCGTGGGTAGCGTTAC TTTT
MLP3	slt::cat DN F	CCCAGTTTGTCGCACTGATAACCTTATGAGGGCAT TATT
MLP4	slt::cat DN R	GGGGTTAAACATGTCATT
pKB57		
1938	slt F	AGGAGTCTAGATTCGCTCCTTTAGTGGTGGTTGTT
1939	slt R	TCACTTACCCAGGCGGTTCATTTCGATATCAGTGT ATGGTTTGTCCTTTGAGTTAAGAGTTT
pKB61		
1954	minD::ksb UP F	TAACGCTAAAGCGGTGGCTTTAAG
1955	minD::ksb UP R	ACCTCAAATGGTTCGCCCGGGAAGTAACTACTATT GCCATATGATTCC
1956	minD::ksb DN F	GAAAATGGGCTTATCCGCTCGAGTTTAAAAGGGA TATTTTCATGAGTTTG
1957	minD::ksb DN R	GCGCGATTTCTTGGATCACTTTATG
pKB66		
1970	VSV-G-minD UP F	GAAGAATGCGAGAGCGATTTATTCG

1971	VSV-G-minD UP R	CTTACCCAGGCGGTTCATTTCGATATCAGTGTACA TATGATTCCTTTTAA
1972	VSV-G-minD DN F	TATCGAAATGAACCGCCTGGGTAAGGCAATAGTA GTTACTATCACTTCAG
1973	VSV-G-minD DN R	TCTATCTGTTGCGGTGGCCGCG
pKB21		
1461	csd5-TOPO-F	CGTAATGCGATTTTATTAGTGGTGG
1462	csd5-TOPOR	CGGTATTGTAACCCACAAGCTCAT
pKB100-106	Universal	
2177	Csd5-2xFLAG UP R	TTATTCTTTCACTTATCGTCATCGTC
2178	Csd5-2xFLAG DN F	ATAAAGACGATGACGATAAGTGAAAGAATAATGA AATTAAAATCTTTTG
pKB101		
2185	ΔNT UP R	CATTCTTACCCTCAACGCTTTTCAA
2186	ΔNT DN F	AAGCGTTGAGGGTAAGAATGCGGCCATCATTAGT GGTTTTAGC
pKB102		
2187	ΔTM UP R	CATAAAAAGTTTTAAAGAAGATTTCATCT
2188	ΔTM DN F	CTTCTTTAAAACTTTTTATGGGTTTTTATGTGAAAA AAGACAGCG
pKB104		
2189	ΔCC UP R	TGCGGTCGTTTCTAAAGGTGGC
2190	ΔCC DN F	CACCTTTAGAAACGACCGCACAAGAAAAAGAAAG CGAGCCTAAACAA
pKB105		
2183	Δ DO UP R	CGCTGTCTTTTTTCACATAAAAACCA
2184	Δ DO DN F	TTATGTGAAAAAAGACAGCGATGGGAAAAAAACC TTTAGAGTATAAAG
pKB106		
2179	ΔSH3 UP R	TCCCATTGTAGAGATTGTAGGGG
2180	ΔSH3 DN F	CTACAATCTCTACAATGGGAGATTATAAAGACGAT GACGATAAGGATTAT
pKB113		
2216	VSV-G-murF UP F	GTGTGAAAACGAACGGATGGTTTTA
2251	MurF-3xVSV-G UP R	TCATTTGCCTAATCTATTCATTTCAATATCAGTATA TTTGCCTAATCTATTCATTTCAATATCAGTATATTT GCCTAATCTATTCATTTCAATATCAGTATAAATGT AATTAGGGGCGTCATTAGCG
2250	3xVSV-G::cat F	TATACTGATATTGAAATGAATAGATTAGGCAAATA TACTGATATTGAAATGAATAGATTAGGCAAATATA CTGATATTGAAATGAATAGATTAGGCAAATGATG AGATATAGATTGAAAAGTGGAT
2209	MurF-VSV-G::cat DN F	CCCAGTTTGTCGCACTGATAAGAAATGAACATGCA ACATTTATACGC
2211	MurF-VSV-G::cat DN R2	TCAATGAAGCTCCTTTTCTAATAAGTT
pKB116		
2199	atpX UP F	GTGGGCGATGAATGTTTGGGTGTA
2274	atpX UP R	GATTATAAAGACGATGACGATAAGGATTATAAAG ACGATGACGATAAGGATTATAAAGACGATGACGA TAAGTGAGATATAGATTGAAAAGTGGAT

2248	3xFLAG cat F	CCCAGTTTGTCGCACTGATAATAAATGCAAGATTTAAAAGTGATCTCTA
2275	aptF-3xFLAG DN F	CCAAGTGTCTTGAGCGATTTCTAC
2276	atpF 3xFLAG DN R	CCAAGTGTCTTGAGCGATTTCTAC
Protein Expression Vectors		
pKB7 and pKM5		
1448	Csd5_SH3_BamHI F	AGGAGGGATCCGGAAAAAAACCTTTAGAGTATAAAGTC
1449	Csd5_SH3_XhoI R	CTCCTCTCGAGTCATTCAGCCTTTTTTAAAAGTTTTAAAA
pJF-1480		
JF1480 F	GST-CcmA R BamHI	CGGATCCATGGCAATCTTTGATAACAAT
JF1480R	GST-CcmA F EcoRI	GGAATTCCCTATTTATTTTCAATTTTCTTTTC
Site-Directed Mutagenesis Primers		
pKM1		
1506	Csd5 R146A SDM	CAGTCAGTGGCGTGAATGTGgcCGCTTTTCCTAGCACAAAAG
pKM2		
1507	Csd5-T151A SDM	ATGTGCGCGCTTTTCCTAGCgCAAAAGGTAAAATCTTGGGA
Universal Primers		
828	catR-F	GATATAGATTGAAAAGTGGAT
829	catR-R	TTATCAGTGCGACAAACTGGG
1948	Universal kansacB-F	TCCCGGGCGAACCATTTGAGGT
1949	Universal kansacB-R	ACTCGAGCGGATAAGCCCATTTTC

Table 3.9. Top 20 scoring proteins shared between Csd5-VSV-G and LSH100 (no-bait control) IPs based on Positive Spectral Matches (PSMs)

H. pylori Protein Accession/Description	LSH100		Csd5-VSV-G		Fold Δ[a]
	PSM AVG	(± SD)	PSM AVG	(± SD)	
YP_002266135.1cytotoxin-associated protein A	78.33	7.64	75.67	7.57	0.97
YP_002266768.1translation elongation factor EF-Tu	54.33	2.08	63.33	4.73	1.17
YP_002266196.1putative vacuolating cytotoxin(VacA)-like protein	37.67	10.50	55.33	1.53	1.47
YP_002265704.1urease B	36.33	7.23	35.00	5.20	0.96
YP_002265705.1urease A	29.67	4.73	28.00	2.00	0.94
YP_002267108.1alkylhydroperoxide reductase	26.67	3.51	18.33	0.58	0.69
YP_002267086.1ubiquinol cytochrome c oxidoreductase	20.00	4.36	10.67	0.58	0.53
YP_002266865.1ribosomal protein S5	14.67	1.53	18.33	1.53	1.25

YP_002266763.1ribosomal protein L1	14.33	2.08	10.67	1.53	0.74
YP_002267127.1pyridoxal phosphate biosynthetic protein J	13.33	4.62	7.67	1.53	0.58
YP_002265752.1hypothetical protein HPG27_117	12.33	4.73	6.67	0.58	0.54
YP_002266297.1hypothetical protein HPG27_676	10.33	3.79	6.00	1.73	0.58
YP_002266081.1ribosomal protein L28	8.67	4.93	9.33	5.51	1.08
YP_002267100.1ribosomal protein S2	7.67	3.06	14.33	2.89	1.87
YP_002266058.1oligoendopeptidase F	7.67	1.15	10.00	1.73	1.30
YP_002265651.1chaperone and heatshock protein	7.33	1.53	17.67	3.06	2.41
YP_002265814.1S-adenosyl methionine synthetase	7.00	2.65	6.00	1.73	0.86
YP_002266839.1N-(5'-phospho-ribosyl) anthranilate isomerase	6.33	2.52	35.33	7.57	5.58
YP_002266110.1cag pathogenicity island protein 3	6.33	4.04	3.67	3.21	0.58
YP_002265652.1cochaperone protein	5.67	0.58	5.00	1.00	0.88

[a] vs. LSH100 no bait control

Table 3.10. ATP synthase subunits in CcmA-2xFLAG crosslinking IPs

ATP Synthase Subunits Identified	**Positive Spectral Matches**	
	CcmA IP (-)	CcmA IP (+)
CcmA-2xFLAG	762	209
ATP synthase F1; subunit beta	14	35
ATP synthase F1; subunit gamma	2	12
ATP synthase F1; subunit alpha	11	37
ATP synthase F1; subunit delta	0	2
ATP synthase F0; subunit b	0	1
ATP synthase F0; subunit b'	0	2

No crosslinking, + = with formaldehyde

Experimental Methods

Bacterial Strains and Growth Conditions

Strains used in this work, as well as primers and plasmids used in strain construction are described in Table S2 and S3. *H. pylori* was grown in Brucella broth supplemented with 10%

fetal bovine serum (Invitrogen) without antimicrobials or on horse blood agar plates with antimicrobials as described (Chan *et al.*, 2015) Bacteria were incubated at 37°C under micro-aerobic conditions in a tri-gas incubator (10% CO_2, 10% O_2 , and 80% N_2). For resistance marker selection, horse blood plates were supplemented with 15 μg/ml chloramphenicol; 25 μg/ml kanamycin; or 60 mg/ml sucrose. For plasmid selection and maintenance in *Escherichia coli,* cultures were routinely grown in Luria broth (LB) or agar at 37°C supplemented with 100 μg/ml ampicillin or 25 μg/ml kanamycin.

Generating polyclonal antisera against the Csd5 SH3 domain and CcmA-

We generated clones of GST-CcmA (pJF-1480) and GST-Csd5-SH3 (pKB7) for recombinant expression and purification of CcmA (genomic position:1,607,196-1,607,510) and Csd5 (genomic position: 1,313,682-1,313,888) antigens respectively. Each protein was expressed in *E. coli* BL21 DE3, and purified on glutathione sepharose 4B (GE Healthcare) resin by gravity. For Csd5 but not CcmA, the GST tag was subsequently removed by Thrombin CleanCleave™ protease (Sigma) with the uncut fusion protein and cleaved GST tag removed in a subsequent purification step. The resulting protein preparations were submitted for antibody production (R & R Research LLC). Approximately 0.5 mg of protein antigen was used for primary immunization and 0.25 mg of antigen was used for 3 boosts prior to the final bleed. Resulting anti-sera were used for western blotting against *H. pylori* whole cell lysates from wild-type and mutant (Δ*ccmA* or Δ*csd5*) bacteria to ensure specific detection of CcmA or Csd5.

H. pylori quantitative cell shape analysis-

Phase-contrast microscopy was performed and resulting images were thresholded using the ImageJ software package. Quantitative analysis of thresholded images of ~200 bacteria per strain was used to measure both side curvature and central axis length with the CellTool software package as described previously (Sycuro *et al.*, 2010).

Immunoprecipitation and crosslinking of H. pylori cell shape proteins-

25 optical density at 600 nm units (OD_{600}) of *H. pylori* bacteria expressing functional FLAG or VSV-G fusions to cell shape proteins (Csd5, CcmA, MurF and AtpF) grown to mid-log phase (0.3-0.7 OD_{600}) were harvested by centrifugation for 10 minutes at 8000 x g at room temperature (RT). Fresh or frozen cell pellets were detergent solubilized in 1.9 ml of chilled Buffer A (20 mM Tris pH 8.0, 150 mM NaCl, 2% Glycerol, 1.0% Triton X-100) supplemented with 10 units of Benzonase nuclease (EMD) and EDTA-free protease inhibitors (Pierce). The cells were sonicated with a microtip (Sonic Dismembrator Model 500, Branson) for 5-10 seconds using short pulses at 10% power until visibly cleared, then centrifuged at 20,000 x g for 30 minutes at 4^{o}C. The soluble fraction was then incubated with 40 μl of equilibrated (with Buffer A) α-FLAG M2 (magnetic or agarose) or VSV-G mouse monoclonal agarose beads (Sigma) and incubated for 90 minutes at 4^{o}C. The beads are then subjected to 3 x 15 ml washes (room temp) with cold Buffer B (20 mM Tris pH 8.0, 150 mM NaCl, 2% Glycerol and 0.1% Triton X-100). Beads were then boiled in 2x bead volumes of protein sample buffer (66 mM Tris-HCl, pH 6.8 26% (w/v) glycerol, 2.1% SDS, 0.01% Bromophenol blue) without beta-mercaptoethanol and subjected to SDS-PAGE, silver staining (SilverQuest™, Invitrogen), western blotting and/or mass spectrometry analysis as appropriate. For crosslinking experiments, 25 OD's of whole cells were

incubated on ice with 1.0% formaldehyde in 2.0 ml of cold 1X phosphate buffered saline (PBS) pH7.2 for 30 minutes and then quenched with 0.83M glycine pH 2.2 for 10 minutes. Cells were harvested by centrifugation and frozen at -80°C prior to immunoprecipitation as described above.

Immunoblotting H. pylori whole-cell extracts and immunoprecipitates-

Whole cell extracts were prepared by harvesting 1.0 OD_{600} of log phase (0.3-0.7 OD_{600}) *H. pylori* by centrifugation for 2 minutes at max speed in a microcentrifuge and resuspending in 2x protein sample buffer at 10.0 OD_{600} per ml and boiled for 10 min. Whole cell extracts or input and IP fractions were separated on 4-15% gradient (BioRad TGX) gels by SDS-PAGE and transferred onto PVDF membranes using the BioRad Turbo-transfer system according to the manufacturer's instructions. Membranes were blocked for 1 hour at RT with 5% non-fat milk in TBS-T (0.5 M Tris,1.5 M NaCl, pH 7.6, plus 0.05% Tween 20). Membranes were incubated with primary antibody overnight at 4°C (or 1.5 hours at RT) with primary antibodies; 1:5000 dilution for α-FLAG M2 (Sigma), 1:5000 for α-VSV-G (Sigma), 1:5000 for α-SH3, 1:10,000 for α-CcmA, 1:10,000 for AtpD (Agrisera, Sweden), 1:20,000 dilution for α-Cag3 (Pinto-Santini and Salama, 2009), in TBS-T. Four 10 minute washes with TBS-T were followed by a 1.25 hour incubation at RT with appropriate horseradish peroxidase-conjugated α-immunoglobulin G (Santa Cruz Biotechnology) antibody at 1:20,000 dilution in TBS-T (α-rabbit for Cag3, Csd5, CcmA; α-mouse for FLAG and VSV-G). After four more TBS-T washes, antibody detection was performed with ECL Plus (Pierce) (Cag3, CcmA, CcmA-2xFLAG, SH3, AtpF-3xFLAG) or Immobilon (Millipore) (SH3, Csd5-2xFLAG, MurF-3xVSV-G, AtpF-3xFLAG, AtpD) detection kits, following the manufacturer's protocol and imaged directly on film or with the BioRad Gel Documentation System.

Mass spectrometry analysis of H. pylori immunoprecipitates-

Immunoprecipitated and washed α-VSV-G monoclonal agarose beads (Table 3.5 and Figure 3.5) were stored at -80°C following three washes with 20 mM ammonium bicarbonate and treated with 20 µg/ml trypsin for 5.25 hours at 37°C. For mass spectrometry, trypsinized bead suspensions were thawed and supernatant collected. The beads were washed twice with 20 µl of 20 mM ammonium bicarbonate and combined with the supernatant followed by reduction with 5mM TCEP (Pierce) for 1 hour at 37°C. Samples were next alkylated with 10 mM iodoacetamide for 30 minutes in the dark at RT. Reactions were quenched by addition of 16 mM N-acetyl-L-Cysteine (Sigma). Samples were then purified and concentrated on C18 reverse spin columns (Pierce) according to the manufacturers directions. Samples were quantified by Absorbance (A_{280}) and submitted to Northwestern Proteomics (Chicago, IL) for protein identification as previously described (Whitney *et al.*, 2015).

For α-FLAG IP's (Figure 3.7 and Figure 3.8), samples were run ~ 1cm into the wells of a 4-15% SDS-PAGE (Bio-Rad TGX) 10-well 50µl gels. After electrophoresis a 1cm gel slice containing the entire IP was excised and submitted for total protein identification to the Fred Hutch Proteomics Facility.

Gel slices were washed 15 minutes each with water, 50/50 acetonitrile/water, acetonitrile, 100 mM ammonium bicarbonate, followed by 50/50 acetonitrile/100 mM ammonium bicarbonate. The solution was removed and the gel slices were dried by vacuum centrifugation. The dried gel slices were reduced by covering them with 10 mM dithiothreitol in 100 mM ammonium bicarbonate and heating them at 56^{o}C for 45 minutes. The solution was removed and discarded. The gel slices were alkylated by covering them with a solution of 55 mM

iodoacetamide in 100 mM ammonium bicarbonate and incubating in the dark at ambient temperature for 30 minutes. The solution was removed and discarded. The gel slices were washed with 100 mM ammonium bicarbonate for 10 minutes on a shaker and an equal amount of acetonitrile was added and washing was continued for 10 minutes with shaking. The solution was removed, discarded, and the gel slices were dried by vacuum centrifugation. The gel slices were cooled on ice and an ice cold solution of 12.5 ng/μL trypsin (Promega, Madison, WI) in 100 mM ammonium bicarbonate was added, enough to cover the gel slice. After 45 minutes, the trypsin solution was removed, discarded, and a volume of 50 mM ammonium bicarbonate was added to cover the gel slices and they were incubated overnight (16 hours) at 37^{o}C with mixing on a shaker. Samples were spun down in a microfuge and the supernatant was collected. Peptides were extracted from the gel slices by adding 0.1% trifluoroacetic acid (TFA) to cover the slices and mixed at ambient temperature for 30 minutes. An equal volume of acetonitrile was added and the samples were mixed for an additional 30 minutes. The samples were spun in a microfuge, the digestion supernatant and the extraction supernatant were pooled, and the pool was taken to dryness by vacuum centrifugation. The trypsin digest was desalted using ZipTip C_{18} (Millipore, Billerica, MA) and eluted with 70% acetonitrile/0.1% TFA. The desalted material was dried by vacuum centrifugation.

The proteolytically-digested samples were brought up in 20 μL of 2% acetonitrile in 0.1% formic acid and 18 μL was analyzed by LC/ESI MS/MS with a Thermo Scientific Easy-nLC II (Thermo Scientific, Waltham, MA) coupled to a Orbitrap Elite ETD (Thermo Scientific, Waltham, MA) mass spectrometer using a trap-column configuration as described (Licklider *et al.*, 2002). In-line de-salting was accomplished using a 100 μm × 20 mm reversed-phase trap column packed with Magic C_{18}AQ (5-μm, 200 Å resin; Michrom Bioresources, Auburn, CA)

followed by peptide separations on a 75 μm × 250 mm reversed-phase column packed with Magic C_{18}AQ (5-μm, 100 Å resin; Michrom Bioresources, Auburn, CA) directly mounted on the electrospray ion source. A 65-minute gradient from 5% to 40% acetonitrile in 0.1% formic acid at a flow rate of 400 nL/min was used for chromatographic separations. A spray voltage of 2500 V was applied to the electrospray tip and the Orbitrap Elite instrument was operated in the data-dependent mode, switching automatically between MS survey scans in the Orbitrap (AGC target value 1,000,000, resolution 120,000, and injection time 250 milliseconds) and collision induced dissociation MS/MS spectra acquisition in the linear ion trap (AGC target value of 10,000 and injection time 100 milliseconds). The twenty most intense precursor ions from the OrbiTrap full scan were each consecutively selected for fragmentation in the linear ion trap by CID with a normalized collision energy of 35%. Selected ions were dynamically excluded for 30 seconds.

Data analysis was performed using Proteome Discoverer 1.4 (Thermo Scientific, Waltham, MA). The data were searched against an *H. pylori* database that was appended with protein sequences from the common repository of adventitious proteins (cRAP; www.thegpm.org/crap/). Searches were conducted with the trypsin enzyme specificity. The precursor ion tolerance was set to 10 ppm and the fragment ion tolerance was set to 0.6 Da. Variable modifications were set for oxidation on methionine (+15.995 Da) and carbamidomethyl (+57.021 Da) on cysteine. All search results were evaluated by Percolator for false discovery rate (FDR) evaluation of the identified peptides. Peptide identifications were filtered to a peptide FDR of 1%.

Purification of H. pylori peptidoglycan cell wall sacculi-

Isolation of *H. pylori* cell wall sacculi was performed as described (Glauner, 1988). Briefly, 330 ml of *H. pylori* cells in shaken liquid culture were grown to 0.6-0.8 OD/ml. Cells were harvested

by centrifugation in a TOMY TX-160 centrifuge with a TMA-27 rotor at 5000 rpm for 10 minutes at 4°C and subsequently resuspended in 6 ml chilled Dulbecco's PBS (Gibco). The cell suspension was added dropwise to 6 ml boiling 8% SDS and boiled for a further 6.5 hours. Sacculi were collected the first time by ultracentrifugation in a Beckman-Coulter Optima L-90K ultracentrifuge with a SW41 Ti rotor at 28000 rpm for 60 minutes at 28°C. Sacculi were resuspended in 2.5 ml 4% SDS and boiled for four hours. Sacculi were again collected by ultracentrifugation. This and all further ultracentrifugation steps were performed in a Beckman TL-100 ultracentrifuge with a TLA 100.3 rotor at 70000 rpm for 60 minutes at 25°C. Sacculi were resuspended in 2.5 ml 1% SDS, and boiled for four hours. We then collected sacculi by ultracentrifugation, resuspended in 2.5 ml of 50mM sodium phosphate pH 7.4. We incubated sacculi at 37°C with 0.25 mg α-chymotrypsin (Sigma, C4129) for four hours, added an additional 0.25 mg α-chymotrypsin, and incubated overnight. Following incubation, we added 300 µl of 10% SDS and boiled samples for four hours. Sacculi were collected by ultracentrifugation, resuspended in 2.5 ml 1% SDS, and boiled for four hours. We then washed sacculi by repeated ultracentrifugation and resuspension in ddH_2O until supernatant was demonstrated free from SDS by the Hayashi test (Hayashi, 1975). Purified SDS-free sacculi were resuspended in 500 µl ddH_2O plus 0.02% sodium azide and stored at 4°C until use.

Purification of recombinant GST-SH3 domain variants-

Plasmids containing N-terminal GST fusions to wild-type (pKB7) and mutant SH3 (pKM4-R146A, and pKM5-T151A) were transformed into *E. coli* expression host BL21 DE3. Strains were grown overnight at 37°C in the presence of LB with 0.2% glucose and 100 µg/ml ampicillin. The next day, cells were diluted 1/100 into fresh media without glucose, grown to

mid-log (~0.5 – 0.8 OD) at 37ºC, chilled on ice for 30 minutes, then induction for protein expression by addition of 1.0mM IPTG. Flasks were transferred to 16ºC and incubated overnight with shaking. Cells were harvested by centrifugation and flash frozen in liquid nitrogen or used immediately. Cells were suspended in 1/10 culture volume of lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10 mM $MgCl_2$, 5% Glycerol), supplemented with 10U of Benzonase nuclease (EMD) and 1mM PMSF (Fisher) and sonicated at 20% power using 4 x 30 second pulses on ice (1 minute rest on ice between each pulse). Lysates were then cleared at 20,000 x g at 4ºC. The cleared lysate was applied to equilibrated glutathione sepharose 4B affinity resin (25 µl slurry/ml culture volume) and incubated for 2 hours at RT with mixing. The protein bound GST resin was washed with 3 x 10 column volumes and proteins were eluted (5 x 1 column volume) from the beads using 20mM glutathione in lysis buffer. Fractions were analyzed by SDS-PAGE for purity and yield (Figure 3.12). For some experiments, the SH3 domain was separated from the GST domain by thrombin cleavage as described above for antigen preparation.

Construction of H. pylori epitope-tagged strains-

Strains containing single or multi-copy FLAG or VSV-G epitope gene fusions were generated by PCR SOEing (Horton, 1995) to target integration at the native locus after natural transformation (Clayton and Mobley, 2010). All strains and plasmids were validated by diagnostic PCR, Sanger sequencing and western blotting as appropriate. Strains TRH4 (Csd5-VSV-G), JTH2 (CcmA-2xFLAG) and JTH1 (Csd5-2xFLAG) were generated by PCR SOEing (Horton, 1995) to add either a single VSV-G or a 2xFLAG epitope tag and a stop codon upstream of a chloramphenicol resistance (cat^R) cassette for recombination at the native locus. JTH3 was generated by natural transformation of JTH2 with a PCR product containing a kanamycin resistance cassette with flanking homology to the cat cassette for conversion to kanamycin resistance. For strain TRH4,

H. pylori LSH100 genomic DNA containing a cat^R cassette was used as PCR template to individually amplify the following genomic regions using oligos THP1/THP2 (*csd5-VSV-G UP*), 828/829 (cat^R) and THP3/THP4 (*csd5-VSV-G DN*). For strains JTCF (CcmA-2xFLAG) and JT5F (Csd5-2xFLAG) the same genomic DNA was used to amplify the following regions using oligos JTP6/JTP7 (*ccmA-2xFLAG UP*), JTP5/829 (*2xFLAG-*cat^R) JTP8/JTP9 (*ccmA-2xFLAG DN*), JTP1/JTP2 (*csd5-2xFLAG* UP), JTP5/829 (*2xFLAG-*cat^R), and JTP3/JTP4 (*csd5-2xFLAG DN*). Resulting PCR products (*up*, cat^R, *down*) were gel extracted (Qiagen), stitched together by PCR SOEing, gel extracted and used directly for natural transformation into *H. pylori* strain LSH100. Strain KBH126 was constructed by natural transformation of strain JTH1 with genomic DNA containing *ccmA::kan*. Transformants were selected on HBA plates containing chloramphenicol or kanamycin and validated by diagnostic PCR, Sanger sequencing, and quantitative cell shape analysis (Figure 3.1, Figure 3.3 and Figure 3.4)

Strain MH9 was generated by PCR SOEing to introduce a *catR* cassette into the *slt* gene using purified PCR products generated by oligos MLP1/MLP2 (*slt UP*), 828/829 (cat^R) MLP3/MLP4 (*slt DN*) and recombination at the native locus. Strain KBH89 (*rdxA::slt-VSV-G*) was created by natural transformation of MHH9 (*slt::*cat^R) with plasmid pKB59 (*rdxA::slt-VSV-G*) and selection on HBA plates with metronidazole. To generate pKB59, plasmid pKB57 was digested with SpeI and XhoI (NEB) and the insert fragment purified. The vector pLC292 (Terry *et al.*, 2005) was then digested with SpeI and XhoI and the backbone gel purified. The purified vector backbone and insert fragments were ligated together using standard T4 ligase-based methodology. To generate plasmid pKB57, the *slt* gene was PCR amplified using G27 genomic DNA with oligos 1938/1939 to append a single VSV-G (*slt-VSV-G*) epitope tag and a STOP

codon. This PCR product was purified and TOPO cloned. Transformants were validated by diagnostic PCR, Sanger sequencing, and quantitative cell shape analysis (Figure 3.5).

Strain KBH103 (VSV-G-MinD) was created by allelic replacement of *minD::kansacB* in KBH91 using counter selection (Copass *et al.*, 1997) with purified donor plasmid pKB66 to generate a markerless *VSV-G-minD* integrated at the native locus. KBH91 was generated by natural transformation of LSH100 with pKB61 (*minD::kansacB*). Plasmid pKB61 was generated by PCR SOEing to stitch PCR products generated from LSH100 genomic DNA and oligos 1954/1955 (*minD::kansacB UP*), 1947/1948 (*kansacB*) and 1956/1957 (*minD::kansacB DN*). Plasmid pKB66 was generated using oligos 1970/1971 (*VSV-G-minD UP*) and 1972/1973 (*VSV-G-minD DN*) to append an N-terminal *VSV-G* epitope tag to *minD*. PCR products were stitched together by PCR SOEing, purified, and TOPO cloned. Transformants were validated by diagnostic PCR, Sanger sequencing, and quantitative cell shape analysis (Fig. S3).

Strains KBH127KBH129 and KBH131-KBH133 were created by allelic replacement of *csd5::catsacB* in LSH36 (Sycuro *et al.*, 2010) using counter selection with purified donor plasmids pKB100-102 and pKB104-106 respectively, to generate markerless Csd5 domain deletion variants (Fig. 1) containing a C-terminal 2xFLAG epitope tag integrated at the native locus. To generate plasmid pKB80 JTH1 genomic DNA was used to PCR amplify *csd5-2xFLAG:cat*^{R} using oligos 51/JTP4. This PCR product was purified and TOPO cloned. To generate plasmid pKB100, pKB80 was used as template for PCR with oligos 51/JTP2 (*csd5-2xFLAG UP*) and LSH100 genomic DNA template was used with oligos 2178/JTP4 (*csd5-2xFLAG DN*) to generate PCR products for PCR SOEing. This product was purified and TOPO cloned. To construct pKB101-103 and pKB104-106, pKB100 was used as PCR template to create *csd5*-2xFLAG* domain deletion (*) alleles. We used PCR SOEing to create *csd5* domain

deletions variants by stitching together *csd5*-up* and *2xFLAG-down* PCR products containing identical overlapping central ends that omit the targeted DNA sequences using the following oligo pairs; 51/2185 and 2186/JTP4 (pKB101), 51/2187 and 2188/JTP4 (pKB102), 51/2189 and 2190/JTP4 (pKB104), 51/2183 and 2184/JTP4 (pKB105), 51/2179 and 2180/JTP4 (pKB106). Each paired PCR product was stitched together, individually purified, and TOPO cloned. Transformants were validated by diagnostic PCR, Sanger sequencing, and quantitative cell shape analysis (Figure 3.3).

Strains KBH157 (*csd5-2xFLAG, murF-3xVSV-G*) and KBH159 (*atpF-3xFLAG*) were constructed by natural transformation of KBH127 and LSH100 with pKB113 and pKB116 respectively with transformants selected on HBA plates containing chloramphenicol. Plasmids pKB113 (*murF-3xVSV-G*) and (*atpF-3xFLAG*) were generated by PCR SOEing to add either a 3xVSV-G or 3xFLAG epitope tag and a stop codon upstream of a chloramphenicol resistance (cat^R) cassette for recombination at the native locus. LSH100 genomic DNA with a cat^R cassette was used as DNA template to amplify the following genomic regions with oligos 2216/2251 (*murF-3xVSV-G* UP), 2250/829 (*VSV-G-*cat^R), 2209/2211 (*murF-VSV-G DN*), 2199/2271 (*atpF-3xFLAG UP*), 2248/829 (*3xFLAG::*cat^R), and 2275/2276 (*murF-3xFLAG DN*). Matched *murF*, *atpF*, and cat^R PCR products were gel purified, stitched together, and TOPO cloned. Transformants were validated by diagnostic PCR, Sanger sequencing, and quantitative cell shape analysis (Figure 3.4).

Construction of Csd5 SH3 domain mutants and recombinant protein expression strains-

Strains KBH17 (Csd5-WT), KMH1 (Csd5-T151A), and KMH2 (Csd5-R146A) were created by allelic replacement of *csd5::catsacB* in LSH36 using counter selection (Copass *et al.*, 1997) with

purified donor plasmids pKB21 (*csd5-WT*), pKM1 (*csd5-R146A*), and pKM2 (*csd5-T151A*). To generate pKB21, *H. pylori* G27 DNA was used to PCR amplify the *csd5* gene region (266bp up and 320 bp down) using oligos 1461/1462. The PCR product was purified and TOPO (Invitrogen) cloned. Plasmids pKM1 and pKM2 were generated using a modified quick-change procedure (Stratagene) for site-directed mutagenesis of plasmid pKB21 using oligos 1506 (pKM1 Csd5 R146A) and 1507 (pKM2 Csd5 T151A). For constructing Csd5-SH3 domain recombinant protein expression plasmids pKB7 and pKM5, we used vectors pKB21 and pKM2 respectively as DNA template with oligos 1448/1449 to generate PCR products with BamHI and XhoI sites appended to the wild-type and mutant alleles of *csd5*. pGEX4T3 was linearized with BamHI and XhoI and gel purified. Each PCR product was digested with BamHI and XhoI, purified, and cloned into the linearized pGEX4T3 backbone using standard T4 ligase based cloning.

3.7 BIBLIOGRAPHY

Arnold, I., Pfeiffer, K., Neupert, W., Stuart, R.A., and Schägger, H. (1998) Yeast mitochondrial F1F0-ATP synthase exists as a dimer: identification of three dimer-specific subunits. *EMBO J* **17**: 7170–7178.

Ausmees, N., Kuhn, J.R., and Jacobs-Wagner, C. (2003) The bacterial cytoskeleton: an intermediate filament-like function in cell shape. *Cell* **115**: 705–713.

Baba, T., and Schneewind, O. (1996) Target cell specificity of a bacteriocin molecule: a C-terminal signal directs lysostaphin to the cell wall of *Staphylococcus aureus*. *EMBO J* **15**: 4789–4797.

Bartlett, T.M., Bratton, B.P., Duvshani, A., Miguel, A., Sheng, Y., Martin, N.R., *et al.* (2017) A periplasmic polymer curves *Vibrio cholerae* and promotes pathogenesis. *Cell* **168**: 172–185.e15.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., *et al.* (2000) The Protein Data Bank. *Nucleic Acids Research* **28**: 235–242.

Bonis, M., Ecobichon, C., Guadagnini, S., Prevost, M.C., and Boneca, I.G. (2010) A M23B family metallopeptidase of *Helicobacter pylori* required for cell shape, pole formation and virulence. *Molecular Microbiology* **78**: 809–819.

Brown, C.T., Hug, L.A., Thomas, B.C., Sharon, I., Castelle, C.J., Singh, A., *et al.* (2015) Unusual biology across a group comprising more than 15% of domain bacteria. *Nature* **523**: 208–211.

Cabeen, M.T., Charbon, G., Vollmer, W., Born, P., Ausmees, N., Weibel, D.B., and Jacobs-Wagner, C. (2009) Bacterial cell curvature through mechanical control of cell growth. *EMBO J* **28**: 1208–1219.

Chan, A.C.K., Blair, K.M., Liu, Y., Frirdich, E., Gaynor, E.C., Tanner, M.E., *et al.* (2015) Helical shape of *Helicobacter pylori* requires an atypical glutamine as a zinc ligand in the carboxypeptidase Csd4. *J Biol Chem* **290**: 3622–3638.

Chaput, C., Labigne, A., and Boneca, I.G. (2006) Characterization of *Helicobacter pylori* lytic transglycosylases Slt and MltD. *J Bacteriol* **189**: 422–429.

Clayton, C.L., and Mobley, H.L.T. (2010) *Helicobacter pylori Protocols*. Humana Press.

Davies, K.M., Anselmi, C., Wittig, I., Faraldo-Gómez, J.D., and Kühlbrandt, W. (2012) Structure of the yeast F1FO-ATP synthase dimer and its role in shaping the mitochondrial cristae. *Proc Natl Acad Sci USA* **109**: 13602–13607.

Eppinger, M., Baar, C., Linz, B., Raddatz, G., Lanz, C., Keller, H., *et al.* (2006) Who ate whom? Adaptive *Helicobacter* genomic changes that accompanied a host jump from early humans to large felines. *PLoS Genet* **2**: e120.

Favini-Stabile, S., Contreras-Martel, C., Thielens, N., and Dessen, A. (2013) MreB and MurG as scaffolds for the cytoplasmic steps of peptidoglycan biosynthesis. *Environmental Microbiology* **15**: 3218–3228.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., *et al.* (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* **136**: E359–86.

Finn, R.D., Clements, J., Arndt, W., Miller, B.L., Wheeler, T.J., Schreiber, F., *et al.* (2015) HMMER web server: 2015 update. *Nucleic Acids Research* **43**: W30–8.

Glauner, B. (1988) Separation and quantification of muropeptides with high-performance liquid chromatography. *Anal Biochem* **172**: 451–464.

Gründling, A., and Schneewind, O. (2006) Cross-linked peptidoglycan mediates lysostaphin binding to the cell wall envelope of *Staphylococcus aureus*. *J Bacteriol* **188**: 2463–2472.

Hahn, A., Parey, K., Bublitz, M., Mills, D.J., Zickermann, V., Vonck, J., *et al.* (2016) Structure of a complete ATP synthase dimer reveals the molecular basis of inner mitochondrial membrane morphology. *Molecular Cell* **63**: 445–456.

Hayashi, K. (1975) A rapid determination of sodium dodecyl sulfate with methylene blue. *Anal Biochem* **67**: 503–506.

Horton, R.M. (1995) PCR-mediated recombination and mutagenesis. SOEing together tailor-made genes. *Mol Biotechnol* **3**: 93–99.

Höltje, J.V. (1998) Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*. *Microbiol Mol Biol Rev* **62**: 181–203.

Kamitori, S., and Yoshida, H. (2015) Structure-function relationship of bacterial SH3 domains. In *SH Domains*. Springer International Publishing, Cham. pp. 71–89.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845–858.

Kim, H.S., Kim, J., Im, H.N., An, D.R., Lee, M., Hesek, D., *et al.* (2014) Structural basis for the recognition of muramyltripeptide by *Helicobacter pylori* Csd4, a D,L-carboxypeptidase controlling the helical cell shape. *Acta Crystallogr Sect D Biol Crystallogr* **70**: 2800–2812.

Koch, M.K., McHugh, C.A., and Hoiczyk, E. (2011) BacM, an N-terminally processed bactofilin of *Myxococcus xanthus*, is crucial for proper cell shape. *Molecular Microbiology* **80**: 1031–1051.

Kühn, J., Briegel, A., Mörschel, E., Kahnt, J., Leser, K., Wick, S., *et al.* (2010) Bactofilins, a ubiquitous class of cytoskeletal proteins mediating polar localization of a cell wall synthase in *Caulobacter crescentus*. *EMBO J* **29**: 327–339.

Laddomada, F., Miyachiro, M.M., and Dessen, A. (2016) Structural insights into protein-protein interactions involved in bacterial cell wall biogenesis. *Antibiotics (Basel)* **5**.

Lee, E.-J., Pontes, M.H., and Groisman, E.A. (2013) A bacterial virulence protein promotes pathogenicity by inhibiting the bacterium's own F1Fo ATP synthase. *Cell* **154**: 146–156.

Leon, E., Navarro-Aviles, G., Santiveri, C.M., Flores-Flores, C., Rico, M., Gonzalez, C., *et al.* (2010) A bacterial antirepressor with SH3 domain topology mimics operator DNA in sequestering the repressor DNA recognition helix. *Nucleic Acids Research* **38**: 5226–5241.

Lu, J.Z., Fujiwara, T., Komatsuzawa, H., Sugai, M., and Sakon, J. (2006) Cell wall-targeting domain of glycylglycine endopeptidase distinguishes among peptidoglycan cross-bridges. *Journal of Biological Chemistry* **281**: 549–558.

McGowan, C.C., Cover, T.L., and Blaser, M.J. (1997) Analysis of F1F0-ATPase from *Helicobacter pylori*. *Infect Immun* **65**: 2640–2647.

Mohammadi, T., Karczmarek, A., Crouvoisier, M., Bouhss, A., Mengin-Lecreulx, D., and Blaauwen, den, T. (2007) The essential peptidoglycan glycosyltransferase MurG forms a complex with proteins involved in lateral envelope growth as well as with proteins involved in cell division in *Escherichia coli*. *Molecular Microbiology* **65**: 1106–1121.

Nelson, W.C., and Stegen, J.C. (2015) The reduced genomes of Parcubacteria (OD1) contain signatures of a symbiotic lifestyle. *Front Microbiol* **6**: 713.

Noble, M.E., Musacchio, A., Saraste, M., Courtneidge, S.A., and Wierenga, R.K. (1993) Crystal structure of the SH3 domain in human Fyn; comparison of the three-dimensional structures of SH3 domains in tyrosine kinases and spectrin. *EMBO J* **12**: 2617–2624.

Paumard, P., Vaillier, J., Coulary, B., Schaeffer, J., Soubannier, V., Mueller, D.M., *et al.* (2002) The ATP synthase is involved in generating mitochondrial cristae morphology. *EMBO J* **21**: 221–230.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605–1612.

Pinto-Santini, D.M., and Salama, N.R. (2009) Cag3 is a novel essential component of the *Helicobacter pylori* Cag type IV secretion system Oouter membrane subcomplex. *J Bacteriol* **191**: 7343–7352.

Pohl, E., Holmes, R.K., and Hol, W.G. (1999) Crystal structure of a cobalt-activated diphtheria toxin repressor-DNA complex reveals a metal-binding SH3-like domain. *Journal of Molecular Biology* **292**: 653–667.

Rexroth, S., Meyer zu Tittingdorf, J.M.W., Schwaßmann, H.J., Krause, F., Seelert, H., and Dencher, N.A. (2004) Dimeric H+-ATP synthase in the chloroplast of *Chlamydomonas reinhardtii*. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1658**: 202–211.

Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* **42**: W320–4.

Rolain, T., Bernard, E., Beaussart, A., Degand, H., Courtin, P., Egge-Jacobsen, W., *et al.* (2013) O-glycosylation as a novel control mechanism of peptidoglycan hydrolase activity. *J Biol Chem* **288**: 22233–22247.

Rühle, T., and Leister, D. (2015) Biochimica et Biophysica Acta. *BBA - Bioenergetics* **1847**: 849–860.

Scheffers, D.-J., and Tol, M.B. (2015) LipidII: Just another brick in the wall? *PLoS Pathog* **11**: e1005213–12.

Seelert, H., and Dencher, N.A. (2011) ATP synthase superassemblies in animals and plants: two or more are better. *Biochim Biophys Acta* **1807**: 1185–1197.

Sham, L.-T., Butler, E.K., Lebar, M.D., Kahne, D., Bernhardt, T.G., and Ruiz, N. (2014) Bacterial cell wall. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. *Science* **345**: 220–222.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539.

Smith, C.A. (2006) Structure, function and dynamics in the mur family of bacterial cell wall ligases. *Journal of Molecular Biology* **362**: 640–655.

Sycuro, L.K., Pincus, Z., Gutierrez, K.D., Biboy, J., Stern, C.A., Vollmer, W., and Salama, N.R. (2010) Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori's* helical shape and stomach colonization. *Cell* **141**: 822–833.

Sycuro, L.K., Rule, C.S., Petersen, T.W., Wyckoff, T.J., Sessler, T., Nagarkar, D.B., *et al.* (2013) Flow cytometry-based enrichment for cell shape mutants identifies multiple genes that influence *Helicobacter pylori* morphology. *Molecular Microbiology* **90**: 869–883.

Sycuro, L.K., Wyckoff, T.J., Biboy, J., Born, P., Pincus, Z., Vollmer, W., and Salama, N.R. (2012) Multiple peptidoglycan modification networks modulate *Helicobacter pylori's* cell shape, motility, and colonization potential. *PLoS Pathog* **8**: e1002603.

Typas, A., Banzhaf, M., Gross, C.A., and Vollmer, W. (2012) From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Micro* **10**: 123–136.

Vecchiarelli, A.G., Li, M., Mizuuchi, M., Hwang, L.C., Seol, Y., Neuman, K.C., and Mizuuchi, K. (2016) Membrane-bound MinDE complex acts as a toggle switch that drives Min oscillation coupled to cytoplasmic depletion of MinD. *Proc Natl Acad Sci USA* **113**: E1479–88.

Vollmer, W., and Bertsche, U. (2008) Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1778**: 1714–1734.

Wattam, A.R., Davis, J.J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., *et al.* (2017)

Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Research* **45**: D535–D542.

Weber, J. (2006) ATP synthase: subunit-subunit interactions in the stator stalk. *Biochim Biophys Acta* **1757**: 1162–1170.

White, C.L., Kitich, A., and Gober, J.W. (2010) Positioning cell wall synthetic complexes by the bacterial morphogenetic proteins MreB and MreD. *Molecular Microbiology* **76**: 616–633.

Whitney, J.C., Quentin, D., Sawai, S., LeRoux, M., Harding, B.N., Ledvina, H.E., *et al.* (2015) An interbacterial NAD(P)+ glycohydrolase toxin requires elongation factor Tu for delivery to target cells. *Cell* **163**: 607–619.

Wilkinson, M., Chaban, Y., and Wigley, D.B. (2016) Mechanism for nuclease regulation in RecBCD. *eLife* **5**: 43.

Xu, Q., Mengin-Lecreulx, D., Liu, X.W., Patin, D., Farr, C.L., Grant, J.C., *et al.* (2015) Insights into substrate specificity of NlpC/P60 cell wall hydrolases containing bacterial SH3 domains. *mBio* **6**: e02327–14

Xu, Q., Sudek, S., McMullan, D., Miller, M.D., Geierstanger, B., Jones, D.H., *et al.* (2009) Structural basis of murein peptide specificity of a gamma-D-glutamyl-l-diamino acid endopeptidase. *Structure* **17**: 303–313.

Chapter 4. CSD5 PROMOTES CELL SHAPE VARIATION DURING CHRONIC INFECTION AND HOST ADAPTATION

4.1 PREFACE

The results presented in this section represent the contributions of students from the Fred Hutch Summer Undergraduate Research (SURP) and High School Internship Programs (SHIP): Max Ruben, Jesse Domingo, Alex Shang, and Daniel Wright. Additional contributors: Rick Peek Jr., Laura Sycuro, Ilana Cohen, Jenny Taylor Tina Gall and Laura Jackson.

4.2 INTRODUCTION

The helical cell shape of *H. pylori* is required for efficient colonization of the gastric niche in a murine model, but the role of helical shape during long-term chronic infection is not clear (Sycuro *et al.*, 2010). Unrelated clinical isolates of *H. pylori* display unique helical cell shape profiles that may vary in axis-length, helical-pitch and wavelength. When grown as clonal populations, they also exhibit population-level heterogeneity that may include the presence of non-helical variants (e.g. straight-rod) within an otherwise helical population (Figure 4.1). Thus we hypothesized that cell shape heterogeneity may play a role in gastric niche acquisition and that specific cell shapes might preferentially favor colonization of distinct regions of the stomach.

To investigate this question we turned to a unique collection of *H. pylori* isolates that were collected at two time points over six years during chronic infection of a human host. *H. pylori* strain J99 (ATCC 700824) was isolated from a gastric antral biopsy sweep from an

American Caucasian male presenting with a duodenal ulcer (1994) at the Vanderbilt Medical Center in Nashville, TN (Peek et al, 1995). The source patient declined antibiotic treatment but returned to the clinic six years later (2000) with persistent dyspepsia and individual gastric biopsies were taken from the duodenum, antrum, corpus, and cardia. A total of 30 single colony isolates were obtained from each region and herein are referred to as the J99 recent isolates (J99-Ri) (Israel et al 2001).

The J99-Ri strain collection constitutes a unique set of strains in which to study the role for phenotypic variation of cell shape of *H. pylori* during infection. We show that during chronic infection *H. pylori* has an enormous capacity to alter its shape with an unexpected diversity of unique cell shapes among the isolates including vibroids, short helices, long filaments, and highly curved c-shaped cells. Using directed sequencing of known cell shape determining (*csd*) genes, we identified a subset of isolates that have accumulated nonsense mutations in the peptidoglycan hydrolases *csd4* and *csd6* that result in straight-rod cells in addition to amino acid changes in the PG binding SH3 domain of Csd5 that disrupt protein expression. Homology modeling of J99-Csd5 suggests that Csd5 is related to NlpC/P60 hydrolases involved in Gram-positive cell wall recycling and we demonstrate a preference of the *H. pylori* SH3 domain for tetrapeptide enriched PG substrate. We present evidence of complex genetic regulation of cell shape and putative PG recycling programs to control helical cell shape.

4.3 RESULTS

J99 Recent Isolate strains exhibit diverse cell morphologies and genetic variation in cell shape gene *csd4*, *csd5* and *csd6*-

We investigated cell shape of a subset of the J99-Ri strains for morphological variation by phase contrast microscopy in broth grown cultures and found significant cell shape heterogeneity in the isolates that include helical/vibroid (J99-Ancestral), straight-rods (C3, C12, D3), highly-curved vibrio's (A3, C7), long-filamentous helical cells (C1, C9, C11), and "normal" helical (C2, Ca-1, D1) cell shapes (Figure 4.1).

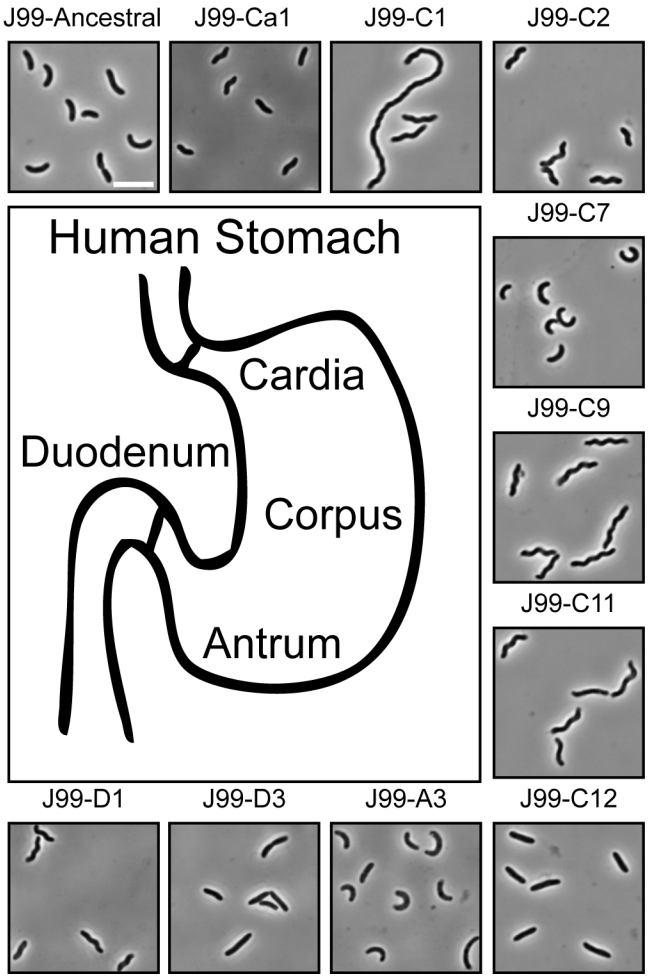

Figure 4.1. Diverse cell shapes of J99 Recent isolates during chronic infection.

Phase-contrast microscopy of the J99 Ancestral and J99-Recent isolate strains showing representative morphologies of single colony isolates obtained from biopsies of 4 different regions of the human stomach (Israel 2000). Antrum (A) Cardia (Ca), Corpus (C), Duodenum (D), number indicates single colony isolate clone ID. Scale-bar = 5 μm

Targeted Sanger sequencing of known cell shape genes *csd1*, *csd2* and *ccmA* revealed no mutations. However, we did identify mutations in *csd4*, *csd5,* and *csd6*. In the case of the *csd4* and *csd6* genes, frameshift mutations in either would abolish expression of the enzyme and its PG hydrolase activity. In each case, the straight-rod cell shape phenotype of these *csd4* and *csd6* mutants is consistent with loss of function phenotypes of these proteins (Figure 4.2).

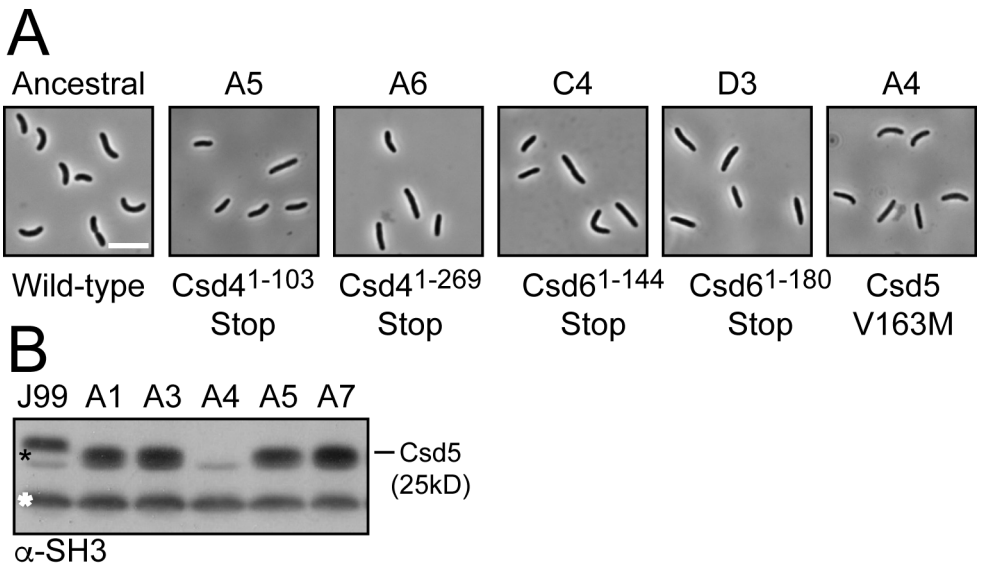

Figure 4.2. Cell shape of J99 strains with mutations in cell shape genes.

A. Phase-contrast microscopy of representative J99-Ri strains containing nonsense (Csd4 and Csd6), point mutations (A4-Csd5-V163M), and in-frame deletions (Csd5-Δ24bp; A5, A6, C4, D3, A4) of known cell shape genes as indicated. Scale bar = 5 μm. B. Western blot of whole-cell extracts of the indicated of J99 Ancestral and J99-Ri strains using an α-Csd5-SH3 domain antibody for detection of Csd5. Black asterisk indicates Csd5 and the white asterisk indicates a cross-reacting non-specific band as a loading control proxy.

We also find that a subset of the J99 Ri's have identical *csd4* (A2, A10 and C12) or *csd6* (C3 and D3) mutations indicating that these particular isolates are more closely linked (Table 4.13). In the case of *csd5*, we found a single amino acid substitution (J99-A4-Csd5^{V163M}) adjacent to the

highly conserved arginine in the RT-loop of the C-terminal SH3 domain that in Chapter 3 (Figure 3.9 B) was demonstrated essential for protein expression and function. We show by western blotting that this Csd5-V163M mutation similarly abolishes expression in this strain.

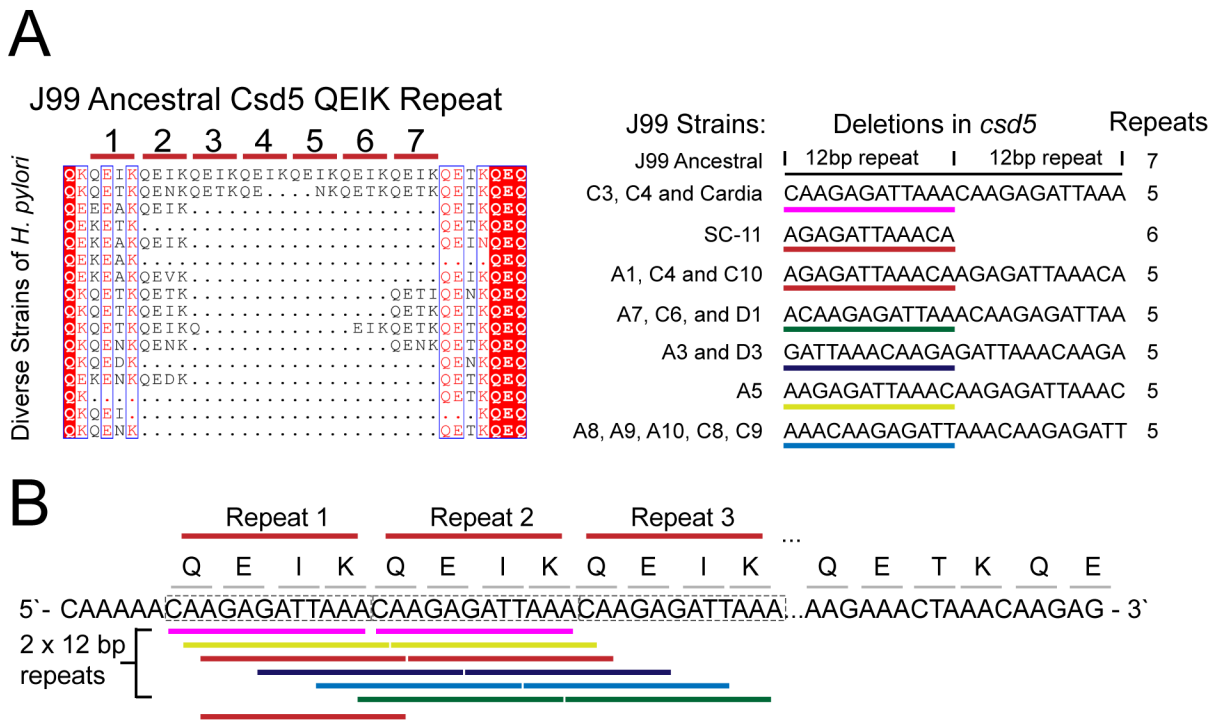

Figure 4.3. The Csd5 QEIK repetitive linker domain can be used to distinguish independent recombination events in the J99-Ri's.

A. Left: Sequence alignment of the J99 Csd5 variable linker domain (QEIK repeat, 7 copies) motif compared to diverse strains of *H. pylori*. Right: The underlined nucleotide sequences represent the either 12 or 24 bp DNA sequences that were deleted by recombination from the indicated strains. B. The DNA sequence of the J99 linker domain and flanking region showing possible sites of nucleotide excision described in panel A.

In addition to finding a missense mutation in Csd5, sequencing of this gene revealed evidence of recombination events that lead to shortening of a variable linker domain (Figure 4.3). In the J99 ancestral strain, the Csd5 linker domain consists of 7 consecutive QEIK sequence

motifs in the disordered domain. In all of the recent isolates tested by Sanger sequencing, the QEIK linker consists of 5 repeats. We queried the 12 single colony isolates from the original J99 antral sweep (1994) and all strains had the ancestral linker of 7 QEIK repeats except J99 SC11 which had 6 repeats (Table 4.13). The nature of the linker contraction is such that either 1 or 2 deletion events of 12 base pairs can occur along the repeat tract, though and it cannot be determined which repeat was lost. We can however group strains by the unique DNA sequence lost from the repeat (Figure 4.3 B). Our results indicate apparently independent deletion events in the infecting population. However, in all cases the consequences are such that no amino acid changes occur other than loss of either 4 or 8 amino acid residues. We also observe no deleterious effect from the 24 base pair deletion on the expression of the Csd5 (Figure 4.3 B) and our work in Chapter 3 demonstrated that this linker domain is not essential for helical shape but that cell shape parameters are subtly affected (Figure 3.3).

Cell wall curvature is influenced by *csd5* genetic variation-

To determine if Csd5 sequence variation is sufficient to alter the intrinsic cell shape of *H. pylori* strains, we complemented Δ*csd5* knockouts in the J99 and G27 strain backgrounds with the non-cognate *csd5* allele at the native locus. We also tested the J99-*csd5*-*QEIK*^{5} variant side-by-side against the J99-*csd5*-*QEIK*^{7} variants in the G27 background and assessed cell wall side curvature. With respect to the wild-type variants, we found that cell shape of the complemented strain, was more consistent with the parental cell shape phenotype of the donated allele (Figure 4.4). In other words, G27 bacteria complemented with the J99 allele looked more like J99 (reduced side curvature), while J99 bacteria complemented with the G27 allele looked more like G27 (increased side-curvature). We did not observe any obvious difference in strains with QEIK^{7}

vs QEIK[5] variant alleles in the G27 background, suggesting that the linker adaptation is either insignificant, subtle, or strain background specific.

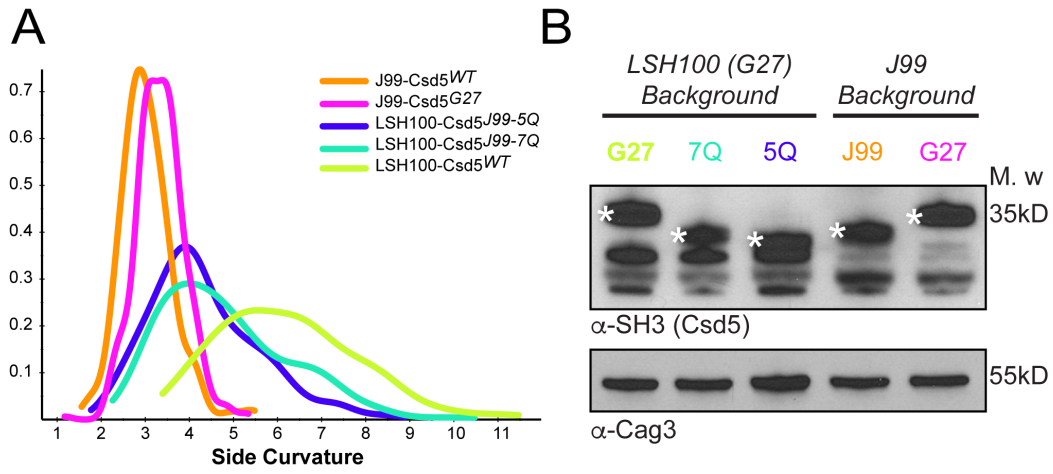

Figure 4.4. Csd5 protein variants alter helical cell shape parameters in different backgrounds. A. Quantitative cell shape analysis using CellTool (Sycuro *et al.*, 2010); smooth histogram showing cell population side-curvature (x-axis) as a density function of strains J99 (WT), J99-Csd5^{G27} (KBH62), LSH100-Csd5$^{J99\text{-}5Q}$ (KBH144), LSH100-Csd5$^{J99\text{-}7Q}$ (KBH34). B. Western blots of *H. pylori* whole-cell extracts of strains using an α-SH3 domain antibody to detect Csd5 and α-Cag3 antibody (Pinto-Santini and Salama, 2009) as a loading control. White asterisks indicate the Csd5 variant protein.

Csd5 is related to NlpC-P60 endopeptidases that recognize free stem-peptides-

Attempts at homology modeling full-length Csd5 from *H. pylori* strain G27 and other strain backgrounds was hampered by the fact that the middle domain of Csd5 is predicted to be disordered and may only fold into it's native structure in a cellular context with associated membranes and interaction partners present. Furthermore, Csd5 is thought to be unique to *H. pylori* making it unlikely that homologs exist in the RCSB database for structural comparisons. Despite this, homology modeling using Phyre2 with the J99 ancestral sequence (jhp1171)

produced a structure in which the Csd5 disordered domain was folded into a structure similar to the overall fold of the C-terminal SH3 domain (Figure 4.5).

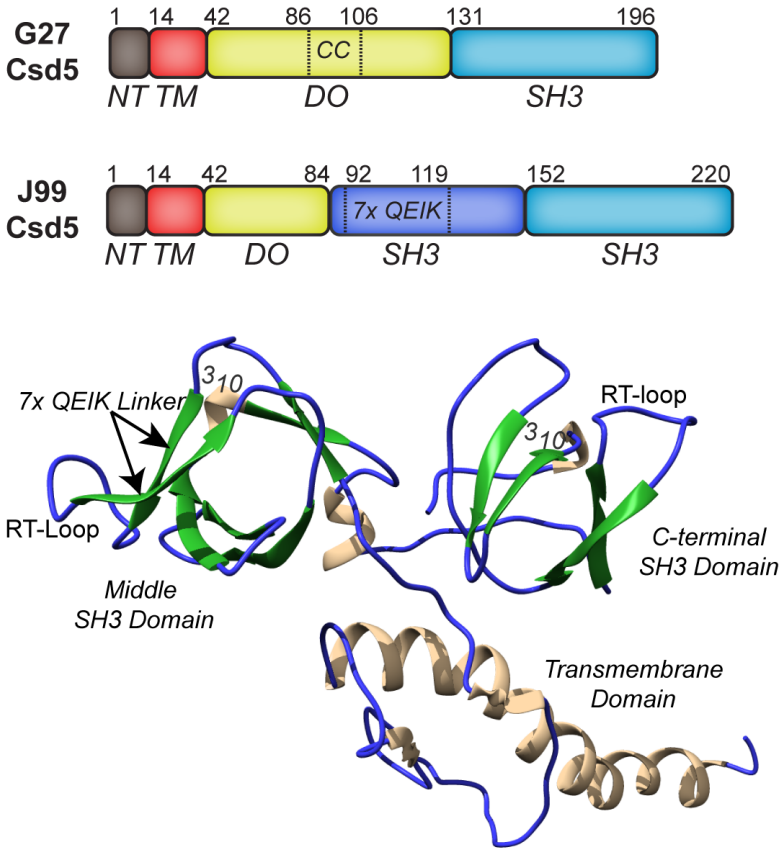

Figure 4.5. Structural modeling of J99 Csd5 indicates the middle domain is a second SH3 domain.

Top: Schematic of Csd5 domain organization from G27 and J99. NT (N-terminal), TM transmembrane), DO (Disordered). Bottom: Phyre2 homology model of J99-Csd5 with β-sheets colored green and loops colored blue. The functional RT-loop and 3_{10} helices of the respective SH3 domains are shown along with the QEIK[7] repeat domain in the first SH3 domain. Image produce using Chimera software.

This model was submitted to the DALI server (Holm and Laakso, 2016) to identify similar structures, returning hits consistent with the homologs used by Phyre2 to build the model (Table 4.11). The top two hits include proteins of the NlpC/P60 family of cell wall lysins from

Clostridium perfringens phage phiSM101 (Tamai *et al.*, 2014) and from *Bacteroides thetaiotaomicron* (Xu *et al.*, 2015) a prominent member of the gut microbiome. These proteins are characterized by tandem SH3 domains and a C-terminal NlpC/P60 cysteine hydrolase domain active on peptidoglycan (Anantharaman and Aravind, 2003). That Csd5 is related to these proteins suggests that Csd5 must have lost the catalytic domain. Interestingly, our proteomic experiments indicate a putative protein interaction between Csd5 and HPG27_80 (Table 5.14), an NlpC/P60 family protein with tandem SH3 domains and associated C-terminal hydrolase domain. However, no cell shape phenotype has been observed for this gene (data not shown) and the putative interaction has not been validated.

Table 4.11. Top 10 DALI hits with the J99 Csd5 homology model.

Rank	%ID	Z Score	PDB	Description	Reference
1	13	19.5	4KRT	*Clostridium perfringens* phage phiSM101 Endolysin	Kamitori, S 2014
2	27	8.4	2KRS	*Clostridium perfringens* Probable Enterotoxin	none
3	5	8.3	4R0K	*Bacteroides thetaiotaomicron* putative dipeptidyl-peptidase VI	none
4	7	8.2	3PVQ	*Bacteroides thetaiotaomicron* putative dipeptidyl-peptidase VI	Xu, 2015
5	3	8.0	3NPF	*Bacteroides ovatus* putative dipeptidyl-peptidase VI	Xu, 2015
6	6	8.0	4Q2W	*Streptococcus pneumonia* Peptidoglycan hydrolase LytB	Bai XH et al 2014
7	5	7.4	2HBW	*Anabaena variabilis* Nlp/P60 Protein	Xu, Q 2009
8	6	7.3	2FG0	*Nostoc punctiforme* gamma-d-glutamyl-l-diamino acid endopeptidase	Xu, Q 2009
9	2	7.2	3H41	*Bacillus cereus* NlpC/P60 Protein	Xu, Q 2009
10	30	7.1	2KT8	*Clostridium perfringens* Probable Surface Protein	none

In Chapter 3 we demonstrated that the Csd5 C-terminal SH3 domain is sufficient for direct binding to peptidoglycan (Figure 3.9). Structural studies of the *Bacteroides thetaiotaomicron*

YkfC (BtYkfC) also indicates a role in the binding of free stem peptides with a preference for tetrapeptides (Xu *et al.*, 2015). In a PG sedimentation assay we show that the SH3 domain exhibits a preference for purified Δ*csd6* mutant *H. pylori* cell wall sacculi (Figure 4.6), which have increased abundance of PG tetrapeptides in their cell walls relative to wild-type (Sycuro *et al.*, 2013) Δ*csd1* (Sycuro *et al.*, 2010), or Δ*csd4* (Sycuro *et al.*, 2012) mutants. YkfC is reported to be involved in peptidoglycan recycling, a process that breaks down PG fragments for import into the cytoplasm and eventual incorporation back into the PG biosynthetic pathway (Park and Uehara, 2008). In these particular proteins the role for the tandem SH3 domains is proposed to define the substrate specificity. In *Bacillus cereus, ykfC* is in the same operon as the *oppA* (oligopeptide ABC transporter) gene (Xu *et al.*, 2010), an arrangement similarly shared by *csd5* in *H. pylori* (Figure 4.7) suggesting a possible role for the *H. pylori* SH3 domains as PG sensors with differing affinities for uncross-linked tetra-, tri- and di- stem peptides.

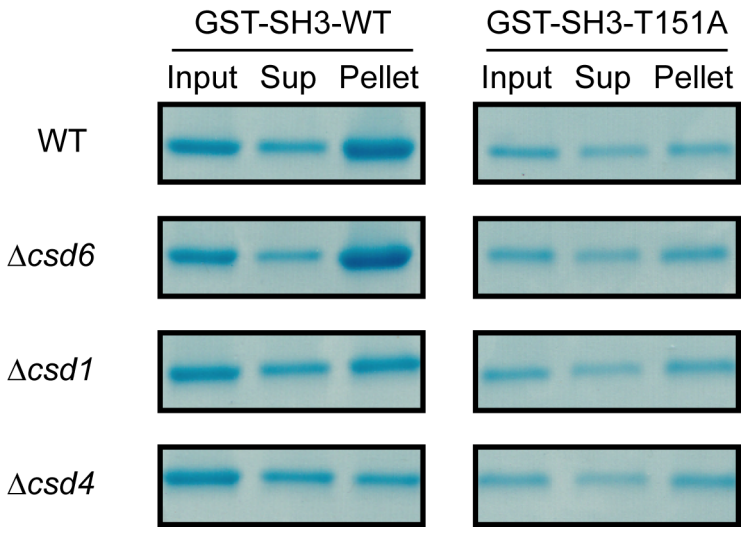

Figure 4.6. The Csd5 SH3 domain preferentially binds to tetrapeptide (Δ*csd6*) enriched PG.

SDS-PAGE of a PG sedimentation assay using purified wild-type (LSH100) and mutant PG sacculi in the presence of wild-type (GST-SH3-WT) or mutant (GST-SH3-T151A) protein as a negative control, stained with Coomasie Blue.

Our model of Csd5^{J99} indicates that the QEIK repeat motifs comprise the RT-loop of the first SH3 domain indicating a reduction in the length of this loop. Our Csd5 domain deletion analysis (Figure 3.3) in Chapter 3 indicates that loss of this region can affect the overall shape parameters of the cell but is not strictly required for generating helical cells. Taken together we conclude that the first SH3 domain of Csd5 may play a role in modulating the affinity or specificity of the second SH3 domain for PG binding to regulate helical cell shape parameters.

Complex regulation and phase variation of the *oppAB* and *csd5* polycistronic operon influences cell shape parameters.

Csd5 is encoded in a polycistronic operon downstream of the oligopeptide transporter genes *oppAB* that together with *oppCD* encoded elsewhere on the chromosome, constitutes one of two putative PG recycling systems in *H. pylori* (Weinberg and Maier, 2007). The second PG recycling program consists of the dipeptide transport system *dppABCDF*. In Gram-negative bacteria these ABC-transporter like systems transport small peptides, including PG degradation products from the periplasm into the cytoplasm for reuse. Transcriptional profiling of gene expression for this locus (Figure 4.7) indicates complex regulation and a variety of alternative transcripts with different 5' start sites (Sharma *et al.*, 2010). Furthermore, the *oppAB* locus is subject to frame-shift mutations that appear to arise from slipped strand mispairing events that result in the conversion of a functional gene product to a non-functional pseudo gene for either *oppA* or in the *H. pylori* PMSS1 to SS1 strain background (Draper *et al.*, 2017). Evidence for this

gene conversion event was also seen during mouse adaptation of *H. pylori* strain G27 to LSH100 (Cohen, 2014) (Figure 4.7) and will be discussed in a subsequent section.

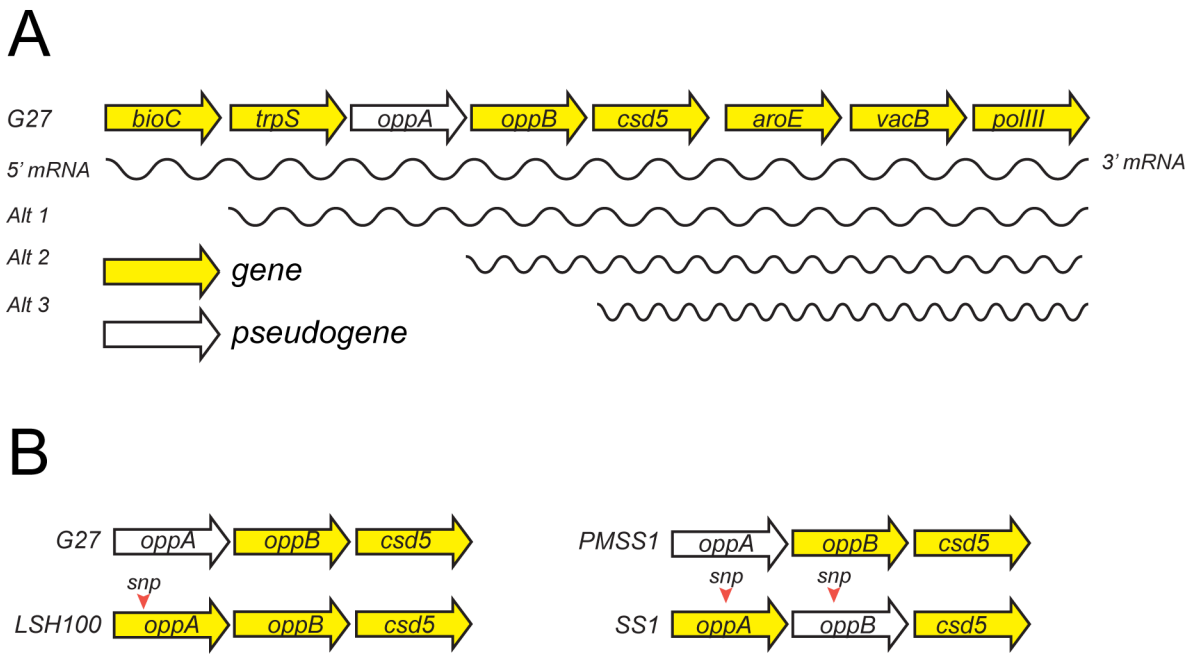

Figure 4.7. The *oppAB* and *csd5* genes are encoded by multiple alternate transcripts with phase variable genes.

A. Yellow colored arrows indicate expressed genes while the white arrow indicates a gene disrupted by phase variation in the G27 strain background. Alternate mRNA transcripts measured by (Sharma *et al.*, 2010) are shown as squiggly lines. B. Pseudogene status of the *oppA* and *oppB* genes before (G27 and PMSS1) and after (LSH100 and SS1, respectively) mouse adaptation.

Overexpression of OppB decreases cell length-

To determine if the putative PG oligopeptide transporter components encoded upstream of *csd5* are involved in cell shape determination we overexpressed *oppB* by placing a second copy of the gene under control of the *rdxA* promoter at the ectopic *rdxA* locus. These cells

displayed a decrease in axis length but no significant change to the side-curvature profile of the cells (Figure 4.8)

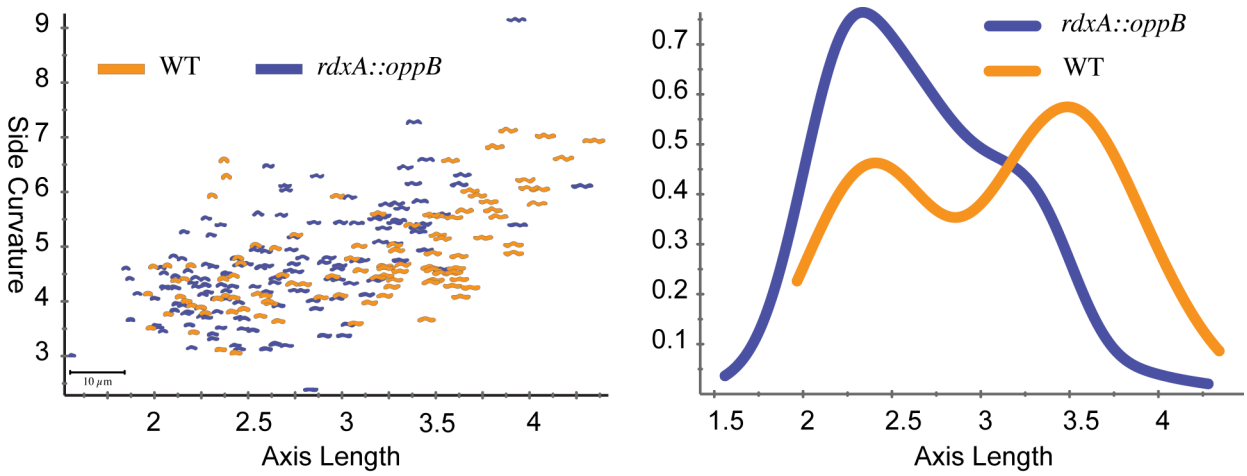

Figure 4.8. Overexpression of *oppB* decreases cell axis length.

Quantitativecell shape analysis using CellTool software (Sycuro *et al.*, 2010). Left: scatter plot showing cell population side-curvature vs. axis-length. Right: smooth histogram showing cell population axis length (x-axis) as a density function (50-100 cell/strain).

Csd5 overexpression mediates cell shape changes during host adaptation-

Previous work has shown that during host adaptation in a mouse model of colonization, a serially passaged isolate of *H. pylori* (LSH100) exhibited cell shape changes manifested as an increase in cell wall side-curvature (Sycuro, 2010). We performed a western blot on whole cell extracts of LSH100 (mouse adapted) and the G27 (parental) and find that Csd5 protein levels are elevated. Furthermore, an examination of Csd5 protein levels in *H. pylori* strain SS1 and its parental strain PMSS1 also exhibit changes in cell shape (change in helical pitch) and an increase in Csd5 protein levels after passaging through mice. To determine if an increase in *csd5* gene expression is sufficient to drive a cell shape change, we generated a series of strain backgrounds with one or two copies of *csd5* at either the native locus, the *rdxA* locus or both in the G27 background.

Increasing copy number of *csd5* increased Csd5 protein expression and there was a concordant increase in cell wall curvature suggesting that the increase in Csd5 protein in the mouse adapted strain is due to the increase in Csd5 protein (Figure 4.9).

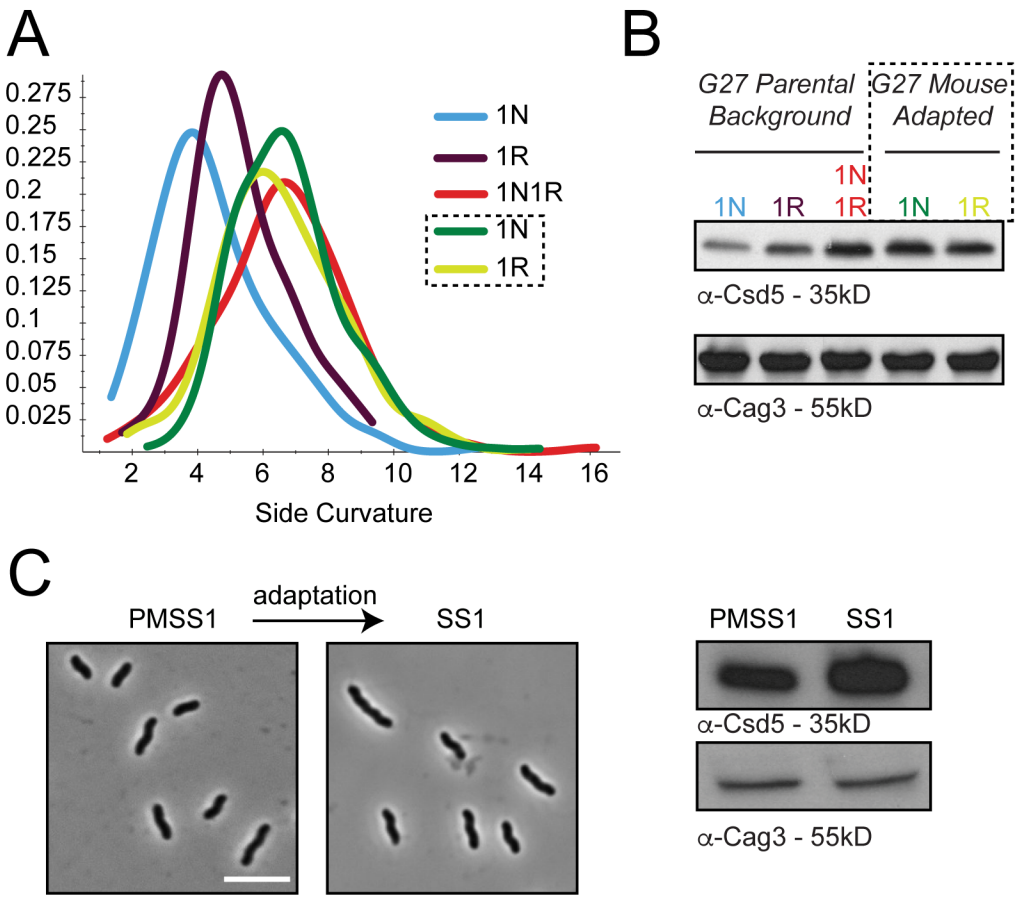

Figure 4.9. Csd5 overexpression increases cell wall curvature.

A. Quantitative cell shape analysis using the CellTool program (Sycuro *et al.*, 2010). Smooth histogram showing cell-population side-cuvature (x-axis) as a density function for *H. pylori* strains expressing *csd5* in single copy at the native locus (1N), oectopic *rdxA* locus (1R) or in 2 copies (1N1R) in parental (G27) and LSH100 genetic backgrounds, as indicated. G27 Background (1N=G27, 1R= LSH42 1N1R=KBH18; LSH100 Background (1N= LSH100, 1R= LSH43). B. Western blot of whole-cell extracts of the indicated strains using α-SH3 domain to detect Csd5. α-Cag3 was used as a loading control. C. Phase-contrast microscopy (left) and western blot analysis of *H. pylori* strains PMSS1 and SS1 using whole cell extracts.

4.4 DISCUSSION

The helical cell shape of *H. pylori* is important for the early stages of host colonization but whether helical shape is required for the maintenance of a chronic infection has not been explored. The J99 Ri strains suggest accumulation of extensive cell shape variation during chronic infection beyond that which we observe in clonal populations passaged in the laboratory. Thus, specifics cell shapes may have been selected for that provide survival benefits or facilitate niche acquisition during long-term colonization of the human stomach. For the isolates we studied in this collection we cannot know for how long they were co-existing in the host, but the isolates collected in year 2000 exhibit much greater variation in their shape phenotypes than among those isolates from the first time point in 1994, suggesting that these variants were selected over time. Interestingly, we observe a subset of strains with a straight-rod phenotype presumably caused by the mutations in the genes for *csd4, csd5* and *csd6.* However, we do not know if this loss of helical shape was selected by use in niche acquisition, or are merely variants that could be lost over time. We are also unable to complement the mutations due to a loss of efficient genetic transformation of J99ri variants.

Our structure-function analysis studies of Csd5 have shown that the C-terminal SH3 domain binds peptidoglycan. Our modeling studies of Csd5 variants resulted in the discovery that Csd5 is related to NlpC/P60 proteins that recognize and hydrolyze free PG stem peptides, similar to the substrates for Csd4 and Csd6. More work is needed to characterize the specific PG ligand of the *H. pylori* SH3 domain(s). Furthermore, these proteins are closely related to human gut bacterium *Bacteroides thetaiotaomicron* indicating a possible host environment from which *H. pylori* acquired this gene long ago. Interestingly, the Csd4 C-terminal domains (of unknown function) also shares remote homology to proteins from this same bacterium (Chan *et al.*, 2015).

YkfC NlpC/P60 homologs are encoded in the opp (oligopeptide permease) and dpp (dipeptide permease) operons of *Bacillus cereus* and *B. subtilis* respectively, further evidence of a shared evolutionary function in PG recycling or recognition. In *H. pylori* this operon and its regulation exhibit potential to affect cell shape as under and over expression of cell shape genes and *oppB* alter cell shape parameters.

4.5 SUPPORTING INFORMATION

Table 4.12. *H. pylori* strains used in this study

Strains and Plasmids	**Relevant or Descriptive Genotype**	**Reference or Source**
***H. pylori* Strains**		
G27	Wild-type	
LSH100	Wild-type	
J99	Wild-type	
LSH36	*csd5::catsacB* (LSH100 Background)	
KBH34	*csd5-QEIK*^{7} (LSH100 Background)	This work
KBH61	*csd5::catsacB* (J99 Background)	This work
KBH62	*csd5*^{G27} (J99 Background)	This work
KBH144	*Csd5QEIK*^{5} *(LSH100 Background)*	This work

Table 4.13. J99 ancestral and recent isolate cell shapes and related genotypes

Strain ID /Source	Observed Shape	Verified genotypes and mutations in known cell shape genes
J99 Ancestral Single Colony Isolates - Gastric Antral Biopsy (1994)		
J99 Ancestral	Low Pitch Helix	*csd5WT*
SC1	Curved Low Pitch Helix	*csd5WT*
SC2	Not determined	*csd5WT*
SC3	Not determined	*csd5WT*
SC4	Helical Short (Ca-1)	*csd5WT*
SC5	Curved Low Pitch Helix	*csd5WT*
SC6	Not determined	*csd5WT*
SC7	Not determined	*csd5WT*
SC8	Not determined	*csd5WT*

SC9	Low Pitch Helix	*csd5WT*
SC10	Low Pitch Helix	*csd5WT*
SC11	Low Pitch Helix	*csd5Δ12bp (QEIK linker deletion)*
SC12	Low Pitch Helix	*csd5WT*
J99-Ri's- Corpus and Cardia (2000)		
Ca-1	Helical Short	*csd5Δ24bp (2xQEIK linker deletion)*
C-1	Long Helix Filamentous	
C-2	Typical helix	
C-3	Straight, slight curve	*csd5Δ24bp, csd6+1 FS pos:508751*
C-4	Straight, slight curve	*csd5Δ24bp, csd6+2 FS pos:508561*
C-5	Long Helix Filamentous	
C-6	Long Helix	*csd5Δ24bp*
C-7	Hyper curved	
C-8	Typical helix	*csd5Δ24bp*
C-9	Long Helix Filamentous	*csd5Δ24bp*
C-10	Long Helix Intermediate	*csd5Δ24bp*
C-11	Long Helix Filamentous	
C-12	Straight	*csd4+1FS pos:381904*
J99-Ri's- Antrum (2000)		
A1	Long Helix	*csd5Δ24bp*
A2	Straight	*csd4+1FS pos:381904*
A3	Curved	*csd5Δ24bp*
A4	Straight	*csd5Δ24bp-V163M*
A5	Straight	*csd5Δ24bp, csd4+1FS pos:381326*
A6	Straight	*csd4+1FS pos:381850*
A7	Long Helix Filamentous	*csd5Δ24bp*
A8	Straight	*csd5Δ24bp csd4Y424STOP*
A9	Long Helix Filamentous	*csd5Δ24bp*
A10	Presumed Straight	*csd5Δ24bp csd4 +1FS pos:381904*
A11	Long Helix Filamentous	
A12	Straight	*csd6 +1FS pos:508580*
J99-Ri's-Duodenum (2000)		
D1	Helical Short	*csd5Δ24bp*
D2	Long Helix	
D3	Straight	*csd5Δ24bp, csd6+1 FS pos:508751*
D4		
D5	Helical	*csd4+1 FS pos:381176*

4.6 BIBLIOGRAPHY

Anantharaman, V., and Aravind, L. (2003) Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes. *Genome Biol* **4**: R11.

Chan, A.C.K., Blair, K.M., Liu, Y., Frirdich, E., Gaynor, E.C., Tanner, M.E., *et al.* (2015) Helical shape of *Helicobacter pylori* requires an atypical glutamine as a zinc ligand in the carboxypeptidase Csd4. *J Biol Chem* **290**: 3622–3638.

Cohen, I. (2014). Genetic basis of host adaptation by the human stomach pathogen *Helicobacter pylori*. University of Washington.

Draper, J.L., Hansen, L.M., Bernick, D.L., Abedrabbo, S., Underwood, J.G., Kong, N., *et al.* (2017) Fallacy of the unique genome: Sequence diversity within single *Helicobacter pylori* strains. *mBio* **8**: e02321–16.

Holm, L., and Laakso, L.M. (2016) Dali server update. *Nucleic Acids Research* **44**: W351–W355.

Park, J.T., and Uehara, T. (2008) How bacteria consume their own exoskeletons (turnover and recycling of cell wall peptidoglycan. *Microbiol Mol Biol Rev* **72**: 211–27

Pinto-Santini, D.M., and Salama, N.R. (2009) Cag3 is a novel essential component of the *Helicobacter pylori* Cag type IV secretion system outer membrane subcomplex. *J Bacteriol* **191**: 7343–7352.

Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiß, S., Sittka, A., *et al.* (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464**: 250–255.

Sycuro, L. (2009). The genetic determinants of *Helicobacter pylori's* helical cell shape and their role in virulence. University of Washington.

Sycuro, L.K., Pincus, Z., Gutierrez, K.D., Biboy, J., Stern, C.A., Vollmer, W., and Salama, N.R. (2010) Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori's* helical shape and stomach colonization. *Cell* **141**: 822–833.

Sycuro, L.K., Rule, C.S., Petersen, T.W., Wyckoff, T.J., Sessler, T., Nagarkar, D.B., *et al.* (2013) Flow cytometry-based enrichment for cell shape mutants identifies multiple genes that influence *Helicobacter pylori* morphology. *Molecular Microbiology* **90**: 869–883.

Sycuro, L.K., Wyckoff, T.J., Biboy, J., Born, P., Pincus, Z., Vollmer, W., and Salama, N.R. (2012) Multiple peptidoglycan modification networks modulate *Helicobacter pylori's* cell shape, motility, and colonization potential. *PLoS Pathog* **8**: e1002603.

Tamai, E., Yoshida, H., Sekiya, H., Nariya, H., Miyata, S., Okabe, A., *et al.* (2014) X-ray structure of a novel endolysin encoded by episomal phage phiSM101 of *Clostridium perfringens*. *Molecular Microbiology* **92**: 326–337.

Weinberg, M.V., and Maier, R.J. (2007) Peptide Transport in *Helicobacter pylori*: Roles of Dpp and Opp systems and evidence for additional peptide transporters. *J Bacteriol* **189**: 3392–3402.

Xu, Q., Abdubek, P., Astakhova, T., Axelrod, H.L., Bakolitsa, C., Cai, X., *et al.* (2010) Structure of the γ-D-glutamyl-L-diamino acid endopeptidase YkfC from *Bacillus cereus* in complex with L-Ala-γ-D-Glu: insights into substrate recognition by NlpC/P60 cysteine peptidases. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **66**: 1354–1364.

Xu, Q., Mengin-Lecreulx, D., Liu, X.W., Patin, D., Farr, C.L., Grant, J.C., *et al.* (2015) Insights into substrate specificity of NlpC/P60 cell wall hydrolases containing bacterial SH3 domains. *mBio* **6**: e02327–14.

Chapter 5. A PATH FORWARD: FURTHER PROTEOMIC ANALYSIS OF AN *H. PYLORI* SHAPEOSOME COMPLEX

5.1 INTRODUCTION

In the preceding chapters I have a charted a path forward towards a more holistic view of the helical cell shape program in *H. pylori* having identified and characterized a helical cell shape promoting complex. I presented a structure-function analysis of Csd5, have identified and validated protein interactions between cell shape determining proteins (Csd5 and CcmA), a cytoplasmic peptidoglycan pre-cursor synthase (MurF), the ATP synthase complex, and ultimately to the peptidoglycan cell wall. Our primary conclusion from these findings is that Csd5 if part one more multi-protein cell shape determining complexes. Further supporting evidence for the existence of multi-protein "Shapeosome" complexes is found in Table 5.14.

5.1 RESULTS AND DISCUSSION

Candidate Csd5 Interactions- Despite having already characterized several robust interactions with Csd5 (MurF, CcmA and ATP synthase), I routinely observe other proteins (HPG27_80) that are highly enriched in our proteomic screens from immunoprecipitation of Csd5 and CcmA. It should be noted that these studies are really just preliminary in nature and many experimental approaches could and should be pursued to more carefully characterize these interactions in more detail. Specifically, genetic approaches and buffer optimization should be used to determine if sub-complexes exist and if the strength of the observed interactions can be improved. As an example, by performing an IP of Csd5-2xFLAG in Δ*ccmA* knockout background we observed a

10-fold increase in the presence of a gene of unknown of function, HPG27_590. These data suggest that CcmA may be inhibiting this interaction in vivo and highlights how little we know the mechanism of these shape complexes.

Candidate CcmA Interactions- In the IP-Mass spec data presented below the most compelling candidate interaction with CcmA is with the DnaK chaperone protein. This protein is highly enriched when IP's are separate by SDS-PAGE and silver stained. We have positively identified this band by excision from an SDS-PAGE gel followed by mass spectrometry to identify peptides. In addition, we see evidence for interactions with MurF and also with the ATP synthase. Again, IP experiments using a Δ*csd5* knockout background hinted at an increase in detection of ATP synthase components. These experiments suggest the possible existence of separate complexes and that Csd5 and CcmA are each able to interact with the same proteins, just not at the same time. Much more work would be needed to explore this hypothesis and others generated by these data. This may also includes looking for evidence of ATP synthase dimers in the curved membranes of the *H. pylori*.

Table 5.14. Non validated shapeosome protein interaction candidates based on immunoprecipitation of cell shape proteins and mass spectrometry.

Bait/ Gene ID	**Putative Interactor**	**Description of Protein Function**	**Exp ID**
Csd5-2xFLAG IP			
HPG27_80	HPG27_80	Putative gamma-D-glutamyl-L-diamino acid endopeptidase *Desulfovibrio vulgaris* NlpC P60 like protein	PC-KB1191
HPG27_80	HPG27_80	Putative gamma-D-glutamyl-L-diamino acid endopeptidase *Desulfovibrio vulgaris* NlpC P60 like protein	PC-KB1142
HPG27_1379	Lpp20	Membrane associated lipoprotein; Lpp20 Domain	PC-KB1191

HPG27_1298	Prc	Carboxy-terminal protease; Csd cell shape protein	PC-KB1142
HPG27_80	HPG27_80	Putative gamma-D-glutamyl-L-diamino acid endopeptidase *Desulfovibrio vulgaris* NlpC P60 like protein	PC-KB1098G2
HPG27_119	HPG27_119	Hypothetical protein; putative Csd from shape screen	PC-KB1098G2
HPG27_30	ClpA	ATP-dependent C1p protease; ATP binding subunit	PC-KB1098G2
HPG27_1099	MurG	UDP-N-acetylglucosamine lipid transferase	PC-KB1098G2
HPG27_265	FtsH-1	Cell division protein	PC-KB1098G2
HPG27_1153	Wzk	Putative flippase; ABC-transporter, ATP-binding	PC-KB1098G2
HPG27_280	DppD	Dipeptide ABC transporter; ATP Binding Protein	PC-KB1098G2
HPG27_610	MurA	UDP-N-acetylglucosamine enolpyruvyl transferase	PC-KB1098G2
HPG27_281	DppF	Dipeptide ABC transporter; ATP Binding Protein	PC-KB1098G2
HPG27_1379	Lpp20	membrane-associated lipoprotein	PC-KB1098G2
HPG27_1482	Csd2	ToxR-activated protein; Cell Shape Protein	PC-KB1098G2
HPG27_1481	Csd1	ToxR-activated protein; Cell Shape Protein	PC-KB1098G2
Csd5-2xFLAG Δ*ccmA* IP			
HPG27_590	HPG27_590	Unknown DUF262 and DUF 1524 Domains; HNH nuclease	PC-KB1191
Multiple	ATP Synthase	Multiple Components of ATP synthase	PC-KB1191
CcmA-2xFLAG IP			
Multiple	ATP Synthase	Multiple Components of ATP synthase	PC-KB1191
HPG27_696	MurF	MurF; MurNac pentapeptide D-Ala D-Ala ligase	PC-KB1191
HPG27_101	DnaK	Heat Shock Protein 70	PC-KB1191
multiple	multiple	ATP Synthase	PC-KB1098
HPG27_747	CadA	Cadmium-transporting ATPase	PC-KB1098
HPG27_1099	MurG	UDP-N-acetylglucosamine lipid transferase	PC-KB1098
HPG27_590	_590	Unknown DUF262 and DUF 1524 Domains; HNH nuclease	PC-KB1098
HPG27_101	DnaK	Heat Shock Protein 70	PC-KB1064
CcmA-2xFLAG Δ*csd5* IP			
Multiple	ATP Synthase	Multiple Components of ATP synthase	PC-KB1191

VITA

Kris Blair grew up in central Indiana and attended Indiana University in Bloomington. He earned a BS in Microbiology in 1998 and eventually joined the IU department of Biology as a Research Associate and Lab Manager for nearly seven years in the laboratories of Dr. Clay Fuqua and Dr. Daniel Kearns. Pursing his passion for science and microbial genetics, Kris then attended the University of Washington beginning in 2011, as a student in the Molecular and Cellular Biology Interdisciplinary PhD program where he was mentored in Nina Salama's lab at the Fred Hutchinson Cancer Research Center. During his training in Nina's lab, Kris earned a National Science Foundation Graduate Research Fellowship to support his efforts to better understand the molecular mechanisms and protein interactions that control helical cell shape in *Helicobacter pylori.*