University of Kentucky
UKnowledge

Theses and Dissertations--Microbiology, Immunology, and Molecular Genetics

Microbiology, Immunology, and Molecular Genetics

2023

SUBSTRATE SELECTION AND EFFECTOR GATING IN THE *HELICOBACTER PYLORI* CAG TYPE IV SECRETION SYSTEM

Mackenzie E. Ryan
University of Kentucky, mackenzie.ryan@uky.edu
Digital Object Identifier: https://doi.org/10.13023/etd.2023.417

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation

Ryan, Mackenzie E., "SUBSTRATE SELECTION AND EFFECTOR GATING IN THE *HELICOBACTER PYLORI* CAG TYPE IV SECRETION SYSTEM" (2023). *Theses and Dissertations--Microbiology, Immunology, and Molecular Genetics*. 29.
https://uknowledge.uky.edu/microbio_etds/29

This Doctoral Dissertation is brought to you for free and open access by the Microbiology, Immunology, and Molecular Genetics at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Microbiology, Immunology, and Molecular Genetics by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Mackenzie E. Ryan, Student

Dr. Carrie Shaffer, Major Professor

Dr. Brett Spear, Director of Graduate Studies

SUBSTRATE SELECTION AND EFFECTOR GATING IN THE *HELICOBACTER PYLORI CAG* TYPE IV SECRETION SYSTEM

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Medicine at the University of Kentucky

By

Mackenzie Ryan

Lexington, Kentucky

Co- Directors: Dr. Carrie Shaffer, Assistant Professor of Veterinary Science and Microbiology, Immunology, and Molecular Genetics

and Dr. Kenneth Fields, Professor of Microbiology, Immunology, and Molecular Genetics

Lexington, Kentucky

2023

Copyright © Mackenzie Ryan 2023

ABSTRACT OF DISSERTATION

SUBSTRATE SELECTION AND EFFECTOR GATING IN THE *HELICOBACTER PYLORI CAG* TYPE IV SECRETION SYSTEM

Helicobacter pylori is a recognized carcinogen and gastric colonization by strains that harbor the *cag* type IV secretion system (1) is the strongest known risk factor for stomach malignancy. Gastric adenocarcinoma is the fourth leading cause of cancer-related deaths world-wide (>700,000 deaths annually), with *H. pylori* directly contributing to the development of more than one million new cases of cancer per year accounting for 5.5% of all malignancies. *H. pylori* exploits *cag* T4SS activity to alter the mucosal microenvironment by delivering diverse immunostimulatory cargo into target gastric epithelial cells. Currently, the mechanisms by which the *cag* T4SS transports substrates across the bacterial envelope are undefined.

The work presented here provides mechanistic insight into *H. pylori cag* T4SS structure and function. Using techniques to monitor uncontrolled effector release, I show that the pilin ortholog CagC plays a critical role in regulating cargo release to the bacterial cell surface, suggesting that CagC forms a gating apparatus that governs substrate transport across the outer membrane. This work demonstrates that *cagC* is required for trans-kingdom DNA conjugation (2, 3)and provides evidence that translocated DNA is double-stranded and exposed to the extracellular milieu prior to entering host cells, supporting a two-step DNA secretion model. CagC interacts with components predicted to comprise the *cag* T4SS inner membrane-embedded apparatus, as well as CagX and CagY within the periplasmic ring complex (PRC). Structural analyses revealed intermolecular π-π interactions among aromatic and positively charged residues formed between adjacent CagX subunits within asymmetric regions of the PRC. I identified CagX as a novel DNA binding protein that strongly and preferentially interacts with dsDNA and demonstrate that CagX π-π stacking coordinates substrate selection and enables trans-kingdom DNA conjugation without disrupting translocation of protein and peptidoglycan effector molecules. Collectively, these studies suggest a model whereby architectural symmetry mismatch exposes CagX π-π interfaces within the PRC to facilitate DNA transit through the *cag* T4SS translocation channel.

Furthermore, this work identified several additional components that regulate effector release across the bacterial outer membrane, providing insight into substrate gating mechanisms in the *cag* T4SS translocation channel. Finally, I demonstrate that APEX2 proximity labeling approaches can be adapted to analyze *cag* T4SS architecture. These studies are expected to uncover the composition of novel apparatus complexes that assemble to deliver diverse effector molecules into gastric epithelial cells.

KEYWORDS: Trans-kingdom DNA translocation, effector substrate selection, *Helicobacter pylori,* gastric cancer, secretion system, architectural asymmetry

Mackenzie Ryan

(Name of Student)

08/04/2023

Date

SUBSTRATE SELECTION AND EFFECTOR GATING IN THE *HELICOBACTER PYLORI CAG* TYPE IV SECRETION SYSTEM

By
Mackenzie Ryan

Dr. Carrie Shaffer
Co-Director of Dissertation

Dr. Kenneth Fields
Co-Director of Dissertation

Dr. Brett Spear
Director of Graduate Studies

08/04/2023
Date

DEDICATION

To my family

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Dr. Carrie Shaffer. Carrie, thank you for your unwavering support over the past five years. You have challenged and motivated me throughout my academic journey. I have learned many invaluable lessons from you, both in and out of the lab, that I will carry with me as I continue my career. Your guidance and mentorship have shaped me into the scientist I am today, and for that, I am truly grateful.

I would also like to thank my advisory committee, Dr. Kenneth Fields, Dr. Brian Stevenson, Dr. Erin Garcia, and Dr. Konstantin Korotkov. My committee members have all been excellent, and I appreciate their insightful suggestions and encouragement.

Additionally, I owe a huge thank you to my lab mates, Prashant Damke and Lynn Leedhanachoke. Prashant, thank you for your guidance and advice that made this project possible. Lynn, thank you for being a constant source of conversation, advice, and support and for always being willing to proofread my emails. I also would like to thank my friends and fellow graduate students, Gabi, Chi, and Zaria. Thank you for always being a source of laughter, advice, and support throughout the years. I am lucky to have you all in my life, and I would not have made it through this journey without you.

Finally, I would like to thank my family. I would not be where I am today without your encouragement and belief in me. Thank you for your unwavering support. I could not have done it without you.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

LIST OF FIGURES vi

LIST OF FIGURES vii

CHAPTER 1. Introduction 1

1.1 *Helicobacter pylori* and Gastric Disease 1
1.1.1 Bacterial Risk Factors 2
1.1.2 Host Risk Factors 7
1.1.3 Environmental Risk Factors 8

1.2 Bacterial type IV secretions systems 9
1.2.1 Interbacterial DNA Delivery Systems 10
1.2.2 DNA Export Into The Extracellular Milieu 16
1.2.3 DNA Uptake By Competence Systems 20
1.2.4 Trans-kingdom DNA Transfer via Conjugative Mechanisms 23

1.3 Structural Architecture of T4SSs 27
1.3.1 The *H. pylori cag* T4SS 28
1.3.2 Architectural Asymmetry in T4SSs 31

1.4 Research Objective 33

CHAPTER 2. CagC is a Novel Gating Mechanism in the *cag* T4SS Apparatus 34

2.1 Summary 34

2.2 Introduction 34

2.3 Materials and Methods 36

2.4 Results 43

2.5 Discussion 56

CHAPTER 3. Architectural asymmetry enables DNA transport through the *cag* type IV secretion system 60

3.1 Summary 60

3.2 Introduction 60

3.3 Materials and Methods 62

3.4 Results 67

3.5 Discussion 75

CHAPTER 4. Role of Unique *cag* T4SS components in Effector Trafficking 78
4.1 Summary 78
4.2 Introduction 78
4.3 Material and Methods 80
4.4 Results 83
4.5 Discussion 90
CHAPTER 5. Summary and Future Directions 94
References 106
VITA 126

LIST OF TABLES

Table 2.1 Prototypical *A. tumefaciens vir* T4SS homologs in the *cag* T4SS..... 44

LIST OF FIGURES

Figure 1. Symmetry mismatch in 'minimized' and 'expanded' T4SS architectures. 29
Figure 2. *H. pylori* DNA is delivered to host cells via a novel secretion mechanism. 46
Figure 3. CagC gates the *cag* T4SS translocation channel. 49
Figure 4. CagC is a VirB2-like T4SS pilin ortholog. 52
Figure 5. Pentameric CagC assembles as an 'endopilus' in the *cag* T4SS. 55
Figure 6. CagX binds dsDNA in the PRC. 68
Figure 7. CagX-DNA interaction is unique to the *cag* T4SS. 69
Figure 8. π-π interactions within asymmetric PRC subunits mediate CagX-DNA binding 71
Figure 9. Symmetry mismatch coordinates DNA substrate selection for trans-kingdom conjugation. 73
Figure 10. Proposed model of DNA binding within the PRC. 74
Figure 11. Unique *cag* components contribute to *cag* T4SS gating mechanisms. 84
Figure 12. APEX2 constructs expressed in *H. pylori*. 87
Figure 13. APEX2 proximity labeling in *H. pylori.* 89

.

CHAPTER 1. INTRODUCTION

1.1 *Helicobacter pylori* and Gastric Disease

Helicobacter pylori is a Gram-negative bacterium that chronically colonizes the gastric mucosa of more than half of the global population (4). While the majority of individuals do not develop clinical signs of *H. pylori* infection, *H. pylori* colonization increases the risk for developing severe gastric disease including peptic ulcer disease, gastric mucosa associated lymphoid tissue (MALT) lymphoma (5-7), and gastric adenocarcinoma (8). Colonization by *H. pylori* is the strongest known risk factor for stomach malignancy. Gastric adenocarcinoma is the fourth leading cause of cancer-related deaths world-wide (>700,000 deaths annually), and *H. pylori* directly contributes to the development of more than one million new cases of cancer per year (9, 10). A World Health Organization group 1 carcinogen, *H. pylori* colonization accounts for approximately 75% of the global gastric cancer burden and 5.5% of all malignancies (9, 10). Stomach cancer rates are 2-fold higher in men than in women, with high incidence in South America, Eastern Asian countries (including Japan and Mongolia), and Eastern European countries. Incidence rates in Northern America and Northern Europe are comparably low and similar to rates observed in African regions (9, 11).

H. pylori-induced tissue damage elicits a pro-inflammatory response that recruits a diverse set of immune cells to the gastric mucosa, resulting in chronic gastritis (12). The progression from non-atrophic chronic inflammation to gastric cancer is described by the Correa cascade, which outlines the multistage pathological progression from superficial non-atrophic gastritis to atrophic gastritis, intestinal metaplasia, dysplasia, and gastric

adenocarcinoma (13, 14). While the most significant risk factor for the development of gastric adenocarcinoma is infection by *H. pylori* that harbors the *cag* pathogenicity island (15), a variety of bacterial, host, and environmental factors interact to influence disease progression and outcomes.

1.1.1 Bacterial Risk Factors

Only a small percentage (1-5%) of *H. pylori*-colonized individuals progress to gastric cancer (12, 16, 17), suggesting that additional factors influence gastric disease outcomes. *H. pylori* exhibits a high level of intraspecies genetic diversity resulting from high mutation and genomic recombination rates, and nearly every *H. pylori* strain is unique at the sequence level. Thus, *H. pylori* isolated from unrelated individuals may harbor gene variants, altered gene content, or different chromosomal organization associated with varied risk for developing gastric cancer (18-21).

The *cag* PAI

The strongest risk factor for development of gastric cancer is chronic colonization by *H. pylori* that the harbors the *cag* pathogenicity island (15). The *cag* PAI is a 40 kB chromosomal region that encodes approximately 27 components which assemble to form the *cag* type IV secretion system (*cag* T4SS). *H. pylori* exploits *cag* T4SS activity to alter the mucosal microenvironment through the delivery of diverse immunostimulatory effectors, including the oncoprotein CagA, chromosomally-derived DNA, fragmented peptidoglycan, and LPS metabolites into gastric epithelial cells (22-25). Translocated effector molecules interact with and disrupt several host cell signaling pathways, resulting in significant gastric inflammation and phenotypic changes that stimulate severe disease.

Currently, the bacterial oncoprotein CagA is the only known protein effector secreted by the *cag* T4SS. Upon translocation into the host cell, CagA interacts with several host cell components to induce a variety of cellular alterations, including changes in epithelial morphology, disruption of cell-cell junctions, enhanced proliferation, and inhibition of apoptosis (18, 26-29). Translocated CagA is rapidly tyrosine phosphorylated by Src and Ab1 kinases at EPIYA motifs within the carboxy terminus (30). Phosphorylated CagA binds to SHP2, a cytoplasmic tyrosine phosphatase that contains a Src homology 2 (SH2) domain (31), preventing inhibition of SHP2 phosphatase activity and inducing cell spreading and cell elongation, known as the 'hummingbird phenotype' (32). Phosphorylated CagA also binds the c-Met receptor resulting in CagA-induced cell scattering (33). Additionally, CagA disrupts signaling pathways associated with p53 tumor suppression and gastric stem cell activation and induces growth factor receptors, β-catenin, and other signaling molecules that stimulate proliferative and anti-apoptotic gene regulation signaling (1, 26, 29, 34). Unphosphorylated CagA also interacts with host cell proteins, including the tight junction scaffolding protein ZO-1 and the transmembrane protein junctional adhesion molecule (JAM) which results in the disruption of cellular interfaces (35). The combination of uncontrolled cellular growth and disruption of normal epithelial cell morphology and barrier integrity by CagA is a major contributing factor of *H. pylori*-associated gastric disease.

In addition to the oncoprotein CagA, the *cag* T4SS also delivers microbial DNA into host cells. The observation that significantly more toll-like receptor 9 (TLR9) is observed in gastric biopsies from patients at a high risk for developing cancer highlights an important role of microbial DNA recognition in gastric disease (36), and recent studies

demonstrate that TLR9 stimulation is a direct result of *cag* T4SS activity (36-38). CagA is not required for TLR9 activation, suggesting that CagA translocation and DNA delivery occur through distinct mechanisms. Furthermore, a recent study demonstrates that in addition to TLR9, translocated microbial DNA activates cytosolic nucleic acid reconnaissance pathways to elicit both anti- and pro-inflammatory responses. Thus, activation of multiple DNA signaling systems enable *H. pylori* to evade immune detection and establish persistent infection (36, 39, 40).

The *cag* T4SS also secretes several other effector molecules into gastric epithelial cells that contribute to the development of severe gastric disease. Translocated muropeptide activates the pathogen recognition molecule Nod1 to stimulate NF-kB signaling and induce pro-inflammatory responses (41-44). Lipopolysaccharide metabolites heptose-1,7-bisphosphate (HBP) and ADP-glycero-β-D-manno-heptose (ADP-heptose) delivered into gastric cells via *cag* T4SS mechanisms stimulate NF-kB independent of Nod1 signaling (45-48) to induce the production of proinflammatory cytokines such as IL-8, increase cell proliferation and migration, and promote cancer cell survival (49, 50). Collectively, *cag* T4SS activity significantly contributes to *H. pylori* pathogenesis and inflammation-associated gastric disease.

Vacuolating Toxin A

H. pylori induces intracellular vacuolation in host cells via vacuolating cytotoxin A (VacA) secretion (51, 52). While most *H. pylori* strains harbor *vacA*, toxin isoforms exhibit altered vacuolating activities associated with increased risk for gastric cancer (53, 54). Upon internalization, VacA induces structural and functional cellular changes including disrupted gastric epithelial barrier function and endosomal compartment alterations that

result in vacuole formation. Additionally, VacA targets mitochondria resulting in reduced transmembrane potential and release of cytochrome *c* followed by activation of caspase-8 and caspase-9 and induction of apoptosis (55). VacA also disrupts normal cell proliferation, differentiation, and adhesion (56-58), further highlighting the critical role of VacA intoxication in *H. pylori*-induced gastric disease.

Adhesins

Adhesins and outer membrane proteins (OMPs) also play an important role in persistent *H. pylori* infection and the development of severe gastric disease. Adherence to host cells facilitates *H. pylori* colonization and chronic infection as well as the delivery of virulence factors into the gastric epithelium. The most well-characterized *H. pylori* OMP associated with increased disease risk is the blood group antigen binding adhesion (BabA) which binds to Lewis^{b} antigens to anchor *H. pylori* to the gastric epithelial cell surface (59, 60). Sialic acid-binding adhesin (61) is also associated with an increased risk for developing severe gastric disease. SabA binds sialyl-Lewis^{x} antigen expressed on the gastric epithelial cell surface (62, 63) and increased sialyl-Lewis^{x} levels produced by *H. pylori*-infected cells suggests that chronic colonization may modulate glycosylation patterns to increase bacterial attachment to the gastric epithelium (63). Other membrane proteins, including Hp0305 and Hp1564, increase gastric inflammation and the risk for stomach cancer by promoting bacterial adhesion to host cells and facilitating the *cag* T4SS-dependent delivery of oncogenic effector molecules into the gastric epithelium (64). Expression of the outer inflammatory protein (OipA) leads to increased levels of pro-inflammatory cytokines including, IL-8, IL-1, IL-17, and TNF-α, that drive gastric inflammation (65, 66). OipA is also involved in disrupting regulation of host proteins

associated with gastric cancer such as metalloproteinase 1 (MMP-1) and glycogen synthase kinase 3β (GSK-3β) (67), further demonstrating an important role in the development of severe gastric disease.

Finally, the adhesin molecule HopQ is a critical *H. pylori* virulence factor that binds to human carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) on the cell surface. HopQ-dependent interactions facilitate adhesion of *H. pylori* to host cells and is required for *cag* T4SS-dependent CagA translocation (68, 69). CEACAMS are expressed on a variety of cell types including epithelial cells, endothelial cells, and immune cells and are important regulators of cell differentiation and proliferation, apoptosis, cell migration and invasion, tumor development, and inflammatory responses (68, 70). CEACAM-1, -5, and -6 are highly expressed during *H. pylori*-induced gastritis and HopQ-CEACAM interactions regulate NF-kB activation to enhance *H. pylori* pathogenesis and stimulate gastric inflammation (68, 70-72).

Motility

Flagellar-driven motility is another important virulence factor that is essential for *H. pylori* penetration into and colonization of the stomach mucosal layer (73). *H. pylori* assembles a unipolar bundle of sheathed flagella that control bacterial movement via chemotaxis through pH and bicarbonate gradients within the mucosa (74, 75). Specific amino acid alterations within flagellar filaments prevent detection by TLR5, providing another mechanism by which *H. pylori* evades host immunity (76, 77). Other intrinsic survival mechanisms, including cell surface-associated urease that cleaves urea into ammonia and carbon dioxide, enable *H. pylori* survival at low pH encountered during transit through the stomach lumen and into the mucosal layer (12, 78, 79).

1.1.2 Host Risk Factors

Colonization by *H. pylori* is the strongest risk factor for gastric disease; however, several host genetic factors increase the likelihood of developing stomach cancer. Polymorphisms within the IL-1β gene cluster are associated with an increased risk for developing gastric atrophy and gastric adenocarcinoma compared to patients with polymorphisms that limit IL-1β expression (80, 81). *In vivo* studies demonstrate that mucosal IL-1β levels increase after challenge with *H. pylori* concomitant with a decrease in gastric acid production, highlighting an important role for IL-1β in acid secretion inhibition (82). Additionally, IL-1β levels are significantly increased in gastric tissue from *H. pylori* colonized patients with high expression polymorphisms compared to those with low expression polymorphisms, providing a direct link between IL-1β production and the severity of gastric inflammation and atrophy (83). These findings demonstrate that IL-1β acts as both an inhibitor of gastric acid secretion and a pro-inflammatory cytokine that modulates gastric disease outcomes.

Another host factor that influences gastric cancer progression is the pro-inflammatory cytokine TNF-α which is also increased in *H. pylori*-colonized gastric mucosa. Like IL-1β, polymorphisms that increase TNF-α expression are associated with an increased risk of developing gastric cancer (80, 84). Expression of TNF-α is linked to abnormal β-catenin signaling that drives the development of gastric dysplasia, suggesting that β-catenin activation is a potential mechanism through which increased TNF-α levels contribute to the development of gastric cancer (58, 85). Polymorphisms that affect the

expression of innate immune components can also increase the risk for gastric cancer. For example, high expression of the chemokine IL-8 increases inflammation and pre-malignant lesion development, while *TLR4* polymorphisms increase the risk for gastric atrophy and gastric cancer (58). Similarly, polymorphisms that result in reduced production of anti-inflammatory cytokines such as IL-10 may also increase the risk for gastric disease (80).

1.1.3 Environmental Risk Factors

In addition to host genetic determinants, environmental elements are associated with an increased risk of gastric cancer. One such factor is high dietary salt intake which has been shown to be positively correlated with the prevalence of *H. pylori* infection and increased disease risk in *H. pylori*-infected individuals (58, 86-89). Mongolian gerbil studies demonstrate that *H. pylori* and dietary salt intake synergize to potentiate gastric cancer (58, 90-92). Additionally, pro-inflammatory cytokines IL-1, IL-6, and TNF-α are increased in *H. pylori*-infected gerbils fed a high salt diet, and *H. pylori*-induced IL-1β expression is enhanced by hyperosmotic stress (93, 94). The exact mechanisms by which salt increases gastric cancer risk are unknown; however, increased pro-inflammatory cytokine expression likely contributes to the observed disease outcomes. Additionally, studies have shown that high salt concentrations can influence *H. pylori* gene expression. For example, levels of several virulence factors, including the oncoprotein CagA, are increased under high salt conditions, demonstrating a potential mechanism through which diet increases gastric cancer risk in patients colonized by *H. pylori* (95, 96).

1.2 Bacterial type IV secretions systems

Bacterial type IV secretion systems (T4SS) are incredibly versatile molecular devices harbored by phylogenetically diverse Gram-negative and Gram-positive organisms (97-101). Studies exploring the function and architecture of these nanomachines have largely focused on paradigmatic Gram-negative systems that are deployed to translocate protein, DNA, multi-subunit toxin, and other macromolecular substrates to target cells (101). The fascinating T4SS superfamily can be broadly segregated into three functional subtypes that uniquely contribute to pathogenicity and genome plasticity (97-99). Bacterial conjugation machineries represent the largest and most wide-spread T4SS subfamily. These molecular devices enable the export of single-stranded DNA (ssDNA) and associated proteins across the donor bacterial membrane and into target prokaryotic and eukaryotic cells via mechanisms that require direct cell-cell contact (97-102). First described by Lederberg and Tatum nearly 80 years ago (103), T4SSs facilitate rapid microbial adaptation to challenging environments through the acquisition of fitness traits encoded within mobile genetic elements. Notably, conjugative DNA transfer is the predominant mechanism by which antibiotic resistance determinants and virulence factors are transferred within and between phylogenetically diverse species (98, 104-108).

Specialized modifications to 'minimized' conjugation systems may enable novel secretion phenotypes or expanded target cell range (101). In contrast to conjugation machineries, the 'effector translocator' T4SS subfamily is used to deliver assorted protein effectors into the cytoplasm of eukaryotic host cells in a contact-dependent manner (97, 98, 100). In some cases, effector translocators implement the interkingdom transfer of nucleoprotein and polysaccharide substrates (44, 47, 97, 100, 102, 109, 110). Thus, effector

translocator systems play a critical role in establishing host-pathogen interactions that promote host tissue colonization and foster bacterial persistence within distinct niches (97, 98). In unusual cases, T4SS-dependent mechanisms facilitate interbacterial killing via toxin delivery (111, 112) or the contact-independent secretion of multi-subunit protein toxins into the extracellular environment (98, 113). The third T4SS subfamily comprised of 'DNA uptake and release' systems function in a contact-independent manner to import exogenous DNA or to secrete DNA into the extracellular milieu (97, 98, 101, 110, 114-117). Thus, these dedicated genetic exchange systems expand the molecular arsenal employed by Gram-negative bacteria to acquire survival elements during infection (110, 114, 118).

1.2.1 Interbacterial DNA Delivery Systems

Pioneering work performed over the past 75 years has uncovered mechanisms underscoring targeted DNA excision and recruitment of corresponding nucleoprotein substrates to conjugative T4SS machineries. Intricate biochemical details involved in DNA processing have been extensively characterized *in vitro* and *in vivo* (119). The process of conjugative DNA transfer initiates when accessory factors, which are typically members of ribbon-helix-helix DNA binding proteins (such as TraM from F-family plasmids (120), or TrwA from plasmid R388 (121)), bind the cognate origin of transfer (*oriT*) sequence located on either an extrachromosomal plasmid or integrative conjugative element (ICE) . A combination of double-stranded DNA bending and direct protein-protein interactions mediated by *oriT*-bound accessory factors (120, 122, 123) recruits the relaxase (a specialized site- and strand-specific phosphodiesterase (107)) to form the catalytically active relaxosome complex (119, 124). Within the relaxosome, the relaxase initiates

substrate processing by triggering an active site-mediated nucleophilic attack of the *oriT* scissile phosphate group, forming a high-energy bond which covalently links excised ssDNA product to the relaxase (97, 101, 119, 124, 125). This bond enables (i) directional nucleoprotein piloting through the T4SS conjugative apparatus, (ii) protects the phosphate group from subsequent nucleophilic attack in the recipient cell, and (iii) facilitates plasmid ends re-joining in the host cell cytoplasm (119, 124).

Relaxases are large, multidomain proteins that typically contain an N-terminal trans-esterase domain (the 'relaxase') which carries out the phosphodiesterase reaction (105, 107, 125). Domains localized to the C-terminal region may be involved in DNA helicase or primase activities or execute unknown functions (107, 126). While conjugative relaxases are categorized into eight mobility ('MOB') groups (127), recent efforts have uncovered the intricate biochemical details underlying the function of MOB_F nuclease associated with F-family conjugation systems (*e.g.*, TraI associated with F, pKM101, RP4, R1, pED208, and TrwC encoded by R388) (107, 128). In addition to the N-terminal trans-esterase domain that facilitates the nicking and covalent attachment of transfer DNA to the relaxase (129), TraI exhibits a C-terminal vestigial helicase domain that binds ssDNA (130), and an active helicase domain that directionally unwinds DNA 5′ to 3' (131, 132). An additional TraI C-terminal domain of unknown function may be used to recruit relaxosome components to target DNA sequences (131, 132). Recent work determined the structure of TraI complexed to the F plasmid-encoded *oriT* at atomic resolution (133). Cryo-EM analysis revealed that ssDNA binds deep within the structure of TraI, with the 5′ region bound longitudinally to the trans-esterase and active helicase domains and the 3′ region extending through the active and vestigial helicase domains thus facilitating

remarkably high enzymatic processivity (107, 133). Structural analysis also mapped two translocation signals proximal to the TraI C-terminus that likely contribute to relaxosome recruitment to the cognate VirD4 coupling protein for transport through the secretion apparatus (107, 133-135).

Coupling proteins are phylogenetically and functionally related to SpoIIIE and FtsK ATPases that translocate DNA across membranes during sporulation and cell division (105, 125, 136). Structural data and biochemical evidence indicate that VirD4 ATPases assemble into a homohexamer with an N-terminal transmembrane domain, a nucleotide binding/hydrolysis domain (137, 138), and a sequence-variable domain that confers substrate binding and specificity (105, 139-141). Additional variable C-terminal domains may provide additional coupling specificity (120) or control the selection and translocation of substrates through the T4SS (142, 143). Relaxosome docking to VirD4 coupling proteins and the T4SS apparatus is achieved through cooperative interactions among the relaxase and accessory factors (107, 124, 125, 144). For example, the F plasmid accessory factor TraM binds a short C-terminal motif within the VirD4-like coupling protein TraD (120, 123), whereas internal TraI translocation signals are accessible only when the relaxase adopts defined tertiary structures (133, 134). Other relaxases may harbor C-terminal translocation signals that facilitate relaxase-VirD4 receptor interactions (125, 135); however, the structural basis of such interactions remains undefined. Additional specialized accessory factors, including ParA-like DNA partitioning proteins, physically tether the relaxosome to VirD4 receptors via multiple protein-protein interactions (145, 146). Subsequent to VirD4 docking, the relaxase is unfolded via unknown mechanisms (147) and released from associated accessory factors for delivery through the translocation

channel (97, 107, 125). In the recipient cell cytoplasm, the relaxase catalyzes the re-circularization of transferred cargo through a reversal of the strand-breaking reaction, followed by second-strand synthesis and DNA replication (148, 149).

Our understanding of conjugative T4SS architectures primarily stems from structures resolved for the pKM101-encoded OMCC (150-152) and nearly intact R388 systems (153, 154). The OMCC assumes a barrel-shaped structure incorporating 14 copies of the OM-associated VirB7 lipoprotein, VirB9, and the VirB10 C-terminal region. Crystallographic analyses of the barrel outer (O) layer revealed VirB10 domains lining the interior and connected to an outer protective crown composed of VirB7 and VirB9 (152). The barrel inner (10) layer is formed by VirB9 N-terminal domains surrounding VirB10, which extends across the inner membrane (151, 155). The IMC of conjugative nanomachines exhibits remarkable asymmetry and features two side-by-side VirB4 ATPase hexamers extending from the IMC into the cytoplasm. The IMC comprises 12 copies of VirB3, VirB4, VirB5, VirB6, and VirB8, as well as 14 copies of the VirB10 N-terminal region (101, 107, 125, 156). Contacts between VirB10 and additional subunits within the IMC thus form the symmetry mismatch within the conjugative machinery that may enable functional plasticity (101, 125) (**Fig. 1**).

Conjugation systems harbored by Gram-negative bacteria elaborate flexible, dynamic conjugative pili that extend from the donor cell and retract to bring potential recipient cells into close proximity (101, 157). Other conjugative systems assemble brittle pili that are shed from the cell surface to induce bacterial aggregation to enable DNA transfer (125). Our understanding of conjugative pilus biogenesis and assembly has been significantly advanced by recent structural analyses of purified pili and *in situ* cryo-ET

studies of intact T4SS machineries (154, 158, 159). Analyses of assembled machineries within the *E. coli* cell envelope revealed multiple F-encoded pilus configurations docked onto alternative basal platforms (158), providing the first direct evidence that, in contrast to other cell surface pili and type III secretion system needle complexes which nucleate on an inner membrane platform, the T4SS pilus assembles at the outer membrane (158). Recent cryo-EM studies of the R388 system provide a structural model for pilus biogenesis whereby VirB2 pilin pentamers bound by VirB6 subunits in the inner membrane are leveraged into the pilus assembly site through the O-layer channel (154).

Despite major advances in our understanding of conjugative pilus assembly, the role of pili in DNA transfer has not been firmly established. Evidence suggests that conjugative pili can serve as translocation conduits through which ssDNA and unfolded relaxases can be delivered into recipient cells at long range (160). In support of this hypothesis, the *E. coli* F pilus (159) and *A. tumefaciens* T pilus (161, 162) lumens are lined by inner membrane-derived phospholipids which impart a weak negative charge that may facilitate DNA passage (159). Conflicting observations support an indirect role of pili in substrate delivery to target cells. For example, multiple studies identified secretion 'uncoupling' mutations that selectively block T4SS pilus biogenesis but do not obstruct substrate transfer, suggesting that elongated pili are not required for cargo translocation (155, 163, 164). Additionally, T4SSs harbored by Gram-positive bacteria do not assemble conjugative pili, yet these systems transfer DNA to recipient cells at high frequencies (97, 99). Although luminal phospholipids may facilitate low-efficiency conjugation, these lines of evidence suggest that conjugative pili serve primarily as adhesive organelles that

establish tight donor-recipient mating junctions required for cell contact-dependent DNA transfer (101, 125).

Mating channels may have initially served as protein transport systems that were later adapted into conjugation machineries following the recruitment of coupling proteins associated with rolling circle replicases (126). Based on the inferred evolutionary relationships between coupling proteins and cognate VirB4 ATPases, conjugative T4SSs emerged in Gram-negative organisms and were later domesticated by archaea and monoderm Gram-positive species (99, 165). Studies of T4SSs encoded on the *Enterococcus faecalis* sex-pheromone responsive plasmid pCF10, *C. perfringens* plasmid pCW3, and the *S. agalactiae* broad host-range plasmid pIP501 revealed that in contrast to orthologous systems in Gram-negative bacteria, Gram-positive conjugation systems lack paradigmatic OMCC subunits and the VirB11 ATPase (99, 105, 166-168). In addition to employing surface adhesins to mediate donor-recipient cell contacts (as opposed to conjugative pili), Gram-positive machineries utilize VirB1-like lytic transglycosylases with multiple catalytic domains to enable assembly of the machine across the dense peptidoglycan layer (99, 169).

Mounting evidence suggests that substrate transfer across the cytoplasmic membrane is mechanistically and structurally conserved among Gram-negative and Gram-positive bacteria. For example, early substrate processing and coupling protein docking reactions in gram-positive systems closely mirror corresponding mechanisms underscoring DNA delivery in Gram-negative systems. As in Gram-negative systems, Gram-positive plasmid- and ICE-encoded relaxases display strand-specific cleavage of cognate *oriT* regions and may coordinate with accessory factors to mediate docking to the coupling

protein (97, 99, 105, 170). Current models of DNA coupling posit that Gram-positive relaxases form a transient covalent adduct to excised ssDNA substrates via conserved active site tyrosine residues; however, in contrast to Gram-negative relaxases, translocation signals that enable nucleoprotein transfer intermediate delivery to cognate coupling proteins have not been defined in Gram-positive orthologs (97, 105, 170). Likewise, comparison of multiple Gram-positive conjugative systems identified six subunits that exhibit homology to the *A. tumefaciens* VirD4, VirB1, VirB3, VirB4, VirB6, and VirB8 channel components (97, 99). Currently, limited structural and biochemical analyses support the assumption that orthologous Gram-positive T4SS subunits facilitate substrate transfer across the cytoplasmic membrane (97, 99). While high resolution structures of individual components encoded within the pIP501 and pCW3 systems have been determined, Gram-positive conjugation machinery architectures have not been visualized *in situ*. Thus, future work to resolve the structures of Gram-positive T4SS assemblies by cryo-EM will significantly advance our understanding of conjugative DNA transfer through evolutionarily-diverse and minimized machineries.

1.2.2 DNA Export Into The Extracellular Milieu

The obligate human pathogen *Neisseria gonorrhoeae* employs a contact-independent T4SS apparatus to secrete chromosomally-derived ssDNA substrates into the extracellular milieu that facilitate the robust transformation of other highly competent *Neisseria* species (115, 171-173). In addition to contributing to exceptional genetic diversity, secreted ssDNA promotes host colonization by stimulating the initial stages of biofilm formation to establish persistent *N. gonorrhoeae* infection (172). Encoded on the 59 kb gonococcal genetic island (174), the specialized DNA release apparatus is harbored

by approximately 80% of *N. gonorrhoeae* and 17.5% of *N. meningitis* clinical isolates (115, 171, 175, 176). The GGI harbors T4SS structural homologs that share sequence similarities to the *E. coli* F plasmid conjugation system and other T4SS machineries (including TraB, TraK and TraV, which comprise the outer membrane core complex, the inner membrane proteins TraL and TraE, and the periplasmic components TraW, TraU, TraH and TraN) as well as highly conserved components involved in DNA processing and recruitment, including partitioning proteins (ParA and ParB) and the relaxase TraI (115, 171, 177). Additional T4SS components, such as the coupling protein TraD (178) and TraC, a homolog of the ATPase VirB4, are also required for DNA secretion (115, 171, 179-182). Interestingly, the pilin homolog, TraA, and TbrI, which exhibits homology to a serine protease that plays a role in circularization of the pilin subunit, are not required for DNA secretion, indicating that DNA release into the extracellular space is not dependent on formation of an extracellular pilus (180). This suggests that gonococcal homologs of T4SS components involved in pilus assembly and extension (TraW, TraU, TraH, TraF, TrbC and TraN) play unique roles within the *Neisseria* T4SS (181). Other GGI-encoded proteins required for ssDNA secretion include peptioglycanases and the disulfide isomerase DsbC (115, 171, 180, 183).

N. gonorrhoeae ssDNA processing and translocation parallels interbacterial DNA conjugation mechanisms. For example, chromosomal DNA processing relies on TraI relaxase-mediated DNA cleavage at *oriT*-like sequences and recruitment to the secretion apparatus by the cognate coupling protein TraD for transport across the inner and outer membranes (115, 177, 178). Experimental evidence supports a model whereby secreted ssDNA is delivered across the bacterial envelope as a nucleoprotein complex that is 5'

protected (presumably mediated by covalently bound TraI relaxase) and is susceptible to nucleases that directionally cleave ssDNA from 3' free ends (184). Amino acid substitutions identified two tyrosines in TraI that are functionally important for DNA secretion – substitution of Tyr^{93} resulted in the complete loss of DNA secretion while Tyr^{201} mutagenesis resulted in partial loss of DNA secretion, demonstrating that both residues are necessary for TraI-dependent DNA processing (177). These data suggest that Tyr^{93} is important for the initial cleavage of DNA while the Tyr^{201} may be involved in a second cleavage that terminates DNA processing (177). Because tyrosine residues contribute to non-specific DNA interactions, Tyr^{201} may enhance TraI affinity for DNA backbone binding (177). Additional *in vitro* evidence supports a role of Tyr^{212} in DNA cleavage (185). A TraI HD phosphohydrolase domain, which contains highly conserved histidine and aspartate residues (186), is also essential for *Neisseria* T4SS-dependent DNA secretion. Amino acid substitutions of the signature histidine and aspartate within this domain significantly reduce DNA secretion, suggesting that charged residues are important for TraI function; however, the exact role of the phosphohydrolase remains unknown (177). Additionally, *N. gonorrhoeae* TraI harbors a distinct N-terminal domain that mediates membrane interactions required for DNA secretion, representing a unique step in contact-independent DNA transport through T4SS machinery (177).

ParA and ParB, partitioning proteins encoded on the GGI, are also essential for DNA secretion by the *N. gonorrhoeae* T4SS (115, 180). Partitioning proteins play a role in localizing chromosomal and plasmid DNA during cell division (187); however, the role of DNA segregation machinery in T4SS effector translocation is incompletely defined. In the prototypical *A. tumefaciens vir* T4SS, ParA (VirC1) and ParB (VirC2) homologs

spatially coordinate the T-DNA/relaxasome complex to the T4SS machinery localized to the cellular poles to initiate DNA secretion (188). Gonococcal ParA and ParB have been shown to interact with the TraI relaxase, suggesting that gonococcal ParAB may assume a similar role in facilitating *N. gonorrhoeae* DNA secretion via relaxasome cellular positioning and docking to the T4SS apparatus (189). Spatial assembly of the secretion machinery is influenced by GGI-encoded peptioglycanases that generate a localized break within the cell wall to implement T4SS apparatus biogenesis (180, 190). The GGI also encodes additional peptidoglycan-associated proteins including EppA (a protein that severs peptidoglycan crosslinks) and Yag, a OmpA-like protein that binds peptidoglycan and may facilitate machinery assembly by tethering the T4SS apparatus to the cell wall (114, 115, 180).

Gonococcal TraG, a homolog of the *E. coli* F plasmid mating pair formation protein, is also essential for ssDNA secretion through the GGI-encoded T4SS (179, 191). The observation that gonococcal DNA secretion occurs in a contact-independent manner suggests that in contrast to its canonical role in conjugative DNA transfer systems, gonococcal TraG evolved novel features to enable DNA export through the *N. gonorrhoeae* T4SS. TraG is localized to the gonococcal inner membrane via five transmembrane domains which orient the N terminus to the periplasm and the C terminus to the cytosol (191). Inner membrane localization and prominent periplasmic and cytosolic domains indicate that TraG may interact with other T4SS structural components, or possibly with the relaxasome complex. While mapping experiments to determine TraG topology within the T4SS apparatus represent a snapshot of machinery architecture, the possibility that the TraG C-terminal domain is cleaved and delivered to the extracellular milieu remains an

intriguing possibility (191). Additionally, strains harboring truncated TraG variants are unable to secrete DNA substrates, suggesting that TraG may orchestrate effector specificity (114, 191). Alternatively, TraG may provide architectural scaffolding during T4SS apparatus biogenesis. In support of this hypothesis, TraG is produced from a strain-variable operon that encodes additional T4SS components required for DNA secretion including TraH and the lytic transglycosylase AtlA (114, 171, 179, 191). TraH localization is dependent on the presence of TraG, suggesting that TraG may stabilize TraH or facilitate TraH transport to the outer membrane (181). In F plasmid-encoded T4SSs, TraH homologs are involved in conjugative pilus assembly; however, because *Neisseria* ssDNA secretion does not require a T4SS-associated pilus, gonococcal TraH appears to assume a unique role within the GGI-encoded T4SS (180, 181). While biochemical analyses have significantly advanced our understanding of contact-independent ssDNA transport, the architecture of GGI machinery remains unresolved.

1.2.3 DNA Uptake By Competence Systems

Natural transformation (NT) is a widespread mode of horizontal gene transfer that contributes to the evolutionary adaptation of diverse bacteria. Approximately 80 bacterial species, including both Gram-positive and Gram-negative organisms, have been reported to be naturally transformable (192). To achieve NT, bacteria assemble membrane-spanning machinery complexes and develop a physiological state known as competence that allows the uptake of exogenous DNA and its subsequent integration into the recipient chromosome. During DNA uptake, double-stranded DNA (dsDNA) is captured from the extracellular environment and accumulates at the outer surface (Gram-positive bacteria) or the periplasm (Gram-negative bacteria) (193). In a process that is incompletely defined,

exogenous dsDNA is converted into a single-stranded form that is channeled via the inner membrane protein ComEC into the cell cytoplasm (194, 195). Single-stranded DNA (ssDNA) is integrated into the recipient chromosome by cytoplasmic recombination machinery that incorporates DprA (15), RecA (196), and other competence-related proteins such as ComFC (197). While the ComEC and homologous recombination machinery is conserved in most naturally transformable bacteria (192, 198), the composition of DNA uptake machinery varies widely. Many bacteria employ a retractable type IV competence pilus (T4P) for dsDNA capture from the extracellular environment (199). In these cases, exogenous DNA is pulled to the outer membrane or periplasm by T4P retraction and subsequent binding to the competence proteins ComEA (193, 200, 201). Additionally, a recent study demonstrates that cell wall-associated teichoic acid mediates and enhances DNA binding during NT (202).

In contrast to most naturally transformable bacteria, *H. pylori* employs the *comB* T4SS for DNA uptake into the periplasm (203, 204). The *comB* apparatus exhibits homology to corresponding components within the prototypical *A. tumefaciens vir* T4SS and is encoded by genes clustered into two distinct operons (*comB2-comB4* encoding 3 genes, and *comB6-comB10* encoding 5 genes) (203, 204). Structurally, ComB7, ComB9, and ComB10 are proposed to form the outer membrane complex (110, 156) while ComB6-ComB8-ComB10 subassemblies are postulated to span the periplasm to bridge the inner membrane translocation apparatus and the outer membrane-associated complex (203, 204). Comparison to the *A. tumefaciens vir* T4SS reveals that while *comB* harbors a VirB4 ATPase homolog, the apparatus lacks multiple components essential for *vir* T4SS function including VirB1, the minor pilin VirB5, the DNA coupling protein VirD4, and the ATPase

VirB11 (110). Similar to *A. tumefaciens* VirB2, the major pilin ortholog ComB2 is proposed to be involved in cell surface DNA binding (102, 110).

Although the mechanistic details underscoring *comB* T4SS-mediated DNA import are unresolved, a two-step DNA uptake process has been described for *H. pylori* (205, 206). In the first step, the *comB* T4SS mediates exogenous DNA transport across the outer membrane followed by the second step of translocation into the bacterial cytoplasm via ComEC (206). ComH, a periplasmic protein that binds dsDNA, provides an essential link between outer and inner membrane transport (207). During DNA import across the cell envelope, the ComH C-terminus interacts with incoming DNA and bridges substrate exchange with ComEC via the N-terminal domain to facilitate delivery across the inner membrane (207), suggesting that ComH is important for periplasmic DNA loading into the translocation channel. It is thus possible that ComB2 shuttles incoming DNA through a 'relay' mechanism with ComH prior to substrate loading into the inner membrane translocation channel for import into the recipient cytoplasm (110, 207).

DNA uptake is a highly regulated process that does not lead to bacterial growth arrest (208, 209). For example, levels of ComB8 and ComB10 correlate with transformation rates and overexpression of the *comB6-comB10* operon leads to increased competence (209). The kinetics of high velocity *comB* T4SS-mediated DNA import exceed corresponding rates and total DNA import volumes observed for analogous type II secretion system/T4P uptake devices (110, 206, 210). For example, while *Neisseria* T4P retraction systems have the capacity to import around 40 kilobase pairs (200), the *comB* T4SS translocates up to the equivalent of one chromosome into the periplasm (208). In contrast to *Neisseria* T4P systems that transport both ssDNA and dsDNA substrates into

the bacterial cell (210), it is unknown whether the *comB* apparatus supports exogenous ssDNA translocation into the periplasm (110). Finally, in addition to significantly enhancing genetic diversity, competence may represent a defensive strategy in which increased reservoirs of periplasmic DNA protect the chromosome by mitigating oxidative stress encountered in the gastric niche (208). Alternatively, NT may protect the core genome and serve as a mechanism to combat undesirable outcomes associated with hypervariability or genome instability in microbes that undergo frequent intrachromosomal recombination (211). While competence is not required to establish infection, *comB* T4SS-mediated transformation may provide competitive advantages that promote chronic gastric colonization or *H. pylori* dissemination (110, 212, 213). Thus, future studies to elucidate *comB* apparatus architecture and the role of natural competence in bacterial persistence are warranted.

1.2.4 Trans-kingdom DNA Transfer via Conjugative Mechanisms

Interdomain conjugation is a major driving force underlying genetic exchange among varied microbial species. In addition to horizontal gene transfer from donor to recipient bacteria, conjugative mechanisms can facilitate genetic element dissemination from bacteria to archaea or to diverse eukaryotic cells. For example, the broad host range proteobacterial IncP1 and associated derivative transfer systems have the capacity to translocate plasmid-derived DNA to bacteria (214, 215), archaea, yeast, and higher order eukaryotic cells (215-220). Accordingly, multiple studies demonstrate that *E. coli* harboring IncP1 systems deliver plasmid-borne DNA to the methanogenic archaeon *Methanococcus maripaludis* (216), and evolutionarily divergent eukaryotes including yeast (*e.g.*, *Saccharomyces cerevisiae* (217, 221, 222), *S. kluyveri* (223), *Kluyveromyces*

lactis (224), and *Pichia angusta* (224)), unicellular algae diatoms (218), plant cells (219), and cultured human cells (220). Although there is a paucity of evidence supporting the role of IncP-mediated genetic transformation in nature, these studies suggest that trans-kingdom conjugation is not limited to the prototypical *Agrobacterium tumefaciens* system.

Recent evidence supports the hypothesis that inter-kingdom DNA transfer is a vestigial consequence of ancestral mechanisms underlying generalized conjugation through T4SS machinery (99, 105). Thus far, 'dual function' systems that translocate diverse effector proteins and achieve DNA conjugation have been discovered in very few organisms. The most extensively characterized dual function machinery is the *A. tumefaciens vir* T4SS that translocates tumor-inducing DNA (T-DNA) and associated effector proteins into recipient plant cells (97-100, 102, 105). Multiple reports indicate that additional eukaryotic cell types are susceptible to transformation via *A. tumefaciens* trans-kingdom conjugation including filamentous fungi (225, 226), arachnid cells (227), and cultured human cell lines (228). In both the *A. tumefaciens vir* T4SS and prototypical interbacterial conjugation systems, DNA is transmitted cell-to-cell as a nucleoprotein complex in which the 5' end of the ssDNA substrate is covalently complexed to a cognate relaxase protein that is secreted into target cells (100, 101). Thus, the relaxase serves to directionally pilot tethered DNA through the translocation channel (97, 100, 101).

When in contact with target plant cells, *A. tumefaciens* T-DNA is excised from the cognate Ti plasmid at defined T-border sequences by the VirD1 DNA topoisomerase, VirD2 endonuclease, and other accessory proteins that form the catalytically active relaxosome (100, 102, 105). Within the relaxosome, target DNA is cleaved by the VirD2 relaxase which remains covalently tethered to the 5' end of the excised ssDNA via a

phosphotyrosine linkage (100, 105). The relaxase and associated accessory proteins promote relaxosome recruitment to the VirD4 coupling protein for docking to the *vir* T4SS apparatus (100, 102, 104, 105, 229, 230). Although VirD4 serves as the relaxosome receptor, substrate loading into the secretion channel is likely coordinated by two additional ATPases, VirB11 and VirB4 (102, 104, 231, 232). Upon relaxosome docking to VirD4, the relaxase is unfolded via unknown mechanisms (147), resulting in the release of accessory proteins and nucleoprotein translocation through the *vir* T4SS apparatus (100, 102, 105). Thus, ssDNA cargo is piloted into recipient plant cells as a "hitch-hiker" molecule that is coincidentally delivered with relaxase effectors. In the recipient plant cell, VirD2 catalyzes T-DNA recircularization and facilitates nucleoprotein complex trafficking into the nucleus for chromosomal integration (97, 100, 105). In support of a generalized protein secretion model, the *vir* T4SS independently transports additional proteins into plant cells including VirE2, VirE3, VirD5, and VirF (98, 100). The VirE2 ssDNA binding protein coats translocated DNA cargo in the recipient cell cytoplasm and complexes with VirD2 as a mechanism by which to protect effector DNA cargo from cellular nucleases (233). VirE3, VirD5, VirF, and other *vir* T4SS effectors injected into plant cells promote infection and are implicated in crown gall tumorigenesis *in planta* (100, 233, 234).

In addition to *A. tumefaciens*, recent studies demonstrate that other plant and animal pathogens, including *Rhizobium etli* (235), *Bartonella henselae* (135, 236-238), *Coxiella burnetii* (237) *Legionella pneumophila* (237, 239, 240), and *Helicobacter pylori* (36) have the intrinsic capacity to translocate DNA into eukaryotic cells via 'dual purpose' T4SS-dependent mechanisms. For example, *L. pneumophila* and *C. burnetii* harboring artificial *dot/icm* T4SS chimeras integrating the MobA relaxase associated with the mobilizable

plasmid RSF1010 gain the capacity to translocate ssDNA into mammalian cells (237). In support of the hypothesis that conjugation is an inherent and ancient function of 'dual purpose' machineries, early studies demonstrate that the *L. pneumophila dot/icm* T4SS mobilizes RSF1010 plasmids into recipient *Legionella,* as well as some *E. coli* strains, via MobA relaxase-dependent conjugative mechanisms (237, 240). Similarly, the *B. henselae vir* T4SS can co-opt the heterologous TrwC relaxase associated with the R388 self-transmissible plasmid to achieve trans-kingdom ssDNA delivery into endothelial cells (236, 241). Chimeric relaxase fusions harboring a C-terminal translocation signal derived from the cognate effector protein BepD significantly enhanced *B. henselae* ssDNA trans-kingdom conjugation efficiencies (236), highlighting the potential to engineer T4SS machinery that delivers therapeutic designer DNA into target human cells *in vivo*. Importantly, TrwC has been shown to facilitate the random integration of plasmid-derived effector DNA into the recipient genome (238) presumably via VirD2-like nuclear import mechanisms. Alternatively, translocated *B. henselae* DNA integration and transgene expression may occur through a nuclear import-independent pathway that relies on breakdown of the nuclear envelope during cellular division (236, 238, 241).

In contrast to *A. tumefaciens*-like conjugation systems that deliver plasmid-borne ssDNA into mammalian cells, the *H. pylori cag* T4SS delivers chromosomal fragments into gastric epithelial cells in a contact-dependent mechanism (36). *H. pylori cag* T4SS-dependent DNA translocation is partially susceptible to exogenous nucleases and neutralizing anti-DNA monoclonal antibodies (36), suggesting a two-step translocation mechanism that does not appear to require an extracellular T4SS-associated pilus conduit. Translocated *H. pylori* DNA activates gastric TLR9 and additional immunostimulatory

pathways that induce both anti- and pro-inflammatory responses (36, 40). Thus, *cag* T4SS-dependent DNA translocation provides additional mechanisms by which *H. pylori* manipulates the immune response to generate a hospitable environment that fosters chronic colonization of the gastric niche (36, 40). In support of the observation that *cag* T4SS activity stimulates multiple nucleic acid reconnaissance systems (36, 242), recent work demonstrates that *H. pylori* has evolved mechanisms to counterbalance or suppress nucleic acid signaling within the gastric mucosa (40, 243). For example, *H. pylori*-induced NF-kB activation induces the upregulation of TRIM30a to antagonize innate responses elicited by TLR9 and other DNA sensors (243). While these studies provide compelling evidence that *H. pylori* provokes nucleic acid surveillance signaling via *cag* T4SS activity, the exact mechanism by which chromosomal DNA is excised and transported through the secretion apparatus remains unresolved.

1.3 Structural Architecture of T4SSs

Recent progress enabled by single particle cryo-electron microscopy (cryo-EM), *in situ* cryo-electron tomography (cryo-ET), and high-resolution correlative fluorescence microscopy has yielded new insights into T4SS assembly dynamics and spatial organization in the bacterial envelope under near-native cellular conditions (101, 153, 244-249). In Gram-negative organisms, 'minimized' bi-membrane spanning T4SSs are composed of approximately 12 conserved components designated VirB1 through VirB11 and VirD4 in accordance with a unified nomenclature based on the prototypical *Agrobacterium tumefaciens vir* T4SS (100, 101, 163). In these simplified systems, the apparatus inner membrane complex (IMC) subassembly forms part of the periplasm-

spanning translocation channel and connects to the outer membrane-associated core complex (OMCC) assembly via a stalk-like structure (151, 153, 156, 249). In some systems, elongated VirB10 domains bridge the OMCC and IMC (101, 153, 249). A trio of ATPases, including the VirD4 coupling protein that tethers DNA and protein effectors to the secretion channel (101, 231, 250-252), form the energetic apparatus at the cytoplasmic face of the inner membrane. Other bacteria, assemble more complex 'expanded' T4SSs (101, 253). Examples of expanded T4SSs include the *Legionella pneumophila dot/icm* T4SS (245, 254-256), the F-plasmid encoded T4SS(158, 159), and the *H. pylori cag* T4SS (244, 257, 258), which contain more components and are larger in size than the prototypical 'minimized' systems (**Fig. 1**). Recent studies investigating the structures of T4SSs have broadened our understanding of T4SS architectural and biological diversity.

1.3.1 The *H. pylori cag* T4SS

The *cag* T4SS is an 'expanded' T4SS made up of approximately 27 components, the majority of which are unique to *H. pylori* and whose function within the secretion system remains unknown. Protein sequence similarity with components of other T4SSs, the subcellular localization of Cag components, and protein-protein interaction studies have contributed to the development of several proposed models of the overall structure of the *cag* T4SS (259-261). However, because most Cag components do not exhibit primary sequence similarity to T4SS components in other bacterial species, an accurate model of *cag* T4SS architectural organization has not been fully resolved.

Recent cryo-ET and cryo-EM studies have provided further insight into the structural organization of the *cag* T4SS. These studies revealed that the *cag* T4SS is comprised of distinct structural features, including an outer membrane cap (OMC)

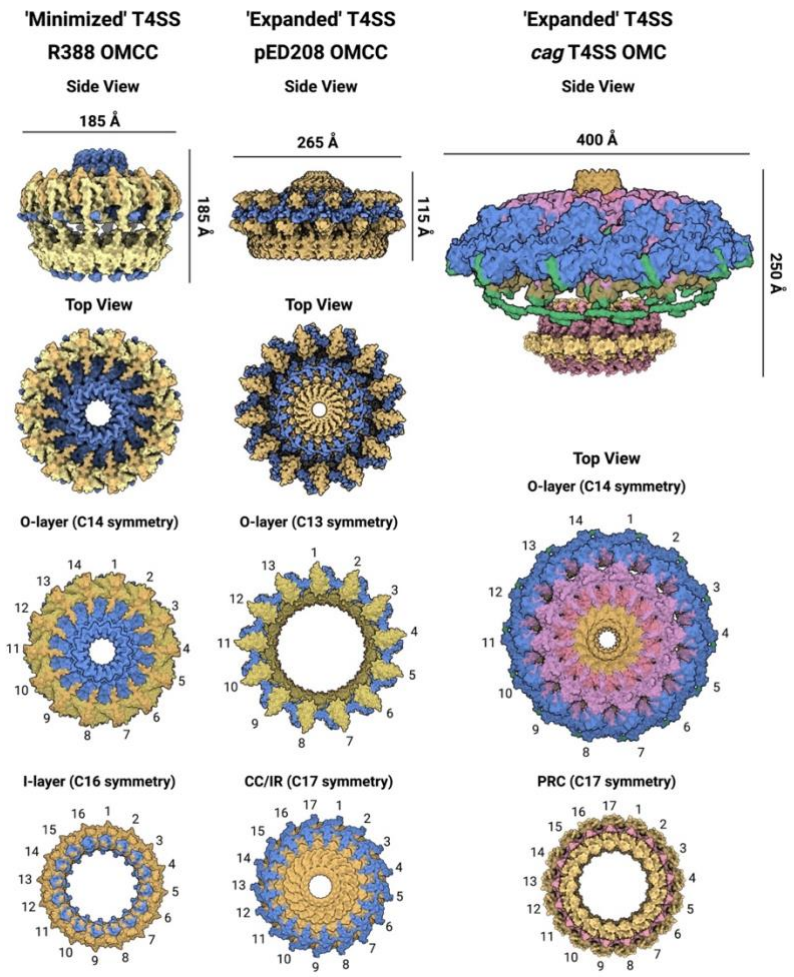

Figure 1. Symmetry mismatch in 'minimized' and 'expanded' T4SS architectures.

High-resolution structures of isolated 'minimized' (R388 (30, 31)) and 'expanded' (pED208 (43, 44) and *cag* T4SS (39, 41)) core complex machineries reveal asymmetric regions within apparatus subassemblies. The 'minimized' R388 core complex is characterized by a tetradecameric O-layer composed of VirB7 and the C-terminal domains of VirB10 ($VirB10_{CTD}$) and VirB9 ($VirB9_{CTD}$) displaying 14-fold symmetry, and a corresponding hexadecameric I-layer composed of $VirB10_{NTD}$ and $VirB9_{NTD}$ exhibiting 16-fold symmetry (78) [PDB: 7O3J and 7O3T]. The iconic F plasmid (encoded by pED208) OMCC forms two radial, concentric rings characterized by a peripheral O-layer ring exhibiting 13-fold symmetry and an interior central cone/inner ring (CC/IR) with 17-fold symmetry (43, 44) [PDB: 7SPB and 7SPC]. The 'expanded' *H. pylori cag* T4SS outer membrane complex (OMC, 14-fold rotational symmetry) and corresponding periplasmic ring complex (PRC, 17-fold rotational symmetry) exhibit striking symmetry mismatch (39, 41) [PDB: 6X6S and 6X6J]. The figure illustrates conserved architectural asymmetry among diverse T4SS machineries.

consisting of an outer layer (O-layer) and inner layer (I-layer), a periplasmic ring complex (PRC) and a stalk region (244, 250, 257, 258). *In situ* cryo-ET of the *cag* T4SS revealed several additional densities not identified by complex purification and single particle cryo-EM approaches, including central periplasmic cylinders and wing-like collars anchored to large inner membrane complexes (244-246, 250). Currently, the composition of the structures is unknown.

Single particle cryo-electron microscopy analysis identified three distinct subassemblies within the *cag* T4SS including the outer membrane core complex (OMCC), a periplasmic ring complex (PRC), and a central stalk (262). The *cag* T4SS OMCC is composed of 5 components – CagX, CagY, CagT, Cag3 and CagM (263). CagX, CagY and CagT are *vir* T4SS homologs (VirB9, VirB10, and VirB7 respectively), while Cag3 and CagM are unique to *H. pylori*. Of these 5 components, CagX, CagY and CagM are required for OMCC formation and stability. In the absence of Cag3 and CagT, which are localized to the periphery of the complex, the OMCC diameter is narrower compared to wild-type complexes (263-265). The OMCC forms a mushroom cap-like structure that, like prototypical T4SSs, has an O-layer made up of CagT, Cag3 and the C-terminal portion of CagX and CagY, and an I-layer made up of CagM as well as other unidentified components (262, 264). The *cag* T4SS core complex forms a 270 Å pore that extends from the bottom of the PRC to the top of the OMCC where it is reduced to a 35 Å opening. CagY is localized to the top of the OMCC and is organized in a ring of α-helices that are predicted to form the channel across the outer membrane (262, 264).

The *cag* T4SS PRC stretches from the OMC to the stalk of the *cag* T4SS. This structural feature is not found in minimized T4SSs and so far, has only been identified in

the *H. pylori* and *L. pneumophila* T4SSs (244, 245, 265). Cryo-EM analysis identified the two components of the PRC as portions of CagX and CagY (258). Cryo-ET studies have also identified a stalk structure that extends from the inner membrane to the PRC. Models created based on Cryo-ET analysis of the *cag* T4SS and the *dot/icm* T4SS propose that a channel spans the stalk structure. Indeed, analysis of a central slice through the longitudinal plane of 3D density revealed a possible channel through the center of the stalk. However, because high resolution visualization of the stalk has not been possible, the putative channel could not be clearly visualized (262). Finally, cryo-ET studies revealed the composition of the *cag* T4SS IMC. The *cag* T4SS IMC is composed of the three ATPases CagE, Cagα, and Cagβ and exhibits a 6-fold symmetry (244, 265). In addition to the ATPases, two Vir-like IM proteins (CagW and CagV) and several cytoplasmic subunits specific to the *cag* T4SS are also predicted to localize to the IMC (266). Despite recent advances revealing *cag* T4SS architecture, a large portion of the structure remains unresolved and the mechanisms underscoring substrate transport into host cells remain incompletely defined.

1.3.2 Architectural Asymmetry in T4SSs

Our structural understanding of model T4SS architectures has also revealed intriguing symmetry mismatch at the junction between the IMC and OMCC or between various rings comprising the OMCC that may afford conformational flexibility or secretion channel gating as a mechanism to orchestrate effector selection and directional translocation (101, 153, 249) (**Fig. 1**). Similar symmetry mismatch has been identified in a number of 'expanded' effector translocator systems by *in situ* cryo-electron tomography and single particle reconstruction of isolated T4SS complexes (111, 257, 258, 260, 267-

269), highlighting remarkable structural conservation across the T4SS superfamily (**Fig. 1**).

For example, the *H. pylori cag* T4SS OMCC is comprised of distinct structural features including an outer membrane cap (OMC) consisting of an outer layer (O-layer) and inner layer (I-layer), a periplasmic ring (PR) complex, and a stalk region (244, 250, 257, 258) that assembles with a striking symmetry mismatch occurring between the OMC (14-fold symmetry) and the PRC (17-fold symmetry) (257, 258) (**Fig. 1**). Similarly, the *L. pneumophila dot/icm* T4SS OMC (13-fold symmetry) and PR (18-fold symmetry) displays remarkable mismatch (101, 256). Complex and subassembly asymmetry is noteworthy considering that 'minimized' conjugation systems effectively translocate DNA or proteins to other bacteria and eukaryotic cell targets despite displaying symmetry mismatch between the OMCC and IMC (101). The biological significance of symmetry mismatch between large T4SS subassemblies is unknown; however, asymmetry may afford dynamic structural transitions associated with channel gating or effector secretion that are induced in response to intra- or extracellular stimuli. Within the expanded 'dual purpose' systems, nanomachine conformational flexibility may be required for binding to specific eukaryotic receptors, the recruitment of non-protein substrates to the secretion channel, or orchestrating effector protein delivery across the bacterial envelope (101). Continued structural and mechanistic studies will deepen our understanding of how the fascinating T4SS nanomachine evolved extreme functional diversity and expanded recipient cell range.

1.4 Research Objective

These studies aimed to investigate *cag* T4SS architecture and to identify the function of novel machinery components to deepen our understanding of the mechanisms underscoring trans-kingdom cargo delivery. Here, I identify a novel gating mechanism within the *cag* T4SS translocation channel that regulates the passage of substrates across the bacterial outer membrane. Additionally, I demonstrate that the VirB9 ortholog CagX binds dsDNA via π-π interactions to facilitate secondary DNA substrate selection and transport through the distal *cag* T4SS translocation channel. Finally, I identify several *cag* T4SS components that control the release of effector DNA and protein substrates through the secretion channel and demonstrate that APEX2 proximity labeling techniques can be used to map *cag* T4SS protein-protein interactions *in situ*.

Collectively, this work identifies several novel structural innovations within the *cag* T4SS apparatus periplasmic ring complex that orchestrate DNA substrate selection and enable trans-kingdom conjugation. To our knowledge, these studies represent the first identified function of architecture asymmetry in any microbial nanomachine. These results therefore have far-reaching implications for understanding the molecular mechanisms underpinning interbacterial and trans-kingdom cargo delivery through evolutionarily divergent machineries exhibiting architectural symmetry mismatch including the bacterial T2SS, T3SS, T4SS, and T6SS. Finally, our results provide insight into *cag* T4SS architecture and significantly advance our understanding of how *H. pylori cag* T4SS activity stimulates gastric disease.

CHAPTER 2. CAGC IS A NOVEL GATING MECHANISM IN THE *CAG* T4SS APPARATUS

2.1 Summary

The goal of this work was to identify structural and architectural features that facilitate substrate transfer through *cag* T4SS machinery. *H. pylori* employs *cag* T4SS activity to inject diverse substrates into epithelial cells that alter the mucosal microenvironment and manipulate the inflammatory response. However, the molecular mechanisms underlying substrate transport through the *cag* T4SS apparatus remain incompletely defined. Using techniques to monitor uncontrolled effector release, I show that the pilin ortholog CagC plays a critical role in regulating cargo delivery to target cells. Inactivation of *cagC* results in the mislocalization of protein and DNA effectors to the bacterial cell surface in the absence of host cell contact, suggesting CagC forms a gating apparatus that governs substrate transport across the outer membrane. Protein-protein interaction studies revealed that CagC assembles into homopolymers and interacts with other T4SS components predicted to comprise the inner membrane-embedded apparatus, as well as CagX and CagY within the periplasmic ring complex (PRC). These studies suggest that CagC forms an endopilus-like gating structure that bridges the gap between the IMC and PRC that regulates substrate passage through the translocation channel.

2.2 Introduction

The cancer-associated *H. pylori cag* T4SS translocates the bacterial oncoprotein CagA as well as LPS metabolite and polysaccharide substrates into gastric epithelial cells (27, 44, 270). Recent studies demonstrate the *cag* T4SS-dependent delivery of

chromosomal DNA fragments into target gastric cells to stimulate multiple endosomal and cytosolic nucleic acid surveillance systems (36, 37, 40). Currently, the process by which DNA is excised and trafficked through the *cag* T4SS translocation channel is incompletely defined. Recent work demonstrated that trans-kingdom DNA conjugation is linked to chromosomal replication and replichore decatenation and identified a VirD2-like endonuclease that is essential for recruitment to the *cag* T4SS cytoplasmic apparatus (38). However, the mechanism by which DNA is transported through the periplasm and across the bacterial outer membrane is unresolved.

The *cag* T4SS comprises approximately 27 components, the majority of which are unique to *H. pylori* and are required for *cag* T4SS activity (1, 22, 24). Several *cag* T4SS proteins share homology with components of the prototypical *Agrobacterium tumefaciens vir* T4SS (1, 266) that translocates tumor-inducing DNA (T-DNA) and associated effector proteins into recipient plant cells in a contact-dependent manner (97-100, 102, 105). Previous studies using a transfer DNA immunoprecipitation (TrIP) assay (102) to trace the effector translocation pathway through the *vir* T4SS secretion channel significantly advanced our understanding of how apparatus components coordinate nucleoprotein delivery into eukaryotic cells. Additional studies in *A. tumefaciens* identified gating mechanisms within the *vir* T4SS that regulate substrate transport to the bacterial cell surface. Due to the well-defined mechanism of *vir* T4SS-mediated DNA transport and the presence of several homologous components, the *A. tumefaciens vir* T4SS serves as a reliable model for investigating the mechanism of *cag* T4SS substrate translocation.

In this work, I establish that *vir* T4SS homologs are essential for *cag* T4SS-dependent DNA delivery into host cells and provide evidence that in contrast to *A. tumefaciens*,

translocated *H. pylori* DNA is derived from the chromosome and is double-stranded. Furthermore, I show that effector DNA is exposed to the extracellular milieu before transport into host cells, suggesting that *cag* T4SS-dependent DNA translocation occurs via a two-step mechanism that does not require a pilus-like conduit. Additionally, using methods developed to study *A. tumefaciens vir* T4SS translocation channel integrity, I demonstrate that the pilin ortholog CagC forms a gating mechanism that regulates the passage of DNA and protein effectors across the bacterial outer membrane. I provide evidence that CagC interacts with components of the *cag* T4SS inner membrane (IMC) and periplasmic ring (PRC) complexes and propose a model in which CagC polymers dynamically assemble to form a periplasmic endopilus-like structure that regulates substrate passage through the distal *cag* T4SS translocation channel. Collectively, these findings identify the role of CagC in the secretion system and broadens our understanding of how the *cag* T4SS assembles and functions to deliver diverse effector molecules into host cells.

2.3 Materials and Methods

Bacterial strains and culture conditions

Helicobacter pylori 26695 and isogenic derivatives were grown on trypticase soy agar plates supplemented with 5% sheep blood (BD) at 37˚C with 5% CO_2. Overnight cultures of *H. pylori* were grown in Brucella broth supplemented with 5% fetal bovine serum (FBS) at 37°C in 5% CO_2. *E. coli* strain DH5α (New England Biolabs) was used for plasmid propagation and was grown on lysogeny broth (LB) agar or liquid media supplemented

with the appropriate antibiotics for plasmid maintenance. Bacterial two-hybrid (BACTH) strain BTH101 (a gift from Scot Ouellette, University of Nebraska Medical Center) was maintained on LB, and BACTH screening was performed using LB plates supplemented with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), 100 μg/mL ampicillin, 50 μg/mL kanamycin, and 20 μg/mL 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside (X-Gal)(271). BACTH-dependent growth selection was performed in M63 minimal media [$(NH_4)_2SO_4$ (2 g/L), KH_2PO_4 (13.6 g/L), Thiamine B1 (1 mg/L), 1 mM $MgSO_4$, $FeSO_4.7H_2O$ (5 mg/L), pH 7] supplemented with 0.4% maltose, 0.5 mM IPTG, 50 μg/mL ampicillin, and 15 μg/mL kanamycin. All BACTH experiments were executed at 30°C under constant aeration.

H. pylori mutagenesis

Isogenic mutants (251, 272-274) were complemented in *cis* at the *ureA* locus using plasmids derived from pAD1(275) that were designed to express either the native gene or a protein variant as previously described(64, 275). Point mutants were generated using Q5 Site Directed Mutagenesis (New England Biolabs). Plasmid sequences were confirmed by PCR and whole plasmid sequencing, and constructs were used to transform isogenic mutant strains. Colonies resistant to kanamycin (12.5 μg/ml) or chloramphenicol (10 μg/mL) were selected and complementation at the *ureA* locus was confirmed by PCR and by anti-CagX immunoblotting, when appropriate.

Human cell culture

HEK293-hTLR9 (Invivogen hkb-htlr9), the corresponding parental HEK293 null1 (InvivoGen hkb-null1) cell lines were grown in DMEM media supplemented with 10%

heat-inactivated FBS and 1X GlutaMAX (Life Technologies) in the presence of 5% CO_2 at 37°C. AGS human gastric epithelial cells (ATCC CRL-1739) were grown in RPMI media supplemented with 10% FBS, 2 mM L-glutamine, and 10 mM HEPES in the presence of 5% CO_2 at 37˚C.

TLR9 stimulation assay

TLR9 activation assays were carried out as previous described (276, 277). Briefly, HEK293 cells expressing human TLR9 (HEK293-hTLR9) or the parental HEK293 null1 cells were co-cultured with WT *H. pylori* 26695 or *cag* mutant strains at a MOI of 100. Supernatants were collected at 24 hours post-infection, and TLR9 activation was quantified by measuring secreted embryonic alkaline phosphatase (SEAP) in cell culture supernatant by QuantiBlue™ reagent (Invivogen) using a microplate reader (BioTek Synergy HI) to record the absorbance at 650 nm. TLR9 activation was normalized to levels of SEAP produced by the corresponding infected null1 cells and data were expressed as a percent of the normalized fold change over WT-challenged cells.

Colony immunoblotting

To assess surface localization of bacterial effectors (278), *H. pylori* strains were grown on blood agar plates for 24 h before collection and normalization to OD_{600} of 1. Bacteria were pelleted by centrifugation and resuspended in 50 μL Brucella broth to reach an OD_{600} of 20. Concentrated suspensions were spotted onto fresh blood agar plates (5 μL) and incubated at 37˚C, 5% CO_2. After 24 h, a nitrocellulose membrane was placed on top of the colonies for 5 min at room temperature. The membrane was washed twice in TBST (tris-buffered saline, 0.05% Tween 20) and blocked in 5% milk in TBST. Surface exposure

of bacterial effectors was assessed by immunoblotting using anti-CagA (α-CagA, Santa Cruz Biotechnology), anti-dsDNA (α-dsDNA, EMD Millipore). Immunoblotting with anti-*H. pylori* antibody was performed as a positive control.

Live cell bacterial immunocapture

H. pylori was grown for 24 h on blood agar plates prior to harvesting and normalization to an OD_{600} of 0.5 in Brucella broth supplemented with 6% FBS. Normalized bacteria were incubated with 250 μg/mL anti-dsDNA monoclonal antibody (α-dsDNA, EMD Millipore) or an equivalent amount of IgG isotype control antibody for 1 h in the presence of 5% CO_2 at 37˚C. After incubation, an aliquot was obtained for colony forming unit (CFU) enumeration by serial dilution plating. The remaining sample was incubated with Protein G Dynabeads (Invitrogen) for 10 minutes at RT with constant rotation. Beads were isolated by magnetic separation and washed twice with 5 bead volumes of sterile PBS. Beads were then resuspended in 100 μL sterile PBS followed by serial dilution and plating for CFU enumeration. For DNase treated samples, *H. pylori* pellets were resuspended in PBS and normalized to an OD_{600} of 1 prior to incubation with 5 μL DNase buffer and 1 unit Turbo DNase I (Life Technologies). After 15 min incubation, an equal volume of Brucella broth supplemented with 6% FBS was added to each well to reach an OD_{600} of 0.5, and bacterial immunocapture was performed. Levels of *H. pylori* that display double-stranded DNA on the bacterial cell surface were calculated as a percentage of total bacteria in each sample and normalized to levels of bacteria isolated in IgG isotype control purifications.

Erythromycin susceptibility assays

H. pylori strains cultured for 24 h on blood agar plates were collected, normalized to an OD_{600} of 1, and an equivalent volume of normalized cultures were inoculated onto fresh blood agar by sterile bead plating. Erythromycin-impregnated discs (15 μg) were placed onto dried plates, and the diameter of the zone of inhibition was determined after 24 h incubation at 37˚C and 5% CO_2. Susceptibility assays were performed a minimum of three times for each *H. pylori* strain.

'Transfer DNA' immunoprecipitation

Transfer DNA assays were performed as previously described (38). Briefly, *H. pylori* protein-DNA complexes cross-linked by the addition of 500 μL 1% paraformaldehyde followed by incubation for 10 min at room temperature and quenching with 1 ml 250 mM glycine. Cross-linked cells were lysed in lysis buffer (5 mM EDTA, 1% NP-40 in PBS) supplemented with 2X cOmplete™ inhibitor (63) by sonication. Cell lysates were treated with 2 units of Turbo DNase (Life Technologies) and 10 mM $MgCl_2$ for 30 min at room temperature. Solubilize membranes were separated by centrifugation at 14,000 rpm for 30 minutes at 4˚C. CagY complexes were purified using polyclonal anti-CagY antisera conjugated to Protein G Dynabeads (Invitrogen) isolated by magnetic separation. Purified complexes were washed twice in PBS supplemented with 10 mM $MgCl_2$ and 1 unit Turbo DNase, followed by two washes in high salt buffer (lysis buffer containing 400 mM NaCl), and a final wash in PBS. Following the washes, Protein-DNA complexes were eluted and de-cross-linked in 100 μl 1% SDS in 0.1 M $NaHCO_3$ at 65˚C overnight. Proteins were digested using Proteinase K (10 μg) at 65˚C for 30 minutes, and DNA was precipitated by 100% ethanol in 0.3 M sodium acetate (1:3 v/v) at -20˚C. To serve as a control, the initial 'input' cell pellet was re-suspended in 100 μl of 50 μM NaOH and incubated at 95˚C for

30 min, followed by pH neutralization by the addition of 10 µl of 1M Tris, pH 8. Standard PCR assays using the 'input' and 'IP' samples as templates and targeting a 795 bp chromosomal DNA amplicon were performed to assess bacterial DNA associated with immunopurified CagY complexes. Quantitation of DNA amplification from 'input' and 'IP' samples was conducted by densitometry analysis in Fiji (279) and amplification efficiency of 'IP' samples was calculated as a percent of the corresponding 'input' sample amplification for each biological replicate experiment.

Bacterial adenylate cyclase two-hybrid (BACTH) assay

To screen direct protein-protein interactions, I used a bacterial two-hybrid system in which the bait protein of interest was fused to the either the T18 N-terminus (pUT18C vector) or the T25 N-terminus (pKNT25 vector) as previously described (271, 280-282). Plasmids encoding various prey proteins were cloned into both the pUT18C and pKNT25 vectors. Recombinant plasmids were verified by whole plasmid DNA sequencing. Both bait and prey proteins were cloned into *E. coli* BTH101 (Δ*cya*) and transformants were selected on LB plates supplemented with 100 µg/mL ampicillin and 50 µg/mL kanamycin at 30°C. Empty T18 and T25 vectors were co-transformed to serve as a negative control while T18-Zip/T25-Zip (271) and T18-DprA/T25-DprA (15) were co-transformed to serve as positive controls. Single colonies were selected, grown overnight at 30˚C in LB supplemented with dual antibiotics (100 µg/mL ampicillin and 50 µg/mL kanamycin) and 0.5 mM IPTG, and spotted onto LB plates supplemented with dual antibiotics, 0.5mM IPTG, and 40 mg/mL X-gal. Plates were monitored for changes in colony color (blue indicating positive protein-protein interaction, white indicating negative protein-protein interactions). Growth in M63-0.4% maltose was used to confirm protein-protein interactions. For these studies, 2.5 µL

of LB overnight cultures were inoculated into fresh M63-maltose minimal media and were grown under aerobic conditions at 30˚C for 72 hours. Bacterial growth was determined by measuring the OD_{600} (Biotek Synergy H1) and compared to the positive and negative control strains. Culture growth was expressed as the fold change in OD_{600} over the corresponding negative control culture (BTH101 co-transformed with empty vectors), and data represent a minimum of at least three independent biological replicate experiments for each pairwise co-transformation.

Protein structure modeling

CagC predicted structure was determined using AlphaFold2 colab (2, 283). Protein alignments were performed using MegAlign Pro (DNAstar). Superimposed images of protein structures were generated using ChimeraX Matchmaker (284).

Immunoblotting

To detect bacterial protein expression, normalized samples were lysed in reducing 2X SDS buffer (Bio-Rad), resolved by SDS-PAGE (10%), transferred to a nitrocellulose membrane, and immunoblotted using rabbit polyclonal antisera raised against CagC peptides (273) (a gift from Dr. Tim Cover) as previously described (275). To confirm equal sample loading, membranes were immunoblotted using *H. pylori* outer membrane protein monoclonal antibody (α-OMP, Santa Cruz). Secondary detection by horse radish peroxidase-conjugated anti-rabbit IgG or anti-mouse IgG was achieved using SuperSignal West Pico chemiluminescent substrate (Thermo).

IL-8 quantitation

Secretion of IL-8 by AGS cells co-cultured with *H. pylori* or corresponding isogenic derivatives was determined using the human CXCL8 ELISA (R&D Systems). Briefly, AGS cells were challenged with *H. pylori* or the indicated isogenic mutant strain for 4.5 h at an MOI of 100. Supernatants were collected and analyzed as previously described (64, 275). Biological replicate experiments were performed a minimum of three times.

2.4 Results

Chromosomal DNA is translocated into host cells via a unique mechanism

The *cag* T4SS harbors multiple *vir* T4SS homologs including CagX, CagY, CagT, CagV, CagE, Cagα, CagL, CagC, and Cag5 (VirB9, VirB10, VirB7, VirB8, VirB4, VirB11, VirB5, VirB2, and VirD4, respectively) (1, 105, 110, 125, 250, 251, 257, 258, 274) (Table 1). Prior work identified 'uncoupling' mutations within the *A. tumefaciens vir* T4SS that selectively disarm T-pilus biogenesis without perturbing trans-kingdom conjugation. Similarly, various *H. pylori cag* T4SS-dependent phenotypes, including CagA translocation and IL-8 stimulation, occur via uncoupled mechanisms that require incongruent subsets of *cag* T4SS components, raising the hypothesis that molecularly dissimilar *cag* T4SS substrates are secreted into host cells via distinct mechanisms. To test this hypothesis, I first analyzed the contribution of *vir* T4SS homologs in TLR9 stimulation, an outcome that requires a functional *cag* T4SS (36, 38, 64, 251). Consistent with previous studies analyzing *cag* T4SS-dependent phenotypes (1, 40, 64, 251), TLR9 stimulation occurred in a CagA-independent manner(36, 38) that required recognized *vir* T4SS homologs, with the exception of the VirD4-like coupling protein Cag5 (251),

Table 2.1 Prototypical *A. tumefaciens vir* T4SS homologs in the *cag* T4SS

Gene Number[a]	Protein Name	*vir* T4SS homolog[b]
hp0523	Cag4	VirB1
hp0524	Cag5/Cagβ	VirD4
hp0525	Cagα	VirB11
hp0527	CagY	VirB10
hp0528	CagX	VirB9
hp0532	CagT	VirB7
hp0529	CagW	VirB6
hp0530	CagV	VirB8
hp0539	CagL	VirB5
hp0544	CagE	VirB3/B4
hp0546	CagC	VirB2

[a]Gene number in *H. pylori* strain 26695
[b]Based on the *A. tumefaciens* strain C58 *vir* T4SS nomenclature

suggesting that DNA trafficking through the secretion channel occurs via non-canonical trans-kingdom conjugation (**Fig. 2a**).

To further investigate the mechanism of *cag* T4SS DNA translocation, I next performed studies to determine if translocated *H. pylori* DNA is protected from the extracellular environment by assessing whether exogenous nuclease treatment would compromise DNA transport and impair TLR9 activation. In agreement with a previous study (36), *H. pylori*-driven TLR9 activation decreased by approximately 50% when DNase I was added to co-cultures concomitantly with *H. pylori* compared to levels achieved in the absence of exogenous nuclease (**Fig. 2b**). Because DNase I non-specifically cleaves single- and double-stranded DNA, as well as RNA:DNA hybrids, I next tested whether nucleases that specifically degrade either double-stranded DNA (dsDNase) or single-stranded DNA (exonuclease I) would impact *cag* T4SS-dependent DNA translocation. TLR9 activation was significantly impaired by the addition of exogenous dsDNase but was unaffected by the presence of exonuclease I or RNase A, supporting the hypothesis that translocated *H. pylori* DNA is double-stranded and that it is exposed to the extracellular space prior to transport into host cells. Additionally, I analyzed the ability of monoclonal anti-DNA antibodies to functionally block TLR9 activation. Compared to isotype control treated co-cultures, *H. pylori*-driven TLR9 activation was significantly reduced in the presence of either anti-dsDNA or, to a lesser degree, anti-ssDNA monoclonal antibodies (**Fig. 2c**), a result likely due to inherent antibody cross-reactivity. Bacterial DNA localization to an extracellular site prior to transport into host cells was further analyzed by labeling the *H. pylori* chromosome via bromodeoxyuridine (BrdU) incorporation. Compared to unlabeled WT bacteria, anti-BrdU antibodies significantly

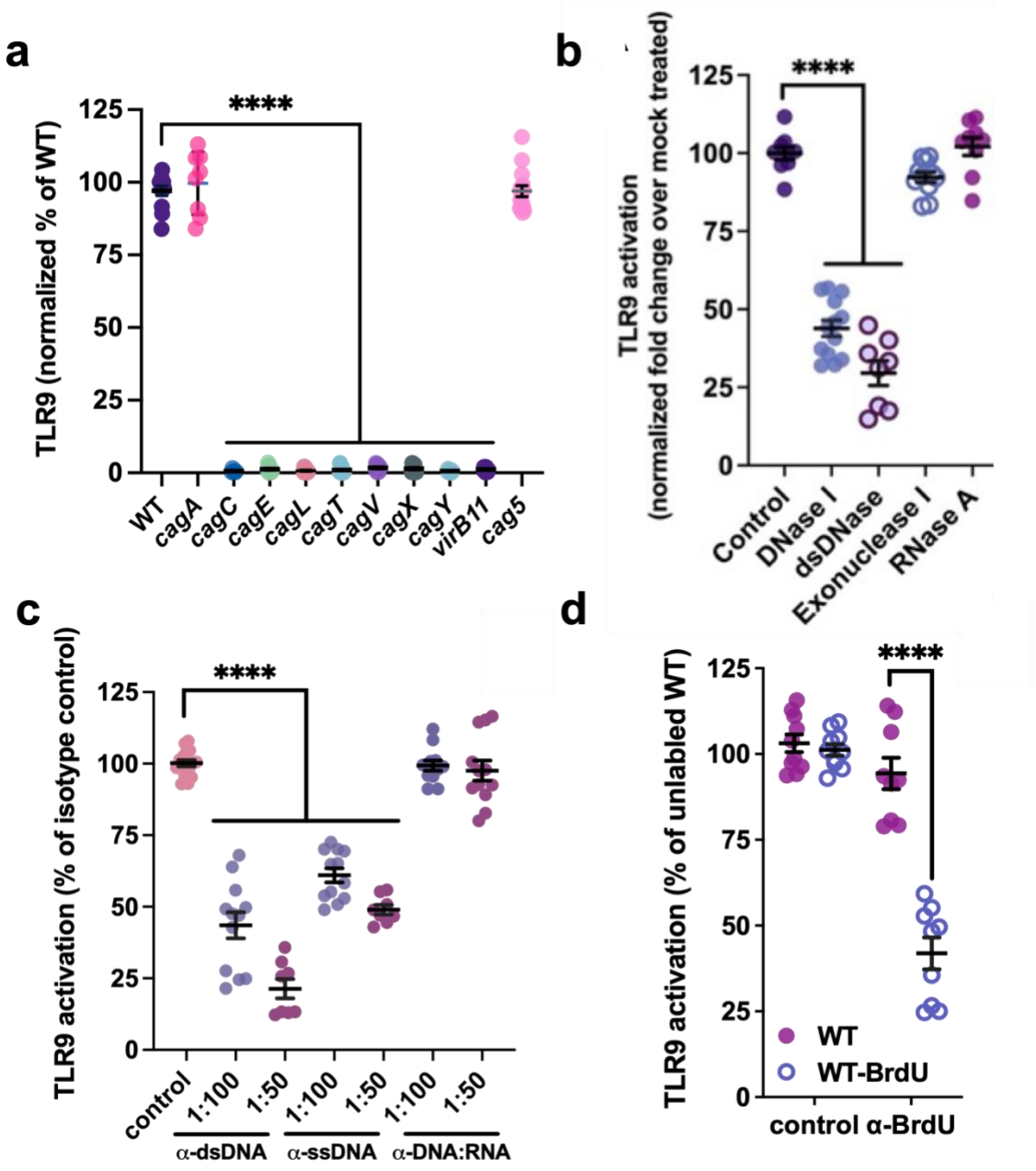

Figure 2. *H. pylori* DNA is delivered to host cells via a novel secretion mechanism.

(**a**) Levels of *cag* T4SS-dependent TLR9 activation induced by the indicated *H. pylori* 26695 isogenic mutant strain. Data are expressed as the normalized fold change over mock infected cells as a percent of TLR9 stimulation achieved by WT. **(b)** TLR9 activation is susceptible to exogenous nucleases. **(c)** *cag* T4SS-dependent TLR9 activation is impaired in the presence of monoclonal antibodies targeting dsDNA and ssDNA, but not DNA:RNA hybrids. **(d)** Anti-BrdU monoclonal antibodies inhibit TLR9 activation by BrdU-labeled *H. pylori*. Significance was determined by one-way ANOVA with Dunnett's post-hoc correction for multiple comparisons to experimental controls; ****, $p<0.0001$ in all panels.

impaired TLR9 activation by *H. pylori* that had incorporated BrdU nucleoside analogs into the chromosome (**Fig. 2d**). Collectively, these results suggest that *H. pylori* DNA cargo is at least partially double-stranded and localized to an extracellular site prior to entering host cells, suggesting that DNA translocation by the cag T4SS is occurring via a unique mechanism. Previous studies demonstrate that CagA translocation is partially blocked in the presence of anti-CagA monoclonal antibodies (285), further supporting a model whereby *cag* T4SS cargo is transported to the bacterial cell surface prior to delivery to the gastric cell cytoplasm.

CagC gates effector translocation through the *cag* T4SS

I next analyzed potential structural features that could facilitate DNA substrate selection for export through the *cag* T4SS apparatus. Previous work identified a single point mutation within the *A. tumefaciens vir* T4SS OMC that confers secretion channel gating defects resulting in the uncontrolled release of VirE2 effectors to the bacterial cell surface (278). Because the *cag* T4SS apparatus contains several *vir* T4SS homologs, I hypothesized that similar gating mechanisms may be occurring within the *cag* T4SS (251, 257, 258, 274). To test whether individual *cag* T4SS components are associated with apparatus gating, I employed non-denaturing colony blot assays developed to interrogate translocation channel integrity in *A. tumefaciens* (278). In contrast to WT *H. pylori*, which does not release detectable levels of *cag* T4SS effector molecules into the extracellular milieu, isogenic mutants deficient in *cagC* accumulated surface-exposed CagA and transfer DNA in a host cell contact-independent manner (**Fig. 3a**). Genetic complementation of the *cagC* isogenic mutant in a heterologous chromosomal locus abrogated effector leakage (**Fig. 3a**) and restored *cag* T4SS-dependent TLR9 activation (**Fig. 3b**) and IL-8 stimulation

phenotypes (**Fig. 3c**), suggesting that CagC regulates channel gating or substrate passage across the outer membrane. In support of the hypothesis that CagC forms a specialized plug or gating mechanism, inactivation of other reported structural (CagX, CagY, CagT, CagV, CagL) or energetic (CagE, Cagα, Cag5) components did not affect *cag* T4SS translocation channel integrity (**Fig. 3a**).

To quantitatively monitor the uncontrolled release of effector DNA into the extracellular milieu, I developed a live cell assay to enumerate *H. pylori* that accumulate surface-exposed transfer DNA in the absence of host cell contact. For these studies, WT *H. pylori* or the indicated isogenic mutant strain was incubated with anti-DNA monoclonal antibodies or an equivalent volume of a corresponding IgG isotype control, and immunocaptured bacteria were isolated by Protein G-conjugated magnetic bead separation. Compared to WT and the *cagY* strain, which does not assemble the *cag* T4SS core complex (257, 258, 274) and thus does not exhibit effector leakage (**Fig. 3a**), significantly more *cagC* bacteria were immunocaptured by anti-dsDNA antibodies (**Fig. 3d**). Consistent with a role in effector gating, genetic complementation of the *cagC* mutant restored immunocaptured bacteria to WT levels (**Fig. 3d**). Likewise, incubating bacterial cultures with DNase I prior to anti-dsDNA immunocapture markedly reduced levels of isolated *H. pylori* (**Fig. 3d**) demonstrating that aberrantly localized DNA is tethered to the bacterial cell surface.

I next sought to determine whether *cag* T4SS pore gating defects afford unrestricted access to the periplasm. I reasoned that loss of either secretion channel integrity or pore gating would license the influx of small molecules, such as erythromycin, that are typically impermeable to the Gram-negative outer membrane (286, 287). To test whether potential

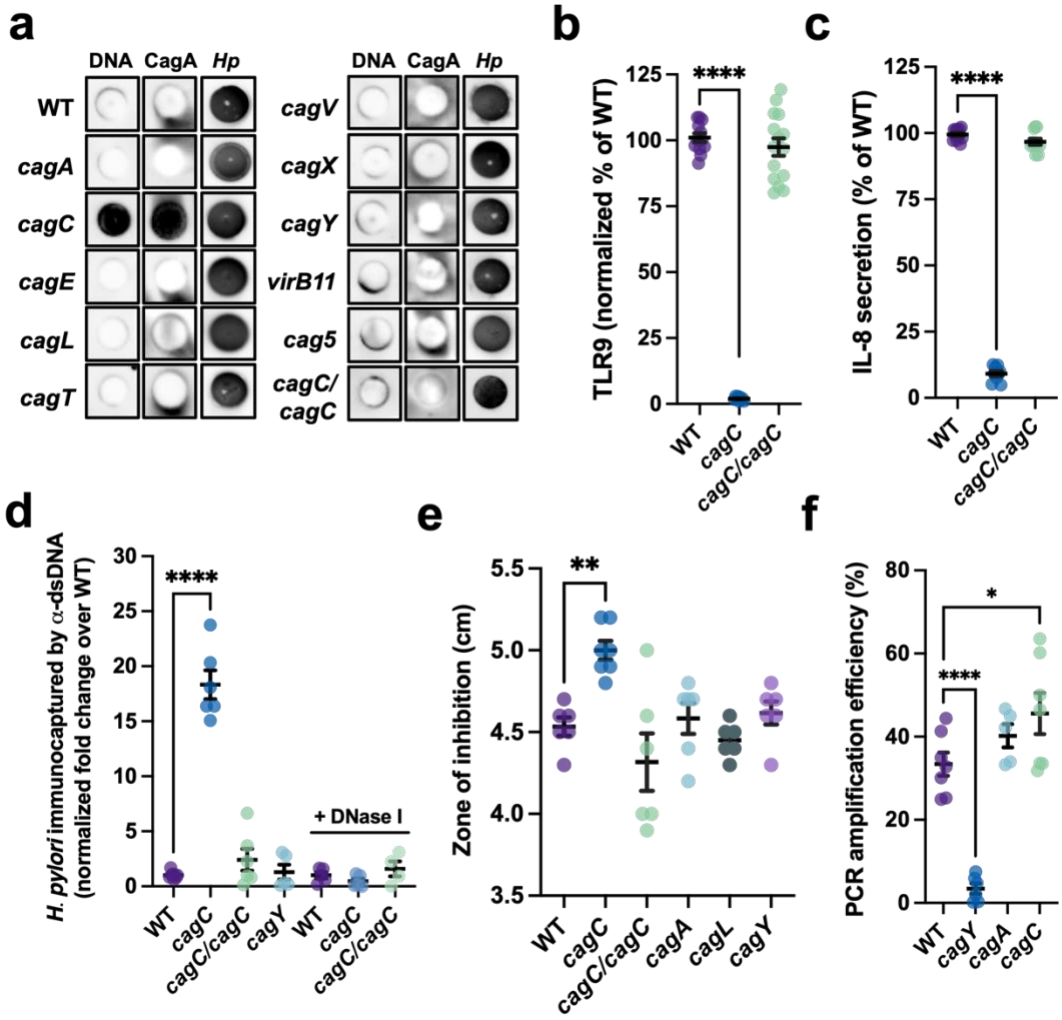

Figure 3. CagC gates the *cag* T4SS translocation channel.

(**a**) Non-denaturing colony blots probing for surface-exposed dsDNA, CagA, and total *H. pylori* antigen (1). (**b**) TLR9 stimulation and (**c**) IL-8 secretion requires *cagC*. (**d**) Levels of live *H. pylori* immunocaptured by anti-dsDNA monoclonal antibodies. Data are expressed as the normalized fold change in recovered colony forming units over levels obtained for WT. (**e**) Erythromycin susceptibility of the indicted strain. Data depict the zone of inhibition for the indicated strain. (**f**) Transfer DNA immunopurification assays demonstrating the presence of *H. pylori* chromosomal DNA fragments within the *cag* T4SS apparatus. Graph depicts the amplification efficiency of a 795 bp fragment in transfer DNA assay preparations purified from the indicated strain. Amplification efficiency of the immunopurification samples is expressed as the percent of the input DNA from at least four biological replicate experiments. In **b-f**, significance was determined by one-way ANOVA with Dunnett's post-hoc correction for multiple comparisons to experimental controls; *, $p<0.05$, **, $p<0.01$, ****, $p<0.0001$.

cag T4SS pore gating mechanisms control periplasm accessibility, I employed erythromycin sensitivity assays to detect changes in outer membrane permeability (286). Compared to WT and isogenic mutants that do not exhibit effector molecule leakage, the *cagC* strain displayed significantly increased pore permeability (**Fig. 3e**) indicating that CagC is important for stabilizing the secretion channel in a tightly closed conformation. Erythromycin susceptibility patterns of the *cagC* complemented mutant phenocopied the parental WT strain (**Fig. 3e**), further demonstrating that CagC controls pore gating and access to the periplasm via leaky secretion channels.

Previous work determined that effector DNA is pre-loaded into the *cag* T4SS apparatus prior to initiating host cell contact to generate a 'ready-to-fire' nanomachine (38). To test whether defective gating modulates DNA loading, I used a modified transfer DNA immunoprecipitation assay developed in *A. tumefaciens* (102) to monitor levels of DNA confined within the *cag* T4SS core complex lumen (38). For these studies, I isolated *cag* T4SS assemblies from chemically cross-linked *H. pylori* via immunopurification targeting CagY (a constituent of the *cag* T4SS outer membrane-associated complex and periplasmic ring structure (257, 258, 274)), and monitored DNA loading by PCR analysis of co-purifying chromosomal fragments that remained shielded from DNase degradation within the secretion channel. In agreement with a previous report (38), DNA fragments were readily amplified from cross-linked CagY preparations obtained from WT and *cagA*-deficient *H. pylori*, but not from mock preparations generated from the *cagY* isogenic strain (**Fig. 3f**). Consistent with a role in regulating DNA passage through the *cag* T4SS OMC, significantly more DNA co-purified with CagY complexes in strains lacking *cagC* (**Fig. 3f**), demonstrating that CagC is not required for effector DNA loading into the secretion

channel. Collectively, these studies demonstrate that loss of CagC confers *cag* T4SS channel leakiness that enables the unrestricted export of effectors and the import of exogenous small molecules into the periplasm.

CagC is a specialized T4SS pilin ortholog

I next performed protein sequence analysis and structural modeling to further investigate the role of CagC as a novel gating mechanism within the *cag* T4SS apparatus. In accordance with a previous study analyzing pilin-like motifs in bacterial cell surface-exposed polypeptides (288), CagC exhibited primary sequence similarity to canonical T4SS pilins including *A. tumefaciens* VirB2 and *E. coli* TraA (**Fig. 4a**). AlphaFold2 (2) modeling of the mature CagC monomer lacking the leader peptide (2, 283, 289) revealed a predicted helical structure that closely resembles the experimentally resolved structures of conjugative pilins VirB2 (161, 162) and TraA (290) (**Fig. 4b**). In contrast to a previous report suggesting that conjugative pilins are cyclic in nature (291), superimposition of predicted CagC structures with monomeric VirB2 and TraA revealed a slightly U-shaped protomer in which the free pilin termini were oriented towards the exterior of assembled T-pilus or F-pilus filaments (159, 161, 162) (**Fig. 4b**). Based on the predicted structural conservation, I hypothesized that orthologous pilins could rescue *cag* T4SS defects exhibited by the *cagC* isogenic mutant. To test this hypothesis, I expressed either *A. tumefaciens virB2* or *E. coli traA* in the *cagC* isogenic mutant and analyzed *cag* T4SS-dependent outcomes elicited by each chimera. Despite predicted structural similarities, genetic complementation of *cagC* with either *virB2* or *traA* did not restore *cag* T4SS pore gating (**Fig. 4c**), TLR9 activation (**Fig. 4f**), or IL-8 stimulation (**Fig. 4g**) defects, suggesting that CagC fulfills a unique role within the *cag* T4SS. One possibility is that CagC forms an

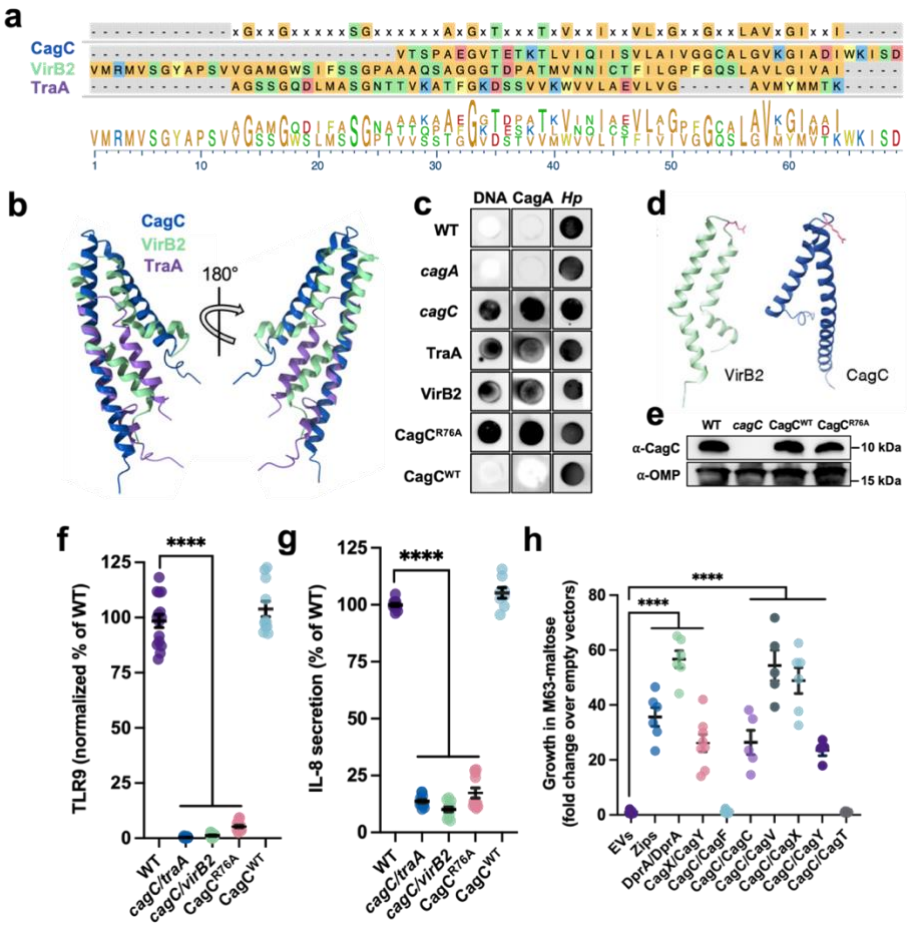

Figure 4. CagC is a VirB2-like T4SS pilin ortholog.

(**a**) Protein sequence alignment of prototypical T4SS pilins VirB2 (*A. tumefaciens vir* T4SS), TraA (*E. coli* pKM101), and CagC. Pilin sequence alignment was generated with MegAlign Pro Software. (**b**) Overlay of CagC, VirB2 (T-pilus), and TraA (F-pilus) using ChimeraX matchmaker function. CagC structure was predicted using AlphaFold2 using the mature protein sequence. (**c**) Colony blots demonstrating dsDNA and CagA cell surface localization in *H. pylori cagC* mutants complemented by either *traA* or *virB2*. (**d**) Structures of VirB2 (PDB: 8CUE) and CagC (AlphaFold2 modeling) highlighting Arg 91 (VirB2) and Arg 76 (CagC) residues in the luminal pilin loop region (pink coloring). (**e**) Western blot analysis of *H. pylori cagC* mutants genetically complemented by $CagC^{WT}$ or the $CagC^{R76A}$ variant. OMP, outer membrane protein. (**f**) Levels of TLR9 stimulation and (**g**) IL-8 secretion induced by the indicated strain. **(h)** Bacterial two-hybrid analysis of CagC protein-protein interactions. Data depict the fold change in the OD_{600} of *E. coli* BTH101 cultures transformed with the indicated plasmid pairs and propagated in M63-maltose. Positive protein-protein interactions are indicated by growth in M63-maltose and is expressed as fold change over BTH101 transformed with empty vectors (EVs). In **f-h**, significance was determined by one-way ANOVA with Dunnett's post-hoc correction for multiple comparisons to experimental controls; ****, $p<0.0001$.

‘endopilus’ within the secretion channel lumen corresponding to the stalk-like density observed in the *in situ cag* T4SS architecture (244, 250). Recent studies analyzing VirB2 protomers identified arginine residues (Arg 91) that protrude into the T-pilus lumen and form extensive electrostatic interactions required for pilus assembly and structural stabilization (161, 162) (**Fig. 4d**). Congruent with a predicted function as a specialized pilin, CagC harbors a single arginine residue in the mature protomer (Arg 76) corresponding to VirB2 Arg 91 (**Fig. 4d**). I reasoned that similar to T-pilus biogenesis, polymerization of a putative ‘endopilus’ could be selectively disrupted by modifying positively charged residues within the predicted CagC luminal loop (**Fig. 4b,d**). To investigate whether Arg 76 is required for CagC function, I generated a CagC variant containing a R76A single point mutation expressed from a heterologous chromosomal locus in the *cagC* mutant. Western blot analysis of *H. pylori* whole cell lysates reveled high levels of $CagC^{R76A}$ expression (**Fig. 4e**), demonstrating that in contrast to VirB2 Arg 91 (161, 162), CagC Arg 76 does not contribute to protomer stability. Consistent with a role as a putative ‘endopilus’ or gating mechanism, $CagC^{R76A}$ exhibited uncontrolled DNA and effector protein release to the bacterial cell surface (**Fig. 4c**). Accordingly, compared to the parental WT strain and *cagC* mutants genetically complemented with $CagC^{WT}$, $CagC^{R76A}$ variants were defective in TLR9 activation (**Fig. 4f**) and IL-8 stimulation (**Fig. 4g**), indicating that positively charged residues within the luminal loop are important for CagC protomer function or multimerization into higher order structures.

I next sought to identify CagC protein interaction partners to determine where CagC gating occurs within the *cag* T4SS apparatus. In the mature form, CagC pilin subunits are approximately 9.2 kDa (**Fig. 4e**) and attempts to generate a functional epitope-tagged

fusion protein for immunopurification studies were unsuccessful. Therefore, I employed a bacterial two-hybrid (BTH) screening approach (271, 292, 293) designed to characterize interactions among both soluble and membrane-associated components to localize CagC complexes within the *cag* T4SS apparatus. Analysis of direct protein-protein interactions revealed that in comparison to negative (empty vector) and positive (Zip-Zip (271), *H. pylori* DprA-DprA (15)) controls, CagC strongly interacted with components predicted to comprise the T4SS apparatus inner membrane complex (IMC) including the bitopic VirB8 homolog CagV (**Fig. 4h**). BTH assays also revealed strong interactions among the CagX N-terminal domain (residues 41–310) and corresponding CagY fragments (residues 1469–1603, designated $CagY_{CT}$) that assemble into the periplasmic ring complex (PRC) (257, 258, 274), demonstrating the ability of our screening approach to detect direct protein-protein interactions occurring in the periplasm. In addition to CagC-CagV, I detected direct CagC-CagC protomer interactions (**Fig. 4h**), suggesting that CagC can self-polymerize into higher order structures or subcomplexes. Whereas interactions among CagC and components localized to either the cytoplasm (the chaperone-like molecule CagF (294)) or the periphery of the *cag* T4SS outer membrane complex (the VirB7 homolog CagT (257, 258, 274)) were not detected, CagC directly interacted with both CagX and $CagY_{CT}$ (**Fig. 4h**). These observations led to the hypothesis that CagC subassemblies bridge the IMC and PRC. Because other T4SS pilin subunits including VirB2 and TraA assemble as pentamers within their respective secretion systems (159, 161), I predict that CagC will similarly assemble as pentamer within the *cag* T4SS. Using AlphaFold-multimer (3, 283), I generated a model of pentameric CagC assembled as a pilus-like structure (**Fig. 5a**). Docking of the assembled structure into the *cag* T4SS apparatus reveals that the predicted

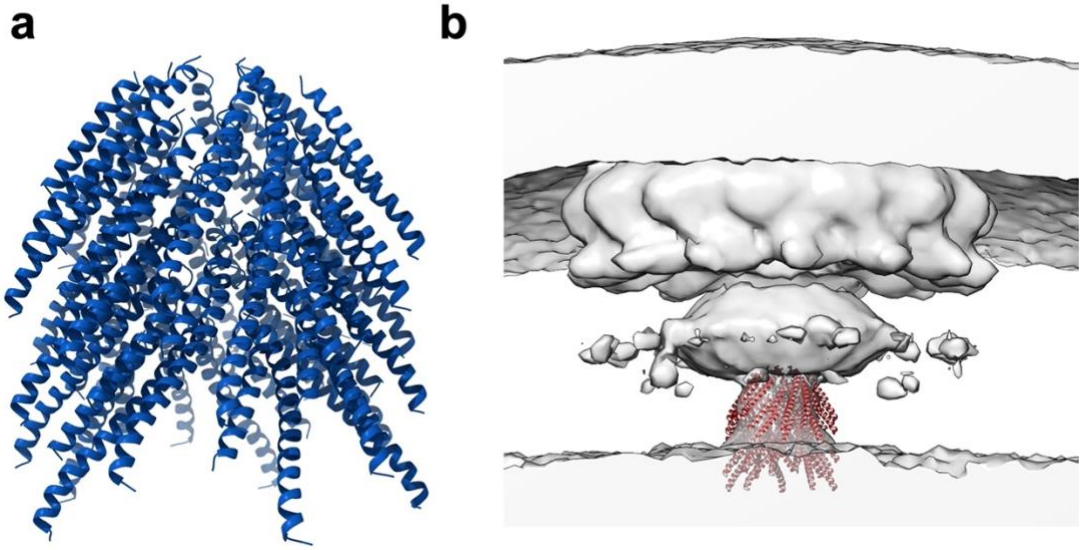

Figure 5. Pentameric CagC assembles into a predicted 'endopilus' in the *cag* T4SS.

(a) Predicted assembly of pentameric CagC as an 'endopilus' generated by AlphaFold-Multimer (3, 283). (b) Predicted 'endopilus' structure docked into the cryo-ET visualized *cag* T4SS (244) (EMBD-7474).

structure is approximately the correct shape and size as the unresolved stalk region visualized *in situ* by cryo-ET (244, 250) (**Fig. 5b**), suggesting the possibility that CagC forms an 'endopilus' corresponding to the stalk-like region that may dynamically assemble and retract to regulate effector gating within the PRC via a piston-like mechanism. Collectively, these data suggest a model whereby CagC polymers into a novel plug or gate-like apparatus to control effector passage through the secretion system pore.

2.5 Discussion

Understanding how the *cag* T4SS delivers immunostimulatory effectors into host cells provides mechanistic detail underscoring how *H. pylori* stimulates severe gastric disease. Compared to minimized conjugative T4SS machineries, the expanded *cag* T4SS architecture and the presence of multiple unique components suggests specialized substrate selection and delivery mechanisms. In prototypical T4SSs, plasmid-borne DNA substrates are single-stranded and are postulated to transit through an enclosed pilus during transport into recipient cells (12, 39, 101, 104). In contrast, I provide multiple lines of evidence that *cag* T4SS effector DNA is double-stranded and derived from the bacterial chromosome and present compelling evidence that *cag* T4SS-dependent DNA translocation is susceptible to exogenous DNA-targeting nucleases and anti-DNA monoclonal antibodies **(Fig. 2).** Jointly, these findings suggest a multi-step translocation mechanism that does not appear to require a conjugative pilus conduit and highlight unique DNA translocation mechanisms employed by the *cag* T4SS.

Additionally, this work revealed a novel CagC-dependent gating mechanism that regulates cargo transfer through the proximal translocation channel (**Fig. 3**), providing further insight into substrate transit through the *cag* T4SS apparatus. Consistent with a proposed role as an endopilus-like structure, previous work demonstrates that CagC is a membrane-associated VirB2 ortholog that is localized to the bacterial cell surface via a process that requires *cag* T4SS apparatus biogenesis (288). Additional studies report the assembly of filamentous structures at the *H. pylori*-gastric epithelial cell interface that were proposed to constitute an extracellular *cag* T4SS-associated appendage analogous to interbacterial conjugative pili or the *A. tumefaciens* T pilus (273, 275, 288, 295, 296). However, the molecular composition of these structures has not been elucidated and multiple reports have identified several inconsistencies that argue against the presence of a canonical *cag* T4SS conjugative pilus. For example, *H. pylori* isogenic mutants deficient in the OMC component CagY, which forms the central translocation channel and is required for core complex assembly (253, 257, 258, 274), produce extracellular filaments at the same level as WT bacteria (273, 295). Thus, it is difficult to envision how a *cag* T4SS-associated pilus nucleates and assembles in the absence of (i) a translocation channel required for pilin transport to the bacterial cell surface and (ii) the OMC platform on which the pilus assembles (158). While a putative *cag* T4SS-associated pilus may nucleate in the periplasm or at the inner membrane complex via mechanisms similar to T3SS needle biogenesis (297), evidence strongly supports a model whereby prototypical conjugative pilus structures polymerize at the outer membrane (158).

Furthermore, extracellular *cag* T4SS-associated filaments have not been observed in frozen hydrated *H. pylori* or in *H. pylori*-gastric cell co-cultures imaged via cryo-electron

tomography (cryo-ET) under near native conditions (244, 250). While a conjugative pilus has not been directly observed in the current study, these results provide compelling evidence for the assignment of CagC as a VirB2 pilin structural ortholog that forms a gating mechanism within the proximal secretion channel (**Fig. 4**). Because CagC forms homomeric complexes and interacts with IMC and PRC components, I hypothesize that CagC constitutes the unresolved 'stalk' structure in the *cag* T4SS that spans the IMC and extends into the periplasm (262, 264, 265). I propose that similar to the conjugative F pilus and the *vir* T4SS T pilus (159, 161), CagC protomers assemble into a pentameric endopilus-like structure that lines the secretion channel to form the cone-shaped stalk density bridging the IMC and PRC (244, 250) (**Fig. 5**). Similar stalk or cylinder structures connecting the IMC and the OMCC are a relatively common feature among T4SSs. In the *Legionella dot/icm* T4SS, the 'stalk' engages the pore of the OMCC and is hypothesized to act as a scaffold for secretion system assembly or as a plug that prevents release of substrates in non-secreting T4SSs (174). In the R388 T4SS, the stalk is comprised of pentameric VirB6 and VirB5 and is predicted to serve as a base for pilus assembly (154). Similarly, assembled CagC pentamers may serve as a scaffold for biogenesis of heteromeric complexes that incorporate the proposed VirB5-like minor pilin CagL (296, 298, 299). In our model, DNA is shuttled through the translocation channel by 'endopilus' ratcheting. Alternatively, CagC may serve as a plug localized within the outer membrane complex that prevents release of substrate by an inactive secretion system.

Taken together, this work provides mechanistic insight into structural features that orchestrate *cag* T4SS substrate translocation. Our findings reveal that *cag* T4SS-mediated trans-kingdom DNA conjugation occurs via a multi-step mechanism and identifies a novel

gating feature whereby CagC facilitates the passage of substrates through the translocation channel. Continuing to elucidate the mechanism of DNA processing and transport through *cag* T4SS and into host cells will further our understanding of how *H. pylori* exploits *cag* T4SS activity to stimulate gastric disease.

CHAPTER 3. ARCHITECTURAL ASYMMETRY ENABLES DNA TRANSPORT THROUGH THE *CAG* TYPE IV SECRETION SYSTEM

3.1 Summary

The goal of this work was to define the role of architectural asymmetry in substrate transport through the *cag* T4SS. Asymmetry between structural subcomplexes is a common feature among bacterial nanomachines including the T4SS; however, the role of symmetry mismatch in secretion system assembly and function remains undefined. This work demonstrates that symmetry mismatch between the *cag* T4SS periplasmic ring complex (PRC) and the outer membrane core complex (OMC) facilitates DNA substrate selection and trafficking through the translocation channel. Structural analyses revealed a cluster of positively charged residues at the interface of adjacent CagX subunits that exhibit extensive π-π stacking interactions. I show that intermolecular π-π stacking enables trans-kingdom DNA conjugation without disrupting translocation of protein or peptidoglycan effector molecules into host cells. Taken together, this work suggests a model in which symmetry mismatch exposes CagX π-π interfaces within the PRC to facilitate DNA cargo passage through the *cag* T4SS translocation channel and provides insight into the biological significance of architectural asymmetry in T4SS function.

3.2 Introduction

Architectural symmetry mismatch is a remarkable evolutionary innovation exhibited by diverse bacterial nanomachines including the dynamic type IV secretion system (T4SS) superfamily (256-258). Recent advances in the structural definition of

paradigmatic T4SS machineries uncovered unexpected asymmetric features at the junction between the inner membrane complex (IMC) and outer membrane complex (OMC) or between concentric rings comprising the OMC in expanded T4SS architectures (253, 256-258). Similar symmetry mismatch has also been observed in other bacterial secretion systems including the T2SS, T3SS, and T6SS (300-303), suggesting an important, yet unresolved, functional significance. In 'minimized' conjugative T4SSs, symmetry mismatch is observed between the IMC and components of the OMC that extend down into the IMC (156, 304-306) which may provide conformational mobility that facilitates the unidirectional ratcheting of selected cargo between adjacent architectural layers. Similarly, 'expanded' effector translocator systems exhibit symmetry mismatch between subassemblies within the OMCC that may allow for conformational flexibility or gating within the secretion channel that serves to orchestrate effector selection and directional translocation (154, 262, 264).

The *cag* T4SS is an 'expanded' effector-translocator system comprised of an IMC harboring a trio of ATPases connected to the OMC via a structurally unresolved stalk-like region (244, 250). Within the OMC, the outer membrane-associated cap incorporates the structural components CagX, CagY, CagT, CagM, and Cag3, and includes a periplasmic ring complex (PRC) comprised of portions of CagX and CagY (244, 262, 264, 265, 274). Analysis of the *cag* T4SS cryo-EM structure revealed striking symmetry mismatch whereby several CagX subunits in the PRC (14-fold symmetry) do not display any obvious connection to corresponding subunits within the OMC (17-fold symmetry) (264). Asymmetry between OMC subassemblies is observed in other 'expanded' systems including the *Legionella pneumophila dot/icm* T4SS (255) and conjugative F-plasmid

machinery (268, 269), however, despite being a common feature among macromolecular complexes, the biological significance of symmetry mismatch between T4SS subassemblies remains unknown.

This chapter aimed to further investigate the mechanism underlying substrate trafficking through the *cag* T4SS and to investigate the functional significance of architectural asymmetry within the *cag* T4SS apparatus. I demonstrate that CagX is a DNA-binding T4SS apparatus component that preferentially binds dsDNA via π-stacked residues within unoccupied CagX regions in the PRC, suggesting that architectural symmetry mismatch in the *cag* T4SS facilitates substrate selection. Disruption of CagX π-π interactions resulted in a significant reduction of DNA binding interactions *in vitro* and diminished DNA substrate translocation *in vivo* without perturbing non-DNA cargo transport into gastric epithelial cells. These data demonstrate secondary DNA substrate selection within the apparatus core complex. Collectively, these studies provide insight into the mechanisms by which the *cag* T4SS orchestrates effector molecule delivery and highlights novel architectural innovations that enable substrate specificity and regulate cargo transport into host cells.

3.3 Materials and Methods

Bacterial strains and culture conditions

Helicobacter pylori 26695 and isogenic derivatives were grown on trypticase soy agar plates supplemented with 5% sheep blood (BD) at 37˚C with 5% CO_2. Overnight cultures of *H. pylori* were grown in Brucella broth supplemented with 5% fetal bovine serum (FBS)

at 37°C in 5% CO_2. *E. coli* strain DH5a (New England Biolabs) was used for plasmid propagation and was grown on lysogeny broth (LB) agar or liquid media supplemented with the appropriate antibiotics for plasmid maintenance. Bacterial two-hybrid (BACTH) strain BTH101 (a gift from Scot Ouellette, University of Nebraska Medical Center) was maintained on LB, and BACTH screening was performed using LB plates supplemented with 0.5 mM isopropyl b-D-1-thiogalactopyranoside (IPTG), 100 mg/mL ampicillin, 50 mg/mL kanamycin, and 20 mg/mL 5-bromo-4-chloro-3-indoyl-b-D-galactopyranoside (X-Gal)(271). BACTH-dependent growth selection was performed in M63 minimal media [$(NH_4)_2SO_4$ (2 g/L), KH_2PO_4 (13.6 g/L), Thiamine B1 (1 mg/L), 1 mM $MgSO_4$, $FeSO_4.7H_2O$ (5 mg/L), pH 7] supplemented with 0.4% maltose, 0.5 mM IPTG, 50 µg/mL ampicillin, and 15 µg/mL kanamycin. All BACTH experiments were executed at 30°C under constant aeration.

H. pylori mutagenesis

Isogenic mutants (251, 272-274) were complemented in *cis* at the *ureA* locus using plasmids derived from pAD1 (275) that were designed to express either the native gene or a protein variant as previously described (64, 275). Point mutants were generated using Q5 Site Directed Mutagenesis (New England Biolabs). Plasmid sequences were confirmed by PCR and whole plasmid sequencing, and constructs were used to transform isogenic mutant strains. Colonies resistant to kanamycin (12.5 µg/ml) or chloramphenicol (10 mg/mL) were selected and complementation at the *ureA* locus was confirmed by PCR and by anti-CagX immunoblotting, when appropriate.

Human cell culture

HEK293-hTLR9 (Invivogen hkb-htlr9), the corresponding parental HEK293 null1 (InvivoGen hkb-null1) cell lines were grown in DMEM media supplemented with 10% heat-inactivated FBS and 1X GlutaMAX (Life Technologies) in the presence of 5% CO_2 at 37°C. AGS human gastric epithelial cells (ATCC CRL-1739) were grown in RPMI media supplemented with 10% FBS, 2 mM L-glutamine, and 10 mM HEPES in the presence of 5% CO_2 at 37˚C.

TLR9 stimulation assay

TLR9 activation assays were carried out as previous described (276, 277). Briefly, HEK293 cells expressing human TLR9 (HEK293-hTLR9) or the parental HEK293 null1 cells were co-cultured with WT *H. pylori* 26695 or *cag* mutant strains at a MOI of 100. Supernatants were collected at 24 hours post-infection, and TLR9 activation was quantified by measuring secreted embryonic alkaline phosphatase (SEAP) in cell culture supernatant by QuantiBlue™ reagent (Invivogen) using a microplate reader (BioTek Synergy HI) to record the absorbance at 650 nm. TLR9 activation was normalized to levels of SEAP produced by the corresponding infected null1 cells and data were expressed as a percent of the normalized fold change over WT-challenged cells.

Recombinant protein purification

To generate maltose binding protein (MBP)-tagged recombinant fusion proteins, genes of interest were amplified from *H. pylori* 26695 and cloned into pMAL-c6T (New England Biolabs) for protein expression and purification via affinity chromatography as previously described(307). Briefly, MBP fusions were expressed in Lemo21 (DE3) *E. coli* by induction with 0.5 mM IPTG for 3 hours at 37˚C. Bacteria were collected via centrifugation

and lysed by sonication in lysis buffer (20 mM Tris pH 7.5, 500 mM NaCl, 1 mM EDTA, 5 mM BME, 1X protease inhibitor cocktail) followed by centrifugation at 16000 x g for 20 min at 4˚C. Soluble MBP fusions were passed through a column containing equilibrated amylose resin 3 times at a low flow rate. Columns were washed to remove unbound protein fractions using one column volume of lysis buffer (20 mM Tris, pH 7.5, 500 mM NaCl, 1 mM EDTA, 5 mM BME), one column volume of high salt lysis buffer (20 mM Tris, pH 7.5, 1 M NaCl, 1 mM EDTA, 5 mM BME), followed by a final wash with one column volume lysis buffer. MBP fusions were eluted in 10 mM maltose, 20 mM Tris, pH 7.5, 100 mM NaCl, 1 mM EDTA, 5 mM BME. Protein size and purity was determined by SDS-PAGE analysis and Coomassie staining, and protein concentration was determined by Bradford assay (Bio-Rad). Purified MBP fusions were stored in elution buffer at -80°C.

DNA binding electromobility shift assays (EMSA)

5′ Cy5-labeled chemically synthesized oligonucleotides containing the *tfs3 oriT*-like recognition sites (308) were incubated with varying concentrations of MBP fusion proteins ranging from 0 µM to 8 µM. DNA binding reactions were assembled in 15 µl total volume containing 0.01 mM DNA substrate in reaction buffer (50 mM Tris-HCl pH 8, 50 mM KCl, 0.5 mM $MgCl_2$, 1 mM DTT, 0.1 µg/µl BSA) for 30 min at 4°C. Free DNA and nucleoprotein complexes were resolved using native PAGE (10%) run in 1X TBE at 70V for 2-3h. For competition assays, unlabeled competitor oligonucleotides (0.01 mM to 0.4 mM) were added to the reaction mixture containing Cy5-labeled oligonucleotides (0.01mM) and MBP-CagX (0.75mM). After incubation on ice for 30 min, reaction mixtures were resolved using native PAGE (10%) run in 1X TBE at 70V for 2-3h at 4˚C. Quantification of free DNA was performed using ImageLab software (Bio-Rad).

Protein structure modeling

Structural analyses of the *H. pylori cag* T4SS PRC were conducted on PDB 6X6J. The electrostatic surfaces of CagX, TraK, and VirB9 were generated as coulombic potential maps in ChimeraX.

Immunoblotting

To detect bacterial protein expression, normalized samples were lysed in reducing 2X SDS buffer (Bio-Rad), resolved by SDS-PAGE (10%), transferred to a nitrocellulose membrane, and immunoblotted using rabbit polyclonal antisera raised against recombinant CagX (antisera a gift from Dr. Tim Cover) as previously described (275). To confirm equal sample loading, membranes were immunoblotted using *H. pylori* outer membrane protein monoclonal antibody (α-OMP, Santa Cruz) or anti-*H. pylori* polyclonal antisera raised against *H. pylori* whole cell lysate (a gift from Dr. Tim Cover). Secondary detection by horse radish peroxidase-conjugated anti-rabbit IgG or anti-mouse IgG was achieved using SuperSignal West Pico chemiluminescent substrate (Thermo).

IL-8 quantitation

Secretion of IL-8 by AGS cells co-cultured with *H. pylori* or corresponding isogenic derivatives was determined using the human CXCL8 ELISA (R&D Systems). Briefly, AGS cells were challenged with *H. pylori* or the indicated isogenic mutant strain for 4.5 h at an MOI of 100. Supernatants were collected and analyzed as previously described (64, 275). Biological replicate experiments were performed a minimum of three times.

3.4 Results

Periplasmic *cag* T4SS assemblies bind transfer DNA

To further investigate the mechanism of substrate transport through the *cag* T4SS, I first sought to better define the mechanism of CagC gating in the translocation channel. One possibility is that a putative CagC 'endopilus' corresponding to the *cag* T4SS stalk-like region visualized *in situ* (**Chapter 2**, (244, 250)) may dynamically assemble and retract to load effector DNA into the secretion channel via a piston-like mechanism. To test the hypothesis that CagC interacts with DNA, I purified recombinant CagC and performed electromobility shift assays (EMSA) to analyze DNA binding capacities. In agreement with the studies described in Chapter 2 suggesting that dsDNA is delivered to host cells in a *cag* T4SS-dependent manner, CagC weakly bound 58-mer effector dsDNA sequences (38), but not corresponding 58-mer single-stranded DNA (ssDNA) targets (**Fig 6a**). Based on the observation that CagC interacts with both CagX and fragments of CagY that comprise the PRC (**Fig 4h, Fig. 6b**), I questioned whether other *cag* T4SS complexes bind DNA to facilitate cargo transfer. Compared to maltose-binding protein (MBP) or recombinant MBP-CagT controls, MBP-CagX robustly bound target *H. pylori* DNA sequences (38) in a concentration-dependent manner (**Fig. 6d,e**). In contrast to CagX, TraK and VirB9 homologs harbored by conjugative DNA transfer systems did not bind DNA (**Fig. 7a,b**), suggesting that DNA binding within the PRC is a unique structural innovation in the *cag*

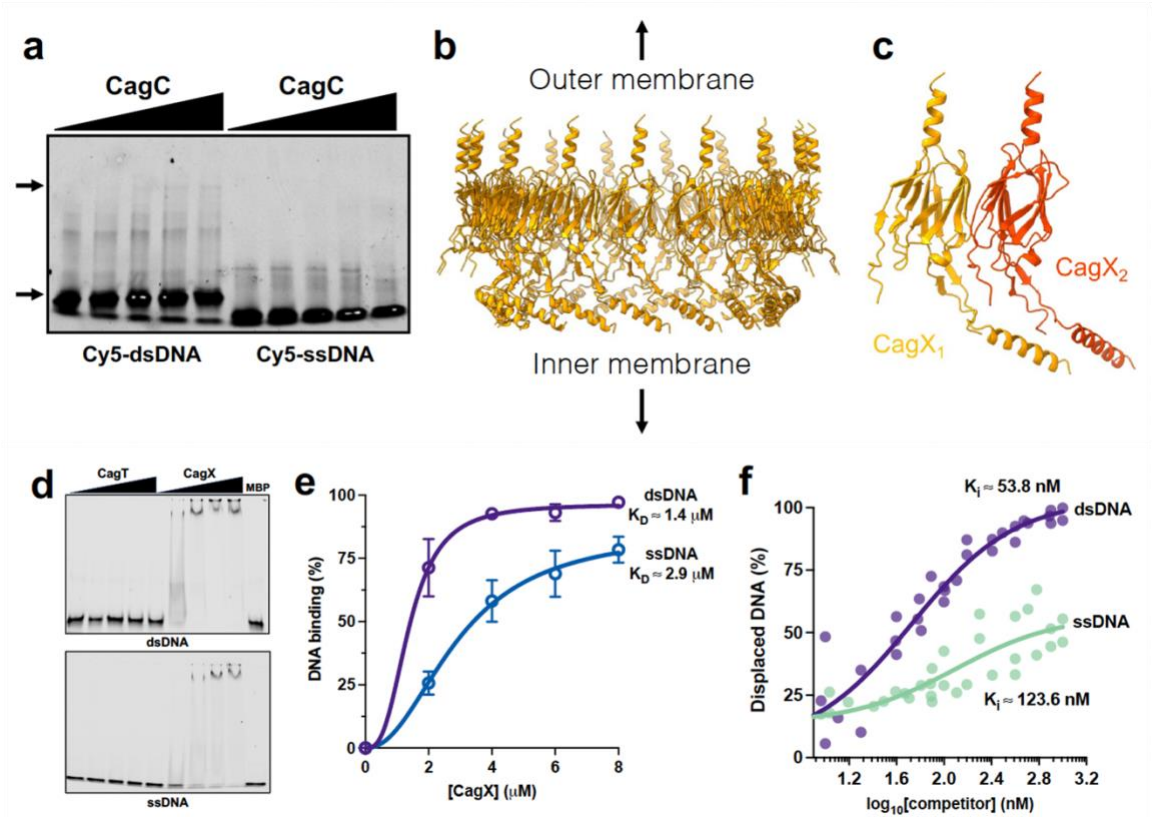

Figure 6. CagX binds dsDNA in the PRC.

(**a**) EMSA analysis of recombinant MBP-CagC binding to 58-mer dsDNA (left) or ssDNA. Arrows indicate recombinant CagC-DNA complexes. (**b**) The position of CagX subunits within the PRC (PDB: 6X6J). (**c**) Adjacent CagX subunits ($CagX_{NT}$, residues 41-310) within the PRC. (**d**) EMSA analysis of recombinant CagX binding to dsDNA (upper panel) and ssDNA (lower panel). MBP-CagX, but not MBP or MBP-CagT, binds target dsDNA and ssDNA. (**e**) Quantitation of MBP-CagX binding to the indicated DNA substrate. Values indicate the experimentally approximated K_D for CagX-DNA binding to 58-mer targets. (**f**) Competition assays demonstrating preferential CagX binding to dsDNA. Data points represent the % displaced DNA obtained in independent biological replicate experiments. Solid line represents the nonlinear regression for estimating the inhibitory constants (K_i) for dsDNA and ssDNA. In **e**, dissociation constants (K_D) were estimated using the specific binding with hill slope nonlinear regression model in and in **f**, inhibitory constants (K_i) were estimated using the one site fit K_i model (heterologous ligand binding affinities using experimentally derived K_D in **e**) in GraphPad Prism 9.5.

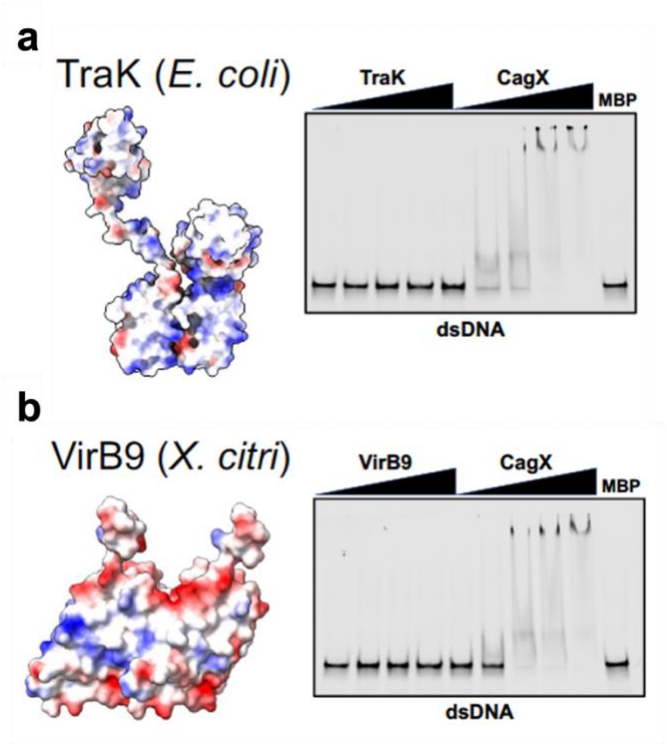

Figure 7. CagX-DNA interaction is unique to the *cag* T4SS.

(**a**) 58-mer dsDNA EMSA analysis of TraK derived from F plasmid R1 (structural model depicts TraK from F plasmid pED208, PDB: 7SPB) and (**b**) VirB9 derived from *A. tumefaciens vir* T4SS (structural model depicts VirB9 from *X. citri vir* T4SS, PDB: 6GYB). MBP-CagX, but not MBP-TraK or MBP-VirB9, binds dsDNA.

T4SS apparatus. Consistent with the observation that CagC displays preferential interactions with dsDNA substrates (**Fig. 6a**), recombinant CagX exhibited a stronger binding affinity for dsDNA compared to corresponding ssDNA sequences (**Fig. 6e**). In support of this observation, DNA competition assays demonstrated CagX preferential binding to dsDNA (**Fig. 6f**). Together, these studies identify CagX as a novel DNA-binding *cag* T4SS apparatus component.

Symmetry mismatch enables DNA loading into the secretion channel.

To better understand the molecular determinants of DNA binding within the PRC, I employed a structure-guided analysis to identify potential DNA binding sites in CagX-CagY assemblies. Close inspection of the CagX structure within the PRC (257, 258) (**Fig. 6c**) revealed a basic patch formed at the interface of adjacent CagX subunits within the PRC (**Fig. 8a**) that represented a potential DNA binding region. I hypothesized that overlapping π orbitals among a series of arginine (Arg 36, 38, 278, and 295) and tryptophan (Trp 58) residues within the basic patch would produce electrostatic interfaces that enable DNA interactions within the secretion channel (**Fig. 8b**). Notably, the diameter and geometry of this proposed DNA binding site supports this assignment (**Fig. 8c**). To test the hypothesis that electrostatic CagX interfaces facilitate DNA loading into the secretion channel, I generated CagX variants by introducing point mutations in adjacent arginine residues within the basic patch region to disrupt intermolecular π-stacking interactions (**Fig. 8b**). In comparison to $CagX^{WT}$, CagX variants in which Arg 36 and Arg 38 were replaced with alanine ($CagX^{R36A,R38A}$) exhibited similar dsDNA binding affinities *in vitro*, whereas CagX variants harboring the Arg 278 single point mutation ($CagX^{R278A}$) exhibited significantly decreased DNA binding capacities that were further reduced when

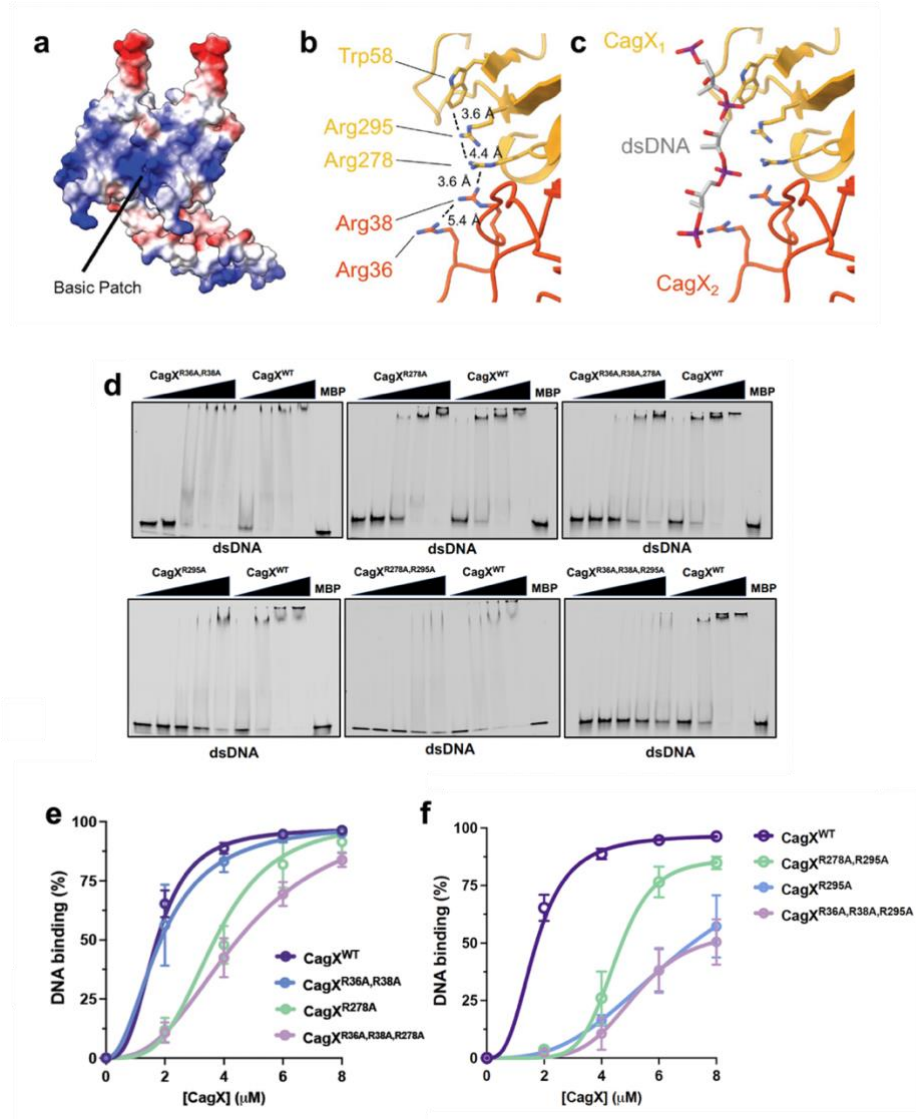

Figure 8. π-π interactions within asymmetric PRC subunits mediate CagX-DNA binding

(**a**) Electrostatic surface potential of two adjacent CagX residues demonstrating the formation of a basic patch at the subunit junction. Regions of acidic and basic potential are shown in red and blue, respectively. (**b**) Overlapping p- p orbital interactions among CagX arginine and tryptophan residues in adjacent subunits. (**c**) Proposed model of dsDNA binding to CagX subassemblies. (**d**) EMSA analyses of recombinant CagX variant binding to dsDNA. (**e**) Quantitation of CagX variant binding to dsDNA. Values indicate the experimentally approximated K_D for $CagX^{WT}$ (1.7 μM), $Cag^{XR36A,R38A}$ (1.9 μM), $CagX^{R278A}$ (3.7 μM), and $CagX^{R36A,R38A,R278A}$ (4.6) μM) binding to 58-mer targets. (**f**) Quantitation of MBP-CagX binding to 58-mer DNA substrates. The experimentally approximated K_D for CagX-DNA binding was determined for $CagX^{WT}$ (1.7 μM), $Cag^{XR278A,R295A}$ (4.5 μM), $CagX^{R95A}$ (6.9 μM), and $CagX^{R36A,R38A,R295A}$ (5.2 μM). In **e** and **f** dissociation constants (K_D) were estimated using the specific binding with hill slope nonlinear regression model in GraphPad Prism 9.

consecutive Arg residues were disrupted in tandem ($CagX^{R36A,R38A,R278A}$) (**Fig. 8d,e**). In agreement with my hypothesis, *H. pylori cagX* mutants expressing $CagX^{R36A,R38A}$, $CagX^{R278A}$, or $CagX^{R36A,R38A,R278A}$ translocated significantly less DNA into host cells, resulting in markedly reduced levels of TLR9 activation compared to WT and *cagX* mutants genetically complemented by $CagX^{WT}$ (**Fig. 9b**). Analysis of additional *cag* T4SS-dependent phenotypes produced by the $CagX^{R36A,R38A}$ variant strain revealed that despite the observed DNA translocation defects, levels of IL-8 secretion were comparable to the parental WT and $CagX^{WT}$ complemented strains (**Fig. 9c**), suggesting that DNA delivery is mechanistically uncoupled from the translocation of other *cag* T4SS substrates. Disrupting Arg 295 via single point mutation or in tandem with Arg 36, Arg 38, or Arg 278 resulted in significantly decreased *in vitro* DNA binding (**Fig. 8d,f**) and destabilized CagX *in vivo* (**Fig. 9a**), leading to significantly reduced *cag* T4SS activity (**Fig. 8d,e).** I thus speculate that Arg 295 constitutes a critical CagY-binding interface in the PRC. Collectively, these studies reveal that CagX subunits assemble to form a DNA-binding pocket within the *cag* T4SS PRC that contributes to trans-kingdom conjugation.

Analysis of *cag* T4SS architecture resolved by single particle reconstruction indicates that portions of CagX and CagY assemble into structures localized in both the OMC and the PRC (**Fig. 10a-c**), producing a striking symmetry mismatch (257, 258). In this study, I identified a putative DNA binding site on CagX that is centered around resides Arg 36, 38, 278, 295 and Trp 58. However, the high-resolution cryo-EM structure indicates that this site is occupied by the periplasmic portion of CagY and thus, DNA binding may be excluded in the assembled PRC. However, the cryo-EM structure was determined through the imposition of symmetry and argue that this map omits asymmetric features. To

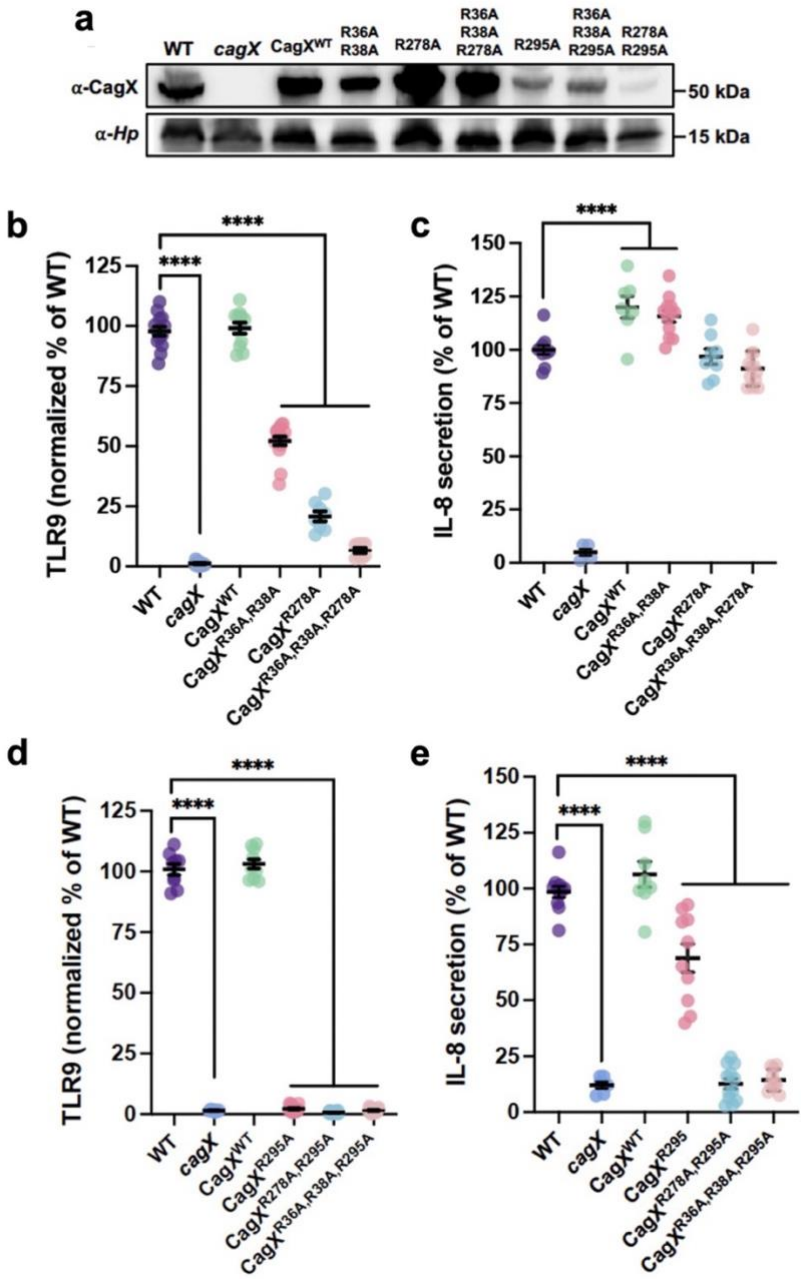

Figure 9. Symmetry mismatch coordinates DNA substrate selection for trans-kingdom conjugation.

a) Levels of CagX produced by *H. pylori cagX* harboring the indicated CagX variant. (**b,d**) TLR9 stimulation and (**c,e**) IL-8 secretion induced by *H. pylori* strains. In **b-e**, significance was determined by one-way ANOVA with Dunnett's post-hoc correction for multiple comparisons to experimental controls; ****, $p<0.0001$

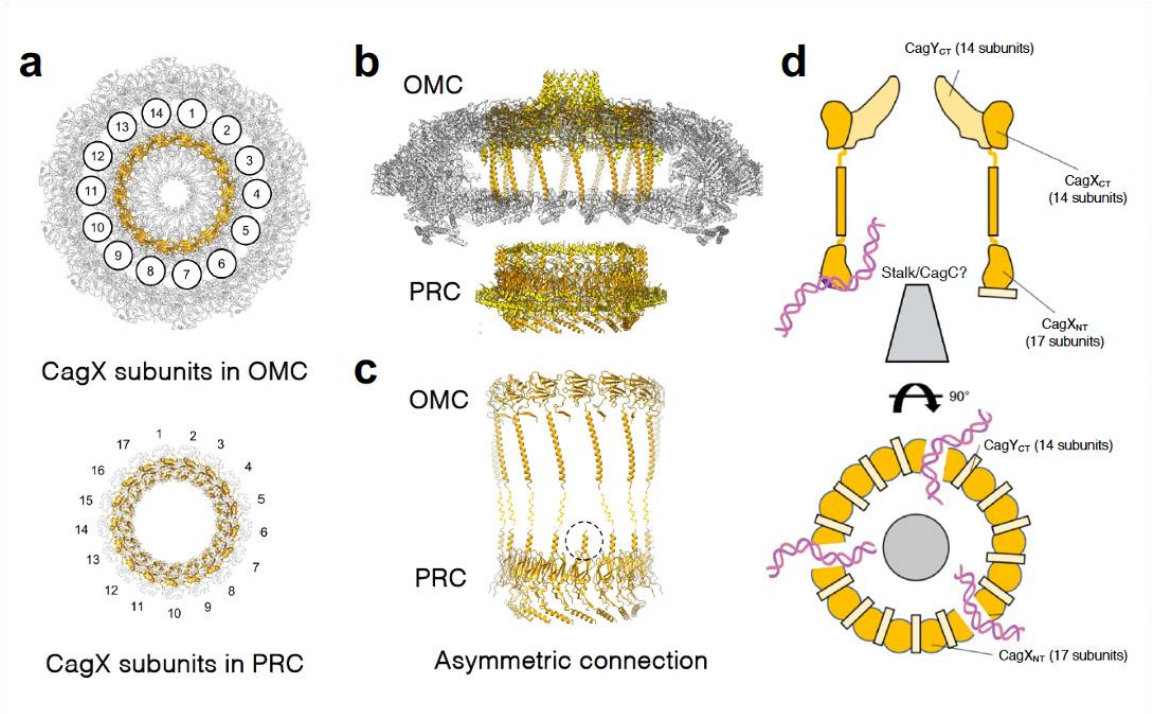

Figure 10. Proposed model of DNA binding within the PRC.

(**a**) Single particle cryo-EM model of CagY (yellow) and CagX (orange) within the *cag* T4SS OMC and the PRC. (**b**) Symmetry mismatch between subunits in the OMC (upper panel, 14-fold symmetry) and the PRC (lower panel, 17-fold symmetry). (**c**) Illustration of one asymmetric connection between the OMC and PRC exposing CagX helices extending into the PRC lumen (dashed circle). Three copies of CagX in the PRC do not show any obvious connection to the OMC. (**d**) Proposed model of dsDNA loading into the *cag* T4SS core complex via p-p stacking within asymmetric PRC connections. Illustration depicts the architectural juxtaposition of CagX N-terminal domains ($CagX_{NT}$), CagX C-terminal domains ($CagX_{CT}$), CagY N-terminal domains ($CagY_{NT}$), and CagY C-terminal domains ($CagY_{CT}$) in the PRC with dsDNA loading into the secretion channel via interaction with asymmetric CagX subunits.

that end, a low-resolution reconstruction of the PRC detected 14 copies of CagX in the OMC and 17 copies of CagX in the PRC (**Fig. 10a**) with 14 tube-like helical densities bridging the OMC (CagX C-terminus) and the PRC (CagX N-terminus) (258) (**Fig. 10b,c**). This observation suggests that the additional copies of CagX in the PRC (**Fig. 10a**) provide DNA interaction sites to facilitate nucleic acid trafficking through the core complex (**Fig. 10d**). I propose a model whereby asymmetric regions of the PRC assemble to form a DNA binding site for substrate loading into the *cag* T4SS secretion channel (**Fig. 10d**).

3.5 Discussion

Trans-kingdom DNA conjugation by the *cag* T4SS profoundly impacts host cell biology and is an important mechanism of *H. pylori* pathogenesis. *H. pylori* trans-kingdom conjugation elicits both anti- and pro-inflammatory responses via TLR9 stimulation (36, 39, 40) and activation of additional nucleic acid reconnaissance systems (38). Consequently, *H. pylori* evolved mechanisms to counterbalance nucleic acid surveillance and suppress STING signaling within the gastric mucosa (40, 243). Thus, *cag* T4SS-dependent DNA translocation is a significant driver of *H. pylori* pathogenesis as it provides additional mechanisms by which *H. pylori* manipulates the immune response to enable persistent gastric niche colonization (36, 40).

Despite its importance to *H. pylori* pathogenesis, the mechanism of *cag* T4SS dependent DNA translocation is incompletely defined. DNA processing for *H. pylori* trans-kingdom conjugation occurs via incompletely defined mechanisms that require evolutionarily divergent VirD2-like relaxases and the DNA translocase/partitioning protein

FtsK that couples nucleic acid cargo to the translocation channel (38). In Chapter 2, I provide multiple lines of evidence demonstrating that *cag* T4SS-dependent DNA translocation is susceptible to exogenous DNA-targeting nucleases and anti-DNA monoclonal antibodies, suggesting a multi-step translocation mechanism that does not appear to require a conjugative pilus conduit. In this study, I reveal a secondary substrate selection site in the periplasmic ring complex facilitated by CagX π-π interactions (**Fig. 8, 9**). As previously suggested for the *A. tumefaciens vir* T4SS (309), the VirB9 ortholog CagX exhibits N-terminal domains that contribute to protein stability and establish inter-subunit contacts required for DNA substrate selection and trafficking through the translocation channel. In *A. tumefaciens,* mutations in the N-terminal domain of VirB9 selectively block translocation of either the IncQ plasmid or T-DNA substrates, suggesting that the N-terminal domain acts as a checkpoint that regulates passage of substrates through the periplasm and to the outer membrane. Similarly, I uncovered multiple 'uncoupling mutations' in CagX that specifically disrupt DNA translocation without perturbing delivery of other substrates into host cells. The capacity of CagX to recognize DNA and mediate selective substrate transfer through the *cag* T4SS OMC provides intriguing evidence that in addition to prototypical coupling proteins, periplasmic architectural features recruit effector cargo to the secretion channel. I speculate that CagX architectural asymmetry thus represents a secondary substrate selection checkpoint that regulates cargo transit through the periplasm and across the distal OMC translocation channel (**Fig. 10**). I hypothesize that symmetry mismatch in other T4SS machineries, and potentially in other diverse secretion system architectures, similarly orchestrates substrate selection to ensure unidirectional translocation into target cells. These observations expand the growing collection of

identified mutants that exhibit selective substrate translocation through the *cag* T4SS (24, 38) and in other evolutionarily divergent nanomachines in which scaffold biogenesis is uncoupled from trans-kingdom DNA conjugation (278, 309, 310).

Taken together, this work provides mechanistic insight into DNA trafficking through the translocation channel. π-π interactions in asymmetric CagX subunits in the PRC facilitate substrate selection within the periplasm, suggesting a novel two-step substrate translocation mechanism. This study identifies the functional significance of architectural symmetry mismatch and deepens our understanding of how the *cag* T4SS orchestrates the release of diverse molecular cargo across the bacterial outer membrane.

CHAPTER 4. ROLE OF UNIQUE *CAG* T4SS COMPONENTS IN EFFECTOR TRAFFICKING

4.1 Summary

The goal of this work was to analyze the contribution of species-specific *cag* T4SS components in effector molecule trafficking through the secretion channel. This work demonstrates that unique *cag* components CagZ (a regulator of ATPase activity and a CagA translocation factor) and CagH (a membrane-associated component that interacts with CagL and CagI) govern CagA and dsDNA transport through the translocation channel. Additionally, disruption of the CagH C-terminal hexapeptide motif results in the mislocalization of both dsDNA and CagA to the bacterial cell surface in the absence of host cell contact, suggesting that similar to CagC, CagH regulates substrate passage through the translocation channel. Furthermore, this work demonstrates that APEX2-proximity labeling can be used to identify direct and transient protein-protein interactions within the *cag* T4SS *in situ. H. pylori* expressing APEX2-CagH labeled a unique set of proximity partners compared to strains expressing either APEX2 or APEX2-CagF, a chaperone-like protein required for delivery of the oncoprotein CagA. This proteomic mapping approach can be used to investigate structural changes in *cag* T4SS that occur to trigger effector cargo translocation when *H. pylori* is in contact with gastric epithelial cells.

4.2 Introduction

Oncogenic effector molecules are trafficked into gastric epithelial cells via *cag* T4SS-dependent mechanisms. Currently, how disparate substrates are moved through the *cag* T4SS secretion channel is unknown. While 12 of the 27 *cag* T4SS components exhibit

homology to prototypical T4SS subunits, the presence of multiple unique apparatus constituents suggests distinct roles *cag* T4SS activity (1, 260). This chapter aims to investigate the function of unique *cag* T4SS components to better understand mechanisms underscoring nanomachine architecture, apparatus assembly, and effector translocation.

Using methods developed to study the *A. tumefaciens vir* T4SS (278), I identified a *cag* T4SS gating mechanism in which CagC controls effector passage to the bacterial cell surface through formation of an 'endopilus' gate that may dynamically assemble and disassemble to plug the secretion channel pore (**Chapter 2**). In this study, I demonstrate that inactivation of *cagH* and *cagZ* results in the mislocalization of CagA to the bacterial cell surface. In addition, I show that deletion of the CagH C-terminal domain results in the uncontrolled release of effector molecules across the bacterial outer membrane, indicating that specific CagH domains are involved in *cag* T4SS translocation channel gating.

To further investigate *cag* T4SS architecture, I adapted APEX2-proximity labeling approaches in *H. pylori*. APEX2 is an ascorbate peroxidase that catalyzes the production of biotin phenoxyl radicals in the presence of hydrogen peroxide (H_2O_2). These biotin-phenoxyl radicals travel short distances (~20 nm) and covalently interact with side chains of nearby proteins, allowing for the detection of both direct and transient protein-protein interactions (311-317). Biotinylated proteins are isolated with streptavidin beads and analyzed by mass spectrometry. In this study, APEX2-dependent proximity labeling was performed both in *H. pylori* cultured in the presence (active secretion) or absence (inactive secretion) of gastric epithelial cell contact to identify changes in secretion system architecture that trigger effector translocation. This proteomic mapping approach can be used to identify novel protein-protein interactions within the assembled machinery and will

provide insight into structural or architectural oscillations that orchestrate effector selection and trigger cargo translocation into the host cell.

4.3 Material and Methods

Bacterial strains and culture conditions

Helicobacter pylori 26695 and isogenic mutant strains were grown on trypticase soy agar plates supplemented with 5% sheep blood (BD) at 37˚C with 5% CO_2. Overnight cultures of *H. pylori* were grown in Brucella broth supplemented with 5% fetal bovine serum (FBS) at 37°C with 5% CO_2. *E. coli* strain DH5α (New England Biolabs) was used for plasmid propagation and was grown on lysogeny broth (LB) agar plates or in LB liquid media supplemented with appropriate antibiotics for plasmid maintenance.

H. pylori mutagenesis

Mutagenesis of *cag* genes was performed as previously described (272, 277). Briefly, *H. pylori* was transformed with a suicide plasmid in which the target gene was replaced by a kanamycin or a chloramphenicol resistance cassette and homologous flanking DNA sequences 500 base pairs (bp) up- and downstream of the target locus. Colonies resistant to kanamycin (12.5 μg/ml) or chloramphenicol (10μg/mL) were selected and correct insertion of the resistance cassette was confirmed by PCR. Complementation of *cag* mutant strains was done using a plasmid derived from pAD1 (272), which allows genes of interest to be inserted into the UreA locus. pAD1 plasmids expressing APEX2 constructs were generated using restriction digest. Plasmid sequences were confirmed by PCR and sanger sequencing, and constructs were used to transform isogeneic *cag* mutant strains. Colonies

resistant to chloramphenicol were selected and correct insertion of complementation constructs into the *ureA* locus was confirmed by PCR amplification.

Colony Immunoblot

Colony immunoblotting was performed as described in Chapter 2 Methods.

Proximity labeling

APEX2 proximity labeling in *H. pylori* was adapted from the proximity labeling experiments performed in *E. coli* (316). *H. pylori* strains expressing APEX2 constructs were grown overnight in 25 mL brucella broth supplemented with 6% FBS and 0.5 mM biotin phenol or with no biotin phenol as a negative control. After overnight growth and normalization to $OD_{600,}$ cultures were pelleted by centrifugation and washed once in brucella broth to remove any remaining biotin phenol. The labeling reaction was catalyzed by resuspending cell pellets in 1mM H_2O_2 for one minute. The reaction was stopped by addition of TSEN quench buffer (20mM Tris-HCl, 30% sucrose, 1mM EDTA, 100mM NaCl; supplemented with 10mM sodium azide, 10mM sodium ascorbate, 10µg/mL egg white lysozyme) and cells were pelleted by centrifugation. The cells were lysed in RIPA buffer (10 mM Tris, 100 mM NaCl, 1% NP-40, 0.25% deoxycholic acid, pH 7.2) supplemented with 2X cOmplete™ inhibitor (63) and lysed overnight at 4˚C. Cell lysates were collected by centrifugation.

APEX2 proximity labeling in *H. pylori* co-cultured with gastric epithelial cells was adapted from chlamydia inclusion membrane proximity labeling experiments (318). AGS cells were seeded into 6 well plates, one plate for each condition/control sample. AGS cells

were infected with *H. pylori* strains expressing APEX constructs either grown with 0.5 mM biotin-phenol or with no biotin phenol as negative control. Infections were performed at an MOI of 100 and 0.5mM biotin-phenol was added to uninfected control wells. After 4 hours, the media was removed, and the labeling reaction was catalyzed by addition of 1 mM H_2O_2 in PBS to each well for one minute at room temperature with gentle rocking. The reaction was quenched by three washes in quench/wash solution (10mM sodium ascorbate, 10mM sodium azide, 5 mM Trolox in PBS). After the final wash, cells were resuspended in RIPA buffer (10 mM Tris, 100 mM NaCl, 1% NP-40, 0.25% deoxycholic acid, pH 7.2) supplemented with 2X cOmplete™ inhibitor (63) and lysed overnight.

Affinity Purification

Biotinylated proteins were affinity purified using streptavidin as previously described (319). For each sample, 25 μL Pierce High Sensitivity Streptavidin beads (Thermo) were washed twice in RIPA buffer. Cell lysates were added to washed beads and incubated for at least 1 hour at room temperature with rotation. After incubation, the beads were washed twice in RIPA buffer and once in 1 M KCl for two minutes at room temperature followed by two quick washes (~10 seconds), once in 0.1 M Na_2CO_3 and once 2 M Urea in 10 mM Tris-HCl, pH 8.0. The beads were then again washed twice in RIPA buffer and after the final wash were transferred to a fresh tube. Streptavidin affinity purified proteins were eluted by boiling at 95˚C in 30 μL 3x protein loading buffer (0.33M Tris (pH 6.8), 34% glycerol, 10% SDS, 0.09% DTT, 0.12% bromophenol blue). Samples were resolved by SDS page in a 4-20% gradient gel (BioRad). Samples were transferred to nitrocellulose membrane and blocked in 3% BSA TBST (0.05% tris buffered saline, 0.1% Tween 20).

Affinity purified biotinylated proteins were detected using Pierce™ High Sensitivity Streptavidin-HRP (Thermo Scientific).

4.4 Results

To investigate the role of Cag components in substrate trafficking through the translocation channel, I screened a panel of *cag* isogenic mutant strains for gating defects similar to defects identified in *H. pylori cagC* **(Chapter 2).** Using colony immunoblotting, I demonstrate that compared to WT and other *cag* isogenic mutant strains, inactivation of *cagZ* results in mislocalization of the protein effector CagA to the bacterial cell surface while inactivation of *cagH* results in the uncontrolled release of both CagA and dsDNA across the outer membrane in the absence of host cell contact (**Fig. 11a**). Previous studies identified CagZ as a regulatory protein that directly interacts with the ATPase Cagβ and is essential for CagA translocation into host cells (320). Binding of CagZ stabilizes Cagβ and traps it in its monomeric state, suppressing endogenous ATPase activity until CagA is tethered to the cytoplasmic complex for translocation (320, 321). Thus, loss of CagZ-dependent Cagβ stabilization and ATPase activity could explain the uncontrolled CagA release observed in the absence of host cell contact.

Additionally, I demonstrate that CagH is involved in substrate gating within the *cag* T4SS. Compared to WT and genetically complemented (CagH-HA) strains, inactivation of *cagH* results in mislocalization of both the CagA and nucleic acid effectors to the bacterial cell surface (**Fig. 11a,c**). CagH is an essential *cag* T4SS component that regulates the formation of pilus-like structures produced by *H. pylori* in contact with gastric cells (272).

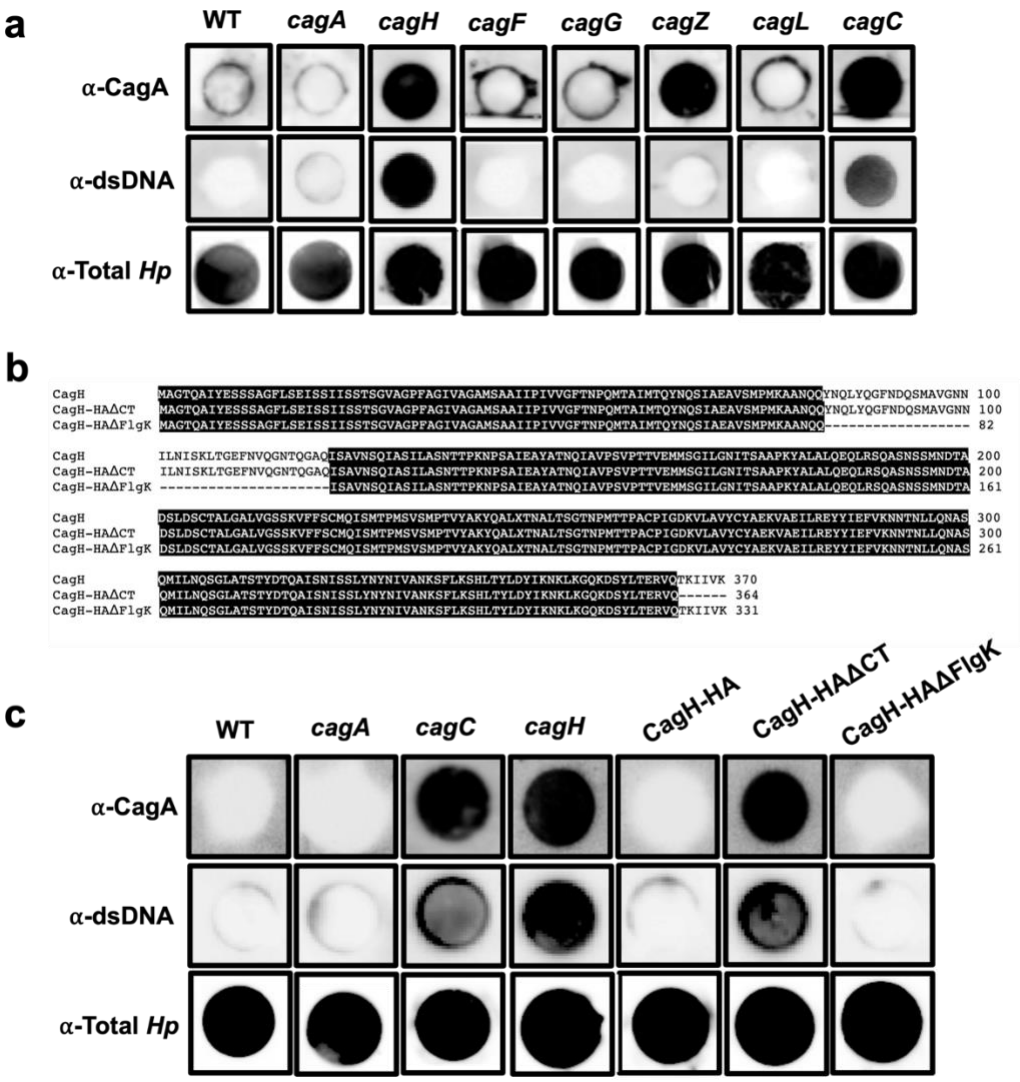

Figure 11. Unique *cag* components contribute to *cag* T4SS gating mechanisms.

(**a**) Non-denaturing colony blots probing for surface-exposed CagA, dsDNA, and total *H. pylori* in *cag* isogenic mutants. (**b**) Protein sequence alignment of CagH, CagH-HAΔCT, and CagH-HAΔFlgK. Sequence alignment was generated using Clustal Omega. (**c**) Colony blots demonstrating dsDNA and CagA cell surface localization in *H. pylori cagH* variants.

As CagC is a pilin subunit ortholog that I predict forms an 'endopilus' in the periplasm of the *cag* T4SS **(Chapter 2),** CagH may be involved in regulating 'endopilus' formation. Inactivation of *cagH* may result in improper endopilus assembly and disrupt gating mechanisms that control of effector transport through the translocation channel.

To further investigate the role of CagH in secretion channel gating, I used colony immunoblotting techniques to test for gating defects in CagH mutant strains expressing forms of CagH where various functional domains were deleted. Strains expressing a CagH variant lacking the C-terminal hexapeptide motif (275) (CagH-HAΔCT) (**Fig. 11b**), exhibit unrestricted release of both protein and DNA effectors to the bacterial cell surface **(Fig. 11c).** In the absence of the C-terminal motif, membrane-associated CagH is mislocalized to the cytoplasm, rendering the *cag* T4SS non-functional (272). This suggests that CagH membrane localization is essential to proper gating within the *cag* T4SS. I also tested a strain expressing a CagH variant in which a flagellar hook-associated protein K (FlgK) domain (39 amino acids) was deleted (CagH-HAΔFlgK) (**Fig. 11b**). Interestingly, while CagH-HAΔFlgK also produces a non-functional secretion system, this mutation does not result in the uncontrolled release of effector molecules to the bacterial cell surface (**Fig. 11c**). This domain is predicted to be involved in terminating *cag* T4SS pilus-like structure assembly (272), suggesting that disruption of the CagH FlgK-like domain is not required for 'endopilus' termination. In combination with previous reports (275), these studies demonstrate that CagH-dependent regulation of pilus-like structure biogenesis is essential for *cag* T4SS translocation channel integrity.

APEX2-proximinty labeling of the *cag* T4SS

To further investigate *cag* T4SS architecture and potential structural changes that trigger cargo translocation, I adapted APEX2-dependent proximity labeling to identify protein-protein interactions within the secretion system. Identification of protein interactions can provide information regarding subcomplex assembly and apparatus biogenesis to improve our understanding of machinery topology and protein function. To perform APEX2-dependent proximity labeling, I first generated *H. pylori* strains expressing epitope tagged APEX2 (APEX2) or epitope tagged APEX2 translational fusions to CagH (APEX2-CagH) or to CagF (APEX2-CagF), a chaperone-like protein required for the translocation of the oncoprotein CagA (272, 322, 323). APEX constructs were inserted into the *ureA* locus via homologous recombination under constitutive expression (**Fig. 12a**). PCR mapping was used to confirm proper insertion of the APEX2 constructs into the *ureA* locus, and immunoblotting analysis using antibody against FLAG and HA epitope tags was used to confirm APEX2, APEX2-CagH, and APEX-CagF production and protein stability (**Fig. 12b**).

I conducted APEX2 proximity labeling experiments by growing *H. pylori* strains expressing APEX2 constructs in the presence of biotin-phenol and catalyzing the labeling reaction with the addition of H_2O_2. The biotin-phenoxyl radicals produced from this reaction travel short distances (approximately 10-20 nm) and interact with electron rich amino acid side chains of nearby proteins, allowing for the detection of both direct and transient protein interactions (314, 324). The reaction was stopped by the addition of quench buffer and after cell lysis, biotinylated proteins were enriched by affinity

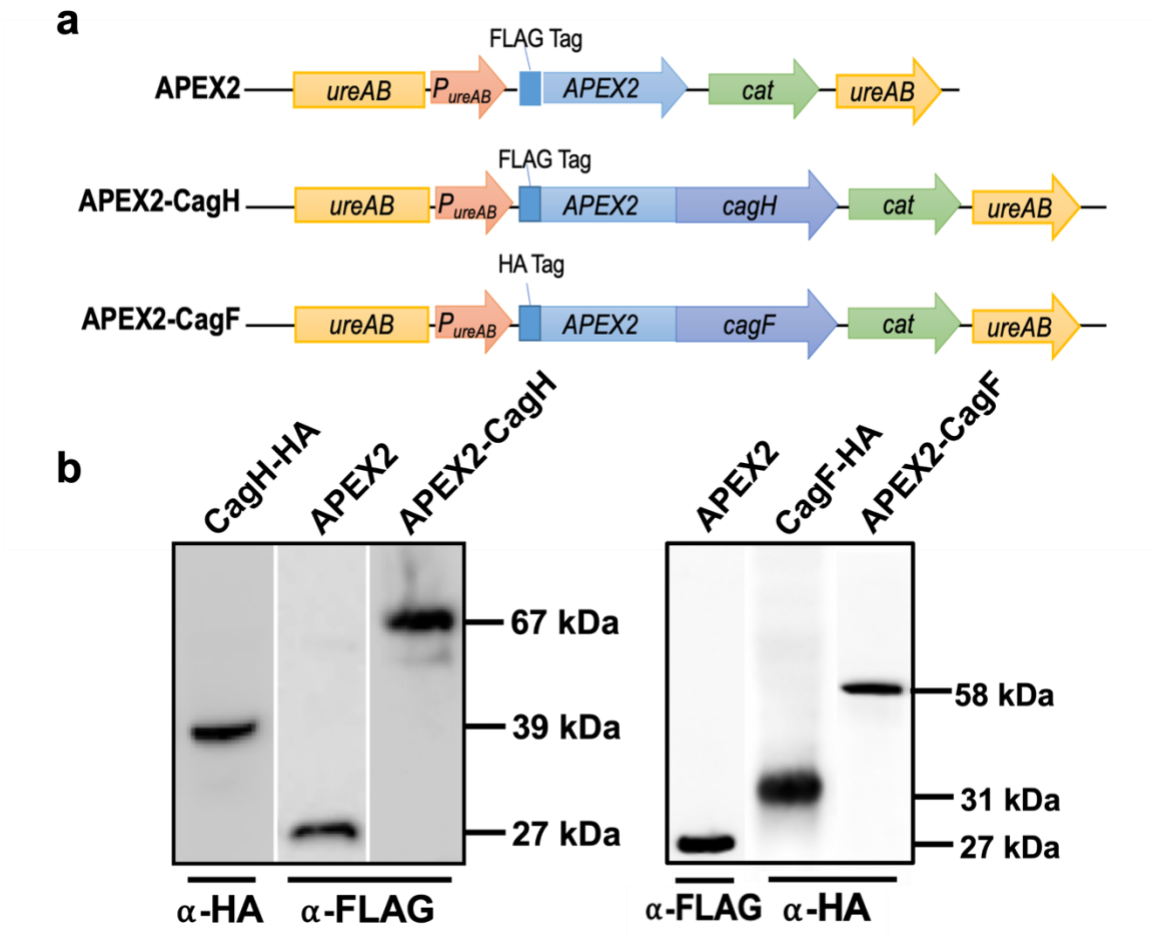

Figure 12. APEX2 constructs expressed in *H. pylori*.

(**a**) Schematic representing APEX2, APEX2-CagH, and APEX-CagF constructs inserted into the *ureAB* locus. APEX2 constructs are constitutively expressed under the *ureA* promoter. (**b**) APEX2, APEX-CagH, and APEX-CagF expression in *H. pylori* detected by immunoblotting with antibody against HA or FLAG epitope tags.

purification using streptavidin beads and analyzed by SDS page and silver staining techniques (313, 314) (**Fig. 13a**). Initial proximity labeling experiments were conducted in bacteria cultured in the absence of host cell contact. In these experiments, the APEX2-CagH strain labeled a unique set of proteins compared to strains expressing APEX2 or APEX2-CagF (**Fig. 13b**). Bands at the correct size of two verified CagH interaction partners [CagL and CagI (272)] were observed (**Fig. 13b**). Furthermore, immunoblotting with antibody against the oncoprotein CagA, which strongly interacts with CagF (322, 323), detected CagA in the streptavidin purified proteins isolated from the APEX2-CagF strain (**Fig. 13b**), further demonstrating that APEX2-dependent proximity labeling can successfully identify *H. pylori* protein-protein interactions.

APEX2 proximity labeling experiments were next performed in *H. pylori* co-cultured with gastric epithelial cells. While more optimization is required to resolve distinct protein-protein interactions, preliminary experiments demonstrate that protein labeling occurs in the presence of biotin phenol and multiple proteins are labeled in AGS cells infected with *H. pylori* expressing APEX2 constructs compared to uninfected cells (**Fig. 13c**). In future studies, the APEX2 proximity labeling approach will be used to identify *cag* T4SS protein-protein interactions that occur in different architectural states to further our understanding of how the *cag* T4SS assembles and functions to deliver immunostimulatory cargo into gastric cells.

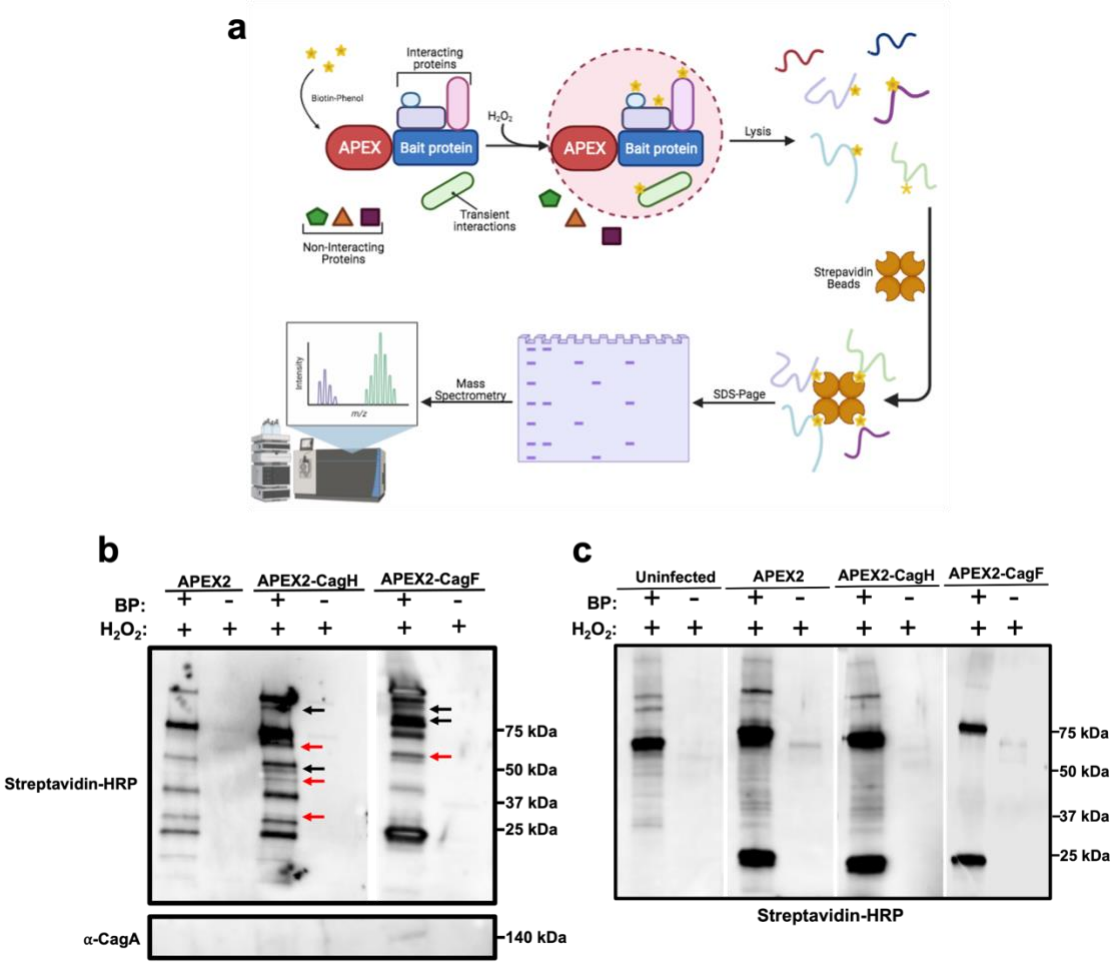

Figure 13. APEX2 proximity labeling in *H. pylori.*

(**a**) Schematic detailing proximity labeling in APEX2-expressing *H. pylori* strains. APEX2 is fused to a bait protein and addition of biotin-phenol and H_2O_2 catalyzes the production of biotin-phenoxyl radicals that label both direct and transient protein interactions. Biotinylated proteins are isolated using streptavidin and analyzed by SDS-page and mass spectrometry approaches. (**b**) Proximity labeling by *H. pylori* expressing APEX2 constructs in *H. pylori* analyzed by SDS-page. Compared to *H. pylori* expressing APEX2 or APEX2-CagF, *H. pylori* APEX2-CagH labeled a unique set of proximity partners. Red arrows indicate APEX-CagH (67 kDa) and APEX-CagF (58 kDa) self-labeling (Top left and right, respectively) and predicted CagH proximity partners CagL (26 kDa, bottom left) and CagI (41 kDa, middle left). Black arrows indicate unique proteins labeled by either APEX2-CagH or APEX2-CagF compared to the APEX2 control. (c) Proximity labeling in the indicated *H. pylori* strain co-cultured with gastric epithelial cells

4.5 Discussion

The *cag* T4SS is a complex nanomachine that integrates approximately 27 components to orchestrate effector transport into host cells. Of these components, several unique subunits are essential to *cag* T4SS function and fulfill a specialized role within the secretion system. In this study, colony immunoblotting techniques revealed that inactivation *cagZ* resulted in mislocalization of the protein effector CagA to the bacterial cell surface, while inactivation of *cagH* results in the uncontrolled release of CagA and dsDNA effectors to the bacterial cell surface (**Fig. 11**). CagZ is essential to *cag* T4SS function and acts as a regulator of Cagβ ATPase activity (320, 321). When CagZ is bound to Cagβ, the ATPase is trapped in its monomeric state, unable to provide the energy necessary for CagA translocation. After CagZ dissociates, the Cagβ hexamer forms, initiating ATPase activity and CagA translocation through the secretion channel. Additionally, in the absence of CagZ there is a loss of Cagβ stability, leading to markedly reduced Cagβ levels in the absence of CagZ (321). While Cagβ is required for CagA translocation, the secretion system OMC still assembles and translocates other effectors into host cells (251, 265, 320). Because a functional secretion apparatus assembles in the absence of CagZ and Cagβ, CagA enters the translocation channel via CagF tethering.

Colony immunoblotting revealed that CagH is also involved in controlling effector release across the bacterial cell surface. CagH is an essential component of the *cag* T4SS and previous studies have shown that it acts as a molecular ruler that governs polymerization and spatial distribution of *cag* T4SS filamentous structures at the cell-cell interface (272). In this study, I show that in the absence of host cell contact, inactivation of

cagH results in the mislocalization of both protein and nucleic acid effectors to the bacterial surface (**Fig. 11**), suggesting an important role in regulating cargo trafficking through the *cag* T4SS. CagC is a VirB2-like pilin subunit that assembles into a putative 'endopilus' in the periplasm to gate the secretion channel pore (**Chapter 2**). Because of CagC similarities to pilin subunits, it is possible CagH may be involved in regulating CagC 'endopilus' assembly (272, 325). Further investigation into the role of CagH in *cag* T4SS translocation channel gating revealed that the C-terminal domain is essential for maintaining channel integrity. Interestingly, another CagH variant in which a FlgK-like domain was disrupted did not exhibit the 'leaky' phenotype observed in the *cagH* and CagH-HA genetically complemented strains. This suggests that while the FlgK domain is essential for *cag* T4SS activity, CagH-dependent channel gating occurs via distinct mechanisms. The CagH FlgK domain shares similarities with multiple T3SS molecular rulers including the *Salmonella* FliK flagellar protein and *Yersinia pestis* YscP (275). These proteins are involved in determining flagellar hook length and T3SS needle length, suggesting a similar role for the CagH FlgK domain in governing the dimensions and spatial distribution of filamentous pilus-like structures (326, 327). If CagH is involved of the biogenesis of the putative CagC 'endopilus', disruption of the FlgK domain may result in improperly assembled CagC structures that cannot mediate effector molecule translocation but maintain gating activity within the *cag* T4SS pore.

This study also demonstrated that APEX2 proximity labeling can be successfully adapted to identify novel protein-protein interactions and to analyze *cag* T4SS architecture. Using localization and protein interaction studies, several subassemblies were identified within the *cag* T4SS including interactions among the ATPase Cagα-Cagα (328),

interaction between CagA and the cognate chaperone-like protein CagF (322, 323), and interactions between CagX, CagY, CagM, and CagT (259). Later studies confirmed that these four proteins along with Cag3 assemble to form the *cag* T4SS OMC (262, 264). Additionally, the non-essential Cag protein CagN has been shown to interact with CagV and CagY suggesting localization to the periplasmic side of the inner membrane. Interactions identified between the periplasmic portions of Cagβ, CagY, and CagL suggest that CagY may span the periplasmic and connect to the inner membrane and indicate that CagL is localized to the periplasmic side of the inner membrane (259). In combination with the current study, these observations highlight the importance of mapping protein-protein interactions within *cag* T4SS architecture.

Several different methods, including co-immunoprecipitation and yeast two-hybrid assays, have been used to identify interactions between Cag components. However, these methods present experimental limitations. Yeast two-hybrid screening approaches are performed in a eukaryotic system which may alter protein folding leading to difficulty in detecting true protein-protein interactions. Co-immunoprecipitation (co-IP) approaches to isolate native *H. pylori* complexes and subassemblies present additional challenges. One disadvantage is the limited availability of antibodies that specifically recognize bait or prey proteins. For many Cag components, monoclonal antibodies that recognize conformational epitopes are unavailable for immunopurification studies, and polyclonal antisersa used to detect co-IP components is often cross-reactive. Additionally, co-IP is unlikely to detect weak affinity or transient protein interactions. To overcome these complications, new protein-protein interaction methods utilizing enzyme catalyzed proximity labeling coupled with mass spectrometry have been developed as a high throughput approach to studying

protein interaction networks. One such method is APEX2-proximity labeling. APEX2 is an engineered ascorbate peroxidase that catalyzes the production of biotin-phenoxyl radicals generated in the presence of H_2O_2 that travel short distances and covalently interact with the side chains of nearby proteins (311-313), allowing for identification of direct, indirect, and transient protein-protein interactions. APEX2-dependent proximity labeling has been used in eukaryotic cells to resolve protein networks in *Saccharomyces cerevisiae,* identify proteins that constitute the mitochondrial matrix in living cells (311), define cell type and subcellular compartment-specific proteomes (329), and to investigate EGF receptor ligand binding, signal transduction, and regulation pathways (330). APEX2 proximity labeling has also been used in prokaryotic cells to map *Chlamydia trachomatis* host-pathogen interactions (318, 331-334), to map periplasmic proteins in *Mycobacteria* and *E. coli* (335), and to investigate T6SS biogenesis (316).

Preliminary studies performed using APEX2 fused to either CagH or CagF revealed direct and transient interactions *in situ*. While *H. pylori* APEX2-dependent labeling approaches require additional experimental optimization, this method will be beneficial to identifying both direct and transient protein-protein interaction within the *cag* T4SS in various architectural states. For example, this approach can be used to map protein interactions within the periplasm to identify 'stalk' and 'wing' components. Additionally, proteomic mapping approaches can be used to investigate structural changes that occur within the *cag* T4SS to trigger effector cargo translocation when *H. pylori* is in direct contact with gastric epithelial cells. These studies will provide insight into *cag* T4SS apparatus topology and will contribute to our understanding of how the *cag* T4SS functions to deliver immunostimulatory cargo into target cells.

CHAPTER 5. SUMMARY AND FUTURE DIRECTIONS

Over half of the global population is colonized by carcinogenic *H. pylori* which is directly responsible for 75% of global gastric cancer burden and approximately 700,000 deaths annually (9). Colonization by *H. pylori* results in chronic gastric inflammation that can progress to more severe gastric disease driven by host, bacterial, and environmental factors (13, 14, 58). Because of the significant cancer burden directly attributable to *H. pylori* colonization, and an alarming rise in antibiotic resistance, understanding the mechanisms underscoring *H. pylori* pathogenesis is an important area of research.

The strongest risk factor for development of gastric cancer is colonization by strains of *H. pylori* that harbor the *cag* type IV secretion system (T4SS) (22, 23). The *cag* T4SS forms a membrane-spanning nanomachine that delivers diverse immunostimulatory molecules, including the oncoprotein CagA, nucleic acid effectors, peptidoglycan fragments and LPS metabolites, into gastric epithelial cells. These effectors interact with various host cell signaling pathways resulting in altered host immune responses and changes to gastric cell morphology and function that stimulate the development of gastric cancer (260, 336). Recent cryo-ET and cryo-EM studies significantly advanced our understanding of *cag* T4SS architecture and structural organization (244, 253, 257, 258). However, the *cag* T4SS periplasmic and inner membrane complexes remain unresolved and how the secretion system delivers cargo into host cells is largely undefined. To better understand how *H. pylori* exploits *cag* T4SS activity to cause disease, the goal of this work was to investigate *cag* T4SS assembly and architecture to explore the mechanisms guiding effector translocation into target cells.

Recent studies demonstrate that in addition to the protein effector CagA, *H. pylori* translocates microbial DNA into host cells via *cag* T4SS activity (36-38, 40). In agreement with findings from earlier studies, I demonstrate that *H. pylori* translocates fragmented chromosomal DNA into host cells in a *cag* T4SS-dependent manner **(Fig. 2)**. A recent study from our lab demonstrated that DNA recruitment to the *cag* T4SS is linked to chromosomal replication and topological isomer disentanglement and identified several proteins including, FtsK and a VirD2-like relaxase, that are essential for DNA recruitment to the translocation channel (38). However, a complete mechanistic view of DNA transport and entry into host cells has not been delineated. In this study, I provide evidence that unlike DNA translocation by other T4SSs (337), the *cag* T4SS transports dsDNA that is exposed to the extracellular milieu before uptake into host cells. Nucleases that specifically target dsDNA or anti-DNA monoclonal antibodies significantly reduce TLR9 activation **(Fig. 2)**, indicating that DNA cargo is delivered to host cells in a novel two-step mechanism. Similarly, anti-BrdU monoclonal antibodies prevented TLR9 activation by labeled *H. pylori* **(Fig. 2)** further supporting a two-step secretion model. Translocated microbial DNA stimulates multiple nucleic acid receptors to elicit both pro- and anti-inflammatory immune responses that modulate gastric inflammation (40). Future studies to understand the mechanism of DNA transport by the *cag* T4SS and to analyze whether trans-kingdom DNA conjugation occurs *in vivo* remain important areas of research.

Additional studies analyzing effector transport through the *cag* T4SS revealed that CagC (a component essential to *cag* T4SS function (24, 273, 288)) plays an important role in translocation channel gating. Inactivation of *cagC* results in the uncontrolled release of both protein and dsDNA effectors to the bacterial cell surface and licenses the import of

exogenous molecules into the periplasm, indicating that CagC is responsible for regulating passage of substrate through the translocation channel **(Fig. 3).** I propose a model in which the VirB2-like pilin CagC polymerizes to form an 'endopilus' gate or plug that corresponds to the stalk region that bridges the IMC and PRC complexes in the *cag* T4SS machinery to control effector transport through the secretion channel.

CagC exhibits primary sequence similarities with other T4SS pilin subunits, including VirB2 from the prototypical *vir* T4SS in *Agrobacterium tumefaciens* and TraA of the *E. coli* F pilin (325). Previous studies report the presence of filamentous structures assembled at the *H. pylori*-gastric epithelial cell interface that are proposed to be an extracellular *cag* T4SS feature analogous to conjugative pili or the *A. tumefaciens* T-pilus (273, 275, 288, 295, 296). However, pilus structures cannot be isolated from bacterial co-cultures and thus, the molecular composition of these structures remains unresolved. Additionally, several inconsistencies have been identified that contradict the presence of a canonical *cag* T4SS pilus. Most notably, extracellular filaments are produced in absence of CagY at the same level as WT *H. pylori* (273, 295). CagY is an OMC structural component that forms the central translocation channel and is required for apparatus assembly (253, 257, 258, 274). It is therefore unlikely that an extracellular *cag* T4SS-associated pilus could form in the absence of a properly assembled OMC.

This study provides strong evidence for the assignment of CagC as a VirB2 pilin structural ortholog that assembles as an 'endopilus' in the *cag* T4SS. Analysis of CagC protein structure demonstrates the strong structural similarities between CagC and canonical T4SS pilins VirB2 and TraA (**Fig. 4**). However, complementation of *H. pylori cagC* with either *virB2* or *traA* did not restore *cag* T4SS function suggesting that CagC

fulfills a unique role within the *cag* T4SS **(Fig. 4)**. Protein interaction studies using a bacterial two-hybrid approach revealed direct interactions between CagC and components of both the IMC and PRC as well as direct CagC-CagC interactions. Taken together, these results suggest that CagC polymerizes in the periplasm to form an 'endopilus' that bridges the IMC and PRC and likely constitutes the unresolved 'stalk' structure identified by cryo-ET and cryo-EM (244, 262, 264) (**Fig. 5**). Additional structural and *in situ* imaging studies are required to test this compelling hypothesis. One possibility is to use cryo-ET to visualize *cag* T4SS architecture in the absence of CagC to define sub-complex structural changes. Future work should focus on determining whether CagC multimerizes to form an 'endopilus' or plug domain in the assembled secretion system. As related T4SS pilins form pentameric polymers within conjugative pilus structures (161, 290), I predict that CagC will form similar higher order sub-complexes. Future studies focused on determining the multimeric state of CagC will deepen our understanding of *cag* T4SS gating and channel integrity. Furthermore, delineating additional CagC interactions could provide supporting evidence that CagC exhibits 'endopilus' activity. For example, determining whether CagC interacts with components such as CagL, a predicted homolog of the *vir* T4SS minor pilin VirB5, would strengthen the argument that CagC assembles into endopilus-like structures. Recent cryo-EM studies of the nearly intact R388 conjugation system revealed that the 'stalk' is composed of VirB5 (CagL) and VirB6 (CagW) (1, 154). Similarly, the *A. tumefaciens* T-pilus uses VirB6 pentamers as the base on which to assemble VirB2 (CagC) structures (161). Determining if CagC interacts with CagL or other *cag* T4SS components will provide insight into how the proposed 'endopilus' assembles. Additionally, future studies should focus on identifying the mechanism of CagC gating within the secretion

pore. One possibility is that CagC directly binds substrate at the inner membrane complex, triggering dynamic assembly and disassembly of an 'endopilus' to facilitate cargo transport between the IMC and PRC. In Chapter 3, I demonstrate that CagC weakly binds dsDNA substrates, strengthening this hypothesis. Future studies to resolve CagC-dsDNA interactions would provide insight effector transit through the translocation channel. Another possibility is that CagC acts plug in the outer membrane pore that effectively prevents effector release by inactive architectures. Protein interaction studies to specifically investigate how subunit interactions shift between active and inactive machineries may pinpoint CagC localization within the assembled apparatus.

The observation that CagC weakly interacts with dsDNA and with components of the *cag* T4SS PRC led to the hypothesis that additional subunits bind DNA to facilitate substrate transport through the translocation channel. In Chapter 2, I provide evidence that π-π interactions between a series of arginine (Arg 36, 28, 278, and 295) and tryptophan (Trp 58) residues in the periplasmic portion of CagX creates a positively charged basic patch predicted to be a possible DNA binding site. Biochemical analyses confirmed that CagX strongly binds both dsDNA and ssDNA and exhibits a preference for dsDNA substrates (**Fig. 6**). Systematic point mutations to ablate π-stacking interactions identified this region as an essential mediator of the CagX-DNA interaction (**Fig. 8**). Single point mutation of Arg 278 resulted in a significant reduction in CagX DNA binding capacities that were further reduced when Arg 36, 38, and 278 were mutated in tandem. Disruption of Arg 36, 38 and 278 abolished TLR9 activation but had no effect on IL-8 stimulation, suggesting distinct translocation mechanisms for DNA versus protein substrates. These studies provide evidence of a secondary substrate selection mechanism in the *cag* T4SS

PRC that is mediated by intermolecular π-π interactions among adjacent CagX subunits. Point mutation of Arg 295 similarly reduced DNA binding *in vitro* and abolished TLR9 activation. However, Arg 295 was critical for CagX stability *in vivo*, resulting in significantly reduced IL-8 stimulation compared to the WT strain. Based on these data, I propose that Arg 295 constitutes a CagY binding site. Future studies will use protein-protein interaction approaches to define the impact of arginine residues on CagX-CagY subcomplex assembly.

Recent studies investigating the architecture of the *cag* T4SS revealed a striking symmetry mismatch between adjacent apparatus subassemblies (257, 262, 264). Architectural symmetry mismatch is a common feature among T4SSs (39); however, the biological significance and role of symmetry mismatch in interkingdom effector translocation remains unresolved. Within the *cag* T4SS, symmetry mismatch between the OMC (14-fold) and the PRC (17-fold) leaves three unoccupied copies of CagX harboring π-π interactions that mediate DNA binding (**Fig. 8 and Fig. 10)**. Thus, these studies identified a novel role for architectural asymmetry in orchestrating substrate selection and DNA transport through the distal secretion channel. Secondary substrate selection mechanisms have also been proposed in the *A. tumefaciens vir* T4SS (309). Mutations in the VirB9 N-terminal domain (the *vir* T4SS structural homolog of CagX) selectively block translocation of either the IncQ plasmid or T-DNA substrates suggesting that VirB9 serves as a secondary checkpoint that governs substrate passage through the translocation channel. This process was predicted to be mediated by VirB9 interaction with the relaxasome or with energy sensing subunits in other secretion system subassemblies (309), raising the possibility that substrate selection by CagX occurs in a similar manner. However, inter-

subunit π-π interactions are not observed among CagX structural homologs including VirB9 (**Fig. 7**), suggesting that PRC-mediated DNA binding is an evolutionary innovation in the *cag* T4SS.

One possibility is that CagX-DNA interactions initiate a conformational change that supports nucleic acid movement through the translocation channel. Similar conformational changes triggered by substrate binding to ATPase machinery has been described in other T4SSs including the *vir* T4SS and the *dot/icm* T4SS (155, 338) and while the data presented here proposes a secondary substrate selection step within the PRC, similar structural changes could be induced. Analysis of CagX-DNA binding capacities revealed a cooperative interaction, indicating that DNA binding affinity increases after the first substrate binds, potentially due to conformational changes that open subsequent binding sites. Cryo-EM studies to visualize CagX subunits complexed to DNA are currently underway. Alternatively, CagX binding may induce DNA conformational changes that promote the binding of a second protein or subunit to increase PRC DNA affinity (339, 340). DNA conformational changes induced by CagX binding could be monitored by circular dichroism (341) or in real-time using single molecule fluorescence resonance energy transfer (smFRET) assays (342, 343). Determining how CagX and/or DNA conformations change during substrate transport will provide valuable insight into secondary selection mechanisms in the translocation channel. Defining how CagX orchestrates substrate selection and transport through the distal secretion channel is an important area of research that will significantly improve our understanding of trans-kingdom DNA conjugation.

In addition to CagC, colony immunoblotting revealed that in the absence of CagH, a unique component that regulates the biogenesis of surface-associated filaments (272), CagA and dsDNA effectors are aberrantly shuttled to the bacterial cell surface in a host cell contact-independent manner (**Fig 11**). In Chapter 2, I hypothesize that CagC forms an 'endopilus' that spans the periplasm and controls effector passage through the secretion channel. Because CagH is associated with cell surface structure biogenesis, loss of CagH may promote improper 'endopilus' assembly, thus preventing proper channel gating. Further studies revealed that the CagH C-terminal domain is essential for channel gating activity (**Fig. 11**). Therefore, I predict that formation of the putative CagC 'endopilus' is dependent on CagH membrane localization and potential interaction with CagC protomers in the inner membrane. Surprisingly, while disruption of the CagH FlgK domain results in a non-functional secretion system, translocation channel gating remains unaffected (**Fig. 11**). Because the FlgK domain is predicted to be involved in terminating pilus production, I hypothesize that strains lacking this functional domain produce a CagC 'endopilus' that is uncapped or an improper size, thus trapping the gating mechanism in a closed state and blocking the release of effectors across the bacterial outer membrane. To further investigate the role of CagH in possible 'endopilus' biogenesis, future studies could analyze the *in situ cag* T4SS architecture by cryo-ET to compare structures assembled by the *cagH* and CagH-HAΔFlgK strains. Future investigations to resolve the role of CagH and CagC in secretion channel biogenesis will continue to provide insight into the mechanisms by which expanded T4SS architectures deliver bacterial effectors into host cells.

To continue my investigation of *cag* T4SS topology, I employed proteomic mapping by APEX2 proximity labeling to identify novel protein-protein interactions in the

cag T4SS. Because biotinylation-based proximity labeling approaches can label both direct and transient interactions (313), this strategy represents a useful tool for studying the architecture of large and dynamic structures like the *cag* T4SS. Preliminary APEX2 proximity labeling experiments demonstrate that this method can be used to label distinct proteins in strains expressing APEX2-CagH compared to APEX2 or APEX2-CagF. Various protein subsets were labeled under conditions in which *H. pylori* was cultured in either the presence or absence of host cell contact (**Fig. 13**) providing the experimental framework for deciphering structural changes that enable effector transport. This method has been previously used to investigate protein interactions in eukaryotic cells (314, 315, 324, 329), and more recently in prokaryotes including *Chlamydia* and *E. coli* (316, 318, 335). The current study is the first report of this system in *H. pylori*, opening a wide variety of new research directions. For example, proximity labeling of the *cag* T4SS periplasmic apparatus will help identify components localized to unresolved structures observed in the cryo-ET and cryo-EM maps, including the stalk (**Chapter 2**) and 'collar' structure on the periphery of the PRC. Similarly, APEX2 proximity labeling could possibly identify the unknown components of the OMC I- and O-layers. This approach can also be used to better understand the role of Cag components with unknown functions in the secretion system. Finally, APEX2 proximity labeling could be used to map CagA interactions during trafficking through the *cag* T4SS channel. Proximity labeling of CagA when *H. pylori* is co-cultured with gastric cells could also identify additional host cell receptors involved in *H. pylori cag* T4SS adhesion and CagA uptake into target cells.

In addition to APEX2 proximity labeling, other biotinylation based proximity labeling systems could be useful tools in investigating *cag* T4SS activity. While this work

and other recent studies have improved our understanding of effector transport through the *cag* T4SS, the mechanism of effector transport across the outer membrane and delivery into host cells remain incompletely defined. The *cag* T4SS injects effector molecules into a variety of host cell lines although our preliminary data suggest that not all effectors are delivered into every human cell line, suggesting the presence of cell type-specific receptors that mediate uptake of different *cag* T4SS effectors. Unsurprisingly, *H. pylori* robustly delivers all known effector molecules into gastric epithelial cells; however, while DNA is translocated into both HEK293 and 293T cell lines, HEK293 cells are resistant to CagA translocation and significantly reduced levels of CagA are transported into 293T cells. Proximity labeling by horseradish peroxidase (HRP) fused to the glycan binding domain of wheat germ agglutinin (WGA-HRP) could be used to compare cell surface receptors on different cell types that mediate effector translocation into host cells. WGA binds N-acetylglucosamine (GlcNAc) in the bacterial cell wall, coating the *H. pylori* membrane and like APEX2, HRP catalyzes the production of biotin phenoxyl radicals that travel short distances and label nearby proteins (344). This proximity labeling method would thus provide insight into differences in host cell surface proteins in cell lines that are permissive or restrictive for CagA translocation. WGA-HRP proximity labeling could also be used to identify extracellular features of the *cag* T4SS including bacterial cell surface proteins that interact with host cells. Additionally, permeabilization of *H. pylori* allows WGA to enter and bind peptidoglycan in the periplasm and label periplasmic *H. pylori* proteins (345), providing another method for mapping unresolved structures in the *cag* T4SS apparatus.

Lastly, bacterial two-hybrid screening is a simple, fast method to identify direct *cag* T4SS protein interactions in a surrogate prokaryotic system. In this study, I show that the

bacterial two-hybrid approach can be used to identify *H. pylori* protein-protein interactions and I used this approach to demonstrate CagC interaction with *cag* T4SS components localized to IMC and PRC which contributed to elucidating the role of CagC as a novel gating mechanism. Expansion of our bacterial two-hybrid screening library to map protein interactions within the *cag* T4SS will continue to provide insight into the function of unknown components. Several proteins including CagD, CagN, CagS, CagP, and CagQ share no homology to prototypical T4SS components and are not required for known *cag* T4SS-dependent phenotypes (1). Nonetheless, these subunits may represent unrecognized effector proteins or may be integrated into the *cag* T4SS apparatus as an accessory factor. Using the bacterial two-hybrid assay to screen interactions with other *cag* T4SS components will help determine where these proteins are localized in the secretion system, thereby expanding our understanding of *cag* T4SS architecture and function.

In conclusion, the data presented here improves our understanding of the molecular mechanisms underlying *cag* T4SS-mediated effector molecule translocation. I used genetic and biochemical approaches to identify a novel gating mechanism that regulates effector substrate passage through the *cag* translocation channel. Additionally, I demonstrate that architectural asymmetry between the outer membrane complex and periplasmic ring complex facilitates DNA substrate selection via intermolecular π-π interactions harbored between adjacent CagX subunits. Further investigation into the role of *cag* T4SS components identified a role of CagH and CagZ in controlling effector release across the bacterial outer membrane. I also provide evidence that proper localization of CagH within the secretion system is essential to *cag* T4SS gating mechanisms. Finally, I adapted APEX2 proximity labeling as a new approach to investigate *cag* T4SS assembly and architecture

in situ. Collectively, this work provides insight into mechanisms governing trans-kingdom cargo transport through expanded T4SS machineries, and the experimental framework for understanding how the remarkable *cag* T4SS nanomachine coordinates diverse effector molecule delivery to stimulate the development of infection-associated malignancies.

REFERENCES

1. S. Backert, N. Tegtmeyer, W. Fischer, Composition, structure and function of the Helicobacter pylori cag pathogenicity island encoded type IV secretion system. *Future Microbiol* **10**, 955-965 (2015).
2. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
3. R. Evans *et al.*, Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 10.1101/2021.10.04.463034, 2021.2010.2004.463034 (2022).
4. J. K. Y. Hooi *et al.*, Global Prevalence of Helicobacter pylori Infection: Systematic Review and Meta-Analysis. *Gastroenterology* **153**, 420-429 (2017).
5. K. Sugano *et al.*, Kyoto global consensus report on Helicobacter pylori gastritis. *Gut* **64**, 1353-1367 (2015).
6. S. Suerbaum, P. Michetti, Helicobacter pylori infection. *N Engl J Med* **347**, 1175-1186 (2002).
7. R. M. Peek, Jr., M. J. Blaser, Helicobacter pylori and gastrointestinal tract adenocarcinomas. *Nat Rev Cancer* **2**, 28-37 (2002).
8. E. J. Kuipers, J. C. Thijs, H. P. Festen, The prevalence of Helicobacter pylori in peptic ulcer disease. *Aliment Pharmacol Ther* **9 Suppl 2**, 59-69 (1995).
9. H. Sung *et al.*, Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209-249 (2021).
10. Anonymous, Schistosomes, liver flukes and Helicobacter pylori. *IARC Monogr Eval Carcinog Risks Hum* **61**, 1-241 (1994).
11. J. Ferlay *et al.*, Global cancer observatory: cancer today. International Agency for Research on Cancer. *Lyon, France* (2020).
12. P. Malfertheiner *et al.*, Helicobacter pylori infection. *Nature Reviews Disease Primers* **9**, 19 (2023).
13. P. Correa, M. B. Piazuelo, The gastric precancerous cascade. *J Dig Dis* **13**, 2-9 (2012).
14. P. Correa, W. Haenszel, C. Cuello, S. Tannenbaum, M. Archer, A model for gastric cancer epidemiology. *Lancet* **2**, 58-60 (1975).
15. I. Mortier-Barriere *et al.*, A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell* **130**, 824-836 (2007).
16. J. M. Liou *et al.*, Screening and eradication of Helicobacter pylori for gastric cancer prevention: the Taipei global consensus. *Gut* **69**, 2093-2112 (2020).
17. J. C. Fann *et al.*, Personalized risk assessment for dynamic transition of gastric neoplasms. *J Biomed Sci* **25**, 84 (2018).
18. T. L. Cover, Helicobacter pylori Diversity and Gastric Cancer Risk. *mBio* **7**, e01869-01815 (2016).
19. S. Suerbaum *et al.*, Free recombination within Helicobacter pylori. *Proc Natl Acad Sci U S A* **95**, 12619-12624 (1998).
20. G. Morelli *et al.*, Microevolution of Helicobacter pylori during prolonged infection of single hosts and within families. *PLoS Genet* **6**, e1001036 (2010).

21. M. J. Blaser, D. E. Berg, Helicobacter pylori genetic diversity and risk of human disease. *J Clin Invest* **107**, 767-773 (2001).
22. N. S. Akopyants *et al.*, Analyses of the cag pathogenicity island of Helicobacter pylori. *Mol Microbiol* **28**, 37-53 (1998).
23. S. Censini *et al.*, cag, a pathogenicity island of Helicobacter pylori, encodes type I-specific and disease-associated virulence factors. *Proc Natl Acad Sci U S A* **93**, 14648-14653 (1996).
24. W. Fischer *et al.*, Systematic mutagenesis of the Helicobacter pylori cag pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8. *Mol Microbiol* **42**, 1337-1348 (2001).
25. P. Olbermann *et al.*, A global overview of the genetic and functional diversity in the Helicobacter pylori cag pathogenicity island. *PLoS Genet* **6**, e1001069 (2010).
26. M. Hatakeyama, Oncogenic mechanisms of the Helicobacter pylori CagA protein. *Nat Rev Cancer* **4**, 688-694 (2004).
27. S. Odenbreit *et al.*, Translocation of Helicobacter pylori CagA into gastric epithelial cells by type IV secretion. *Science* **287**, 1497-1500 (2000).
28. M. Sigal *et al.*, Helicobacter pylori Activates and Expands Lgr5(+) Stem Cells Through Direct Colonization of the Gastric Glands. *Gastroenterology* **148**, 1392-1404 e1321 (2015).
29. L. Buti *et al.*, Helicobacter pylori cytotoxin-associated gene A (CagA) subverts the apoptosis-stimulating protein of p53 (ASPP2) tumor suppressor pathway of the host. *Proc Natl Acad Sci U S A* **108**, 9238-9243 (2011).
30. M. Asahi *et al.*, Helicobacter pylori CagA protein can be tyrosine phosphorylated in gastric epithelial cells. *The Journal of experimental medicine* **191**, 593-602 (2000).
31. H. Higashi *et al.*, SHP-2 tyrosine phosphatase as an intracellular target of Helicobacter pylori CagA protein. *Science* **295**, 683-686 (2002).
32. E. D. Segal, J. Cha, J. Lo, S. Falkow, L. S. Tompkins, Altered states: involvement of phosphorylated CagA in the induction of host cellular growth changes by Helicobacter pylori. *Proc Natl Acad Sci U S A* **96**, 14559-14564 (1999).
33. Y. Churin *et al.*, Helicobacter pylori CagA protein targets the c-Met receptor and enhances the motogenic response. *J Cell Biol* **161**, 249-255 (2003).
34. F. Wang *et al.*, CagA promotes proliferation and inhibits apoptosis of GES-1 cells by upregulating TRAF1/4-1BB. *Mol Med Rep* **16**, 1262-1268 (2017).
35. M. R. Amieva *et al.*, Disruption of the epithelial apical-junctional complex by Helicobacter pylori CagA. *Science* **300**, 1430-1434 (2003).
36. M. G. Varga *et al.*, Pathogenic Helicobacter pylori strains translocate DNA and activate TLR9 via the cancer-associated cag type IV secretion system. *Oncogene* 10.1038/onc.2016.158 (2016).
37. M. G. Varga, R. M. Peek, DNA Transfer and Toll-like Receptor Modulation by Helicobacter pylori. *Curr Top Microbiol Immunol* **400**, 169-193 (2017).
38. P. P. Damke, C. R. Wood, C. L. Shaffer, *Helicobacter pylori* provokes STING immunosurveillance via trans-kingdom conjugation. *bioRxiv* 10.1101/2022.06.29.498044, 2022.2006.2029.498044 (2022).
39. M. E. Ryan, P. P. Damke, C. L. Shaffer, DNA Transport through the Dynamic Type IV Secretion System. *Infect Immun* 10.1128/iai.00436-22, e0043622 (2023).

40. M. G. Varga *et al.*, TLR9 activation suppresses inflammation in response to Helicobacter pylori infection. *Am J Physiol Gastrointest Liver Physiol* **311**, G852-G858 (2016).
41. M. Chamaillard *et al.*, An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid. *Nat Immunol* **4**, 702-707 (2003).
42. S. E. Girardin *et al.*, Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. *Science* **300**, 1584-1587 (2003).
43. S. E. Girardin *et al.*, Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* **278**, 8869-8872 (2003).
44. J. Viala *et al.*, Nod1 responds to peptidoglycan delivered by the Helicobacter pylori cag pathogenicity island. *Nat Immunol* **5**, 1166-1174 (2004).
45. S. Zimmermann *et al.*, ALPK1- and TIFA-Dependent Innate Immune Response Triggered by the Helicobacter pylori Type IV Secretion System. *Cell Rep* **20**, 2384-2395 (2017).
46. L. Pfannkuch *et al.*, ADP heptose, a novel pathogen-associated molecular pattern identified in Helicobacter pylori. *FASEB J* **33**, 9087-9099 (2019).
47. A. Gall, R. G. Gaudet, S. D. Gray-Owen, N. R. Salama, TIFA Signaling in Gastric Epithelial Cells Initiates the cag Type 4 Secretion System-Dependent Innate Immune Response to Helicobacter pylori Infection. *mBio* **8** (2017).
48. S. C. Stein *et al.*, Helicobacter pylori modulates host cell responses by CagT4SS-dependent translocation of an intermediate metabolite of LPS inner core heptose biosynthesis. *PLoS Pathog* **13**, e1006514 (2017).
49. A. Yuan, J. J. Chen, P. L. Yao, P. C. Yang, The role of interleukin-8 in cancer cells and microenvironment interaction. *Front Biosci* **10**, 853-865 (2005).
50. J. Fu *et al.*, TRAF-interacting protein with forkhead-associated domain (TIFA) transduces DNA damage-induced activation of NF-kappaB. *J Biol Chem* **293**, 7268-7280 (2018).
51. S. H. Phadnis, D. Ilver, L. Janzon, S. Normark, T. U. Westblom, Pathological significance and molecular characterization of the vacuolating toxin gene of Helicobacter pylori. *Infect Immun* **62**, 1557-1565 (1994).
52. T. L. Cover, M. J. Blaser, Purification and characterization of the vacuolating toxin from Helicobacter pylori. *J Biol Chem* **267**, 10570-10575 (1992).
53. J. L. Rhead *et al.*, A new Helicobacter pylori vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. *Gastroenterology* **133**, 926-936 (2007).
54. J. C. Atherton, R. M. Peek, Jr., K. T. Tham, T. L. Cover, M. J. Blaser, Clinical and pathological importance of heterogeneity in vacA, the vacuolating cytotoxin gene of Helicobacter pylori. *Gastroenterology* **112**, 92-99 (1997).
55. T. L. Cover, U. S. Krishna, D. A. Israel, R. M. Peek, Jr., Induction of gastric epithelial cell apoptosis by Helicobacter pylori vacuolating cytotoxin. *Cancer Res* **63**, 951-957 (2003).
56. E. Papini *et al.*, Cellular vacuoles induced by Helicobacter pylori originate from late endosomal compartments. *Proc Natl Acad Sci U S A* **91**, 9720-9724 (1994).
57. A. Fujikawa *et al.*, Mice deficient in protein tyrosine phosphatase receptor type Z are resistant to gastric ulcer induction by VacA of Helicobacter pylori. *Nat Genet* **33**, 375-381 (2003).

58. L. E. Wroblewski, R. M. Peek, Jr., K. T. Wilson, Helicobacter pylori and gastric cancer: factors that modulate disease risk. *Clin Microbiol Rev* **23**, 713-739 (2010).
59. J. L. Guruge *et al.*, Epithelial attachment alters the outcome of Helicobacter pylori infection. *Proc Natl Acad Sci U S A* **95**, 3925-3930 (1998).
60. T. Boren, P. Falk, K. A. Roth, G. Larson, S. Normark, Attachment of Helicobacter pylori to human gastric epithelium mediated by blood group antigens. *Science* **262**, 1892-1895 (1993).
61. P. Karimi, F. Islami, S. Anandasabapathy, N. D. Freedman, F. Kamangar, Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. *Cancer Epidemiol Biomarkers Prev* **23**, 700-713 (2014).
62. Y. Yamaoka *et al.*, Helicobacter pylori outer membrane proteins and gastroduodenal disease. *Gut* **55**, 775-781 (2006).
63. J. Mahdavi *et al.*, Helicobacter pylori SabA adhesin in persistent infection and chronic inflammation. *Science* **297**, 573-578 (2002).
64. M. G. Varga *et al.*, Immunostimulatory membrane proteins potentiate H. pylori-induced carcinogenesis by enabling CagA translocation. *Gut Microbes* **13**, 1-13 (2021).
65. M. Sugimoto, T. Ohno, D. Y. Graham, Y. Yamaoka, Gastric mucosal interleukin-17 and -18 mRNA expression in Helicobacter pylori-induced Mongolian gerbils. *Cancer Sci* **100**, 2152-2159 (2009).
66. A. T. Franco *et al.*, Regulation of gastric carcinogenesis by Helicobacter pylori virulence factors. *Cancer Res* **68**, 379-387 (2008).
67. F. H. Tabassam, D. Y. Graham, Y. Yamaoka, Helicobacter pylori activate epidermal growth factor receptor- and phosphatidylinositol 3-OH kinase-dependent Akt and glycogen synthase kinase 3beta phosphorylation. *Cell Microbiol* **11**, 70-82 (2009).
68. A. Javaheri *et al.*, Helicobacter pylori adhesin HopQ engages in a virulence-enhancing interaction with human CEACAMs. *Nat Microbiol* **2**, 16189 (2016).
69. V. Koniger *et al.*, Helicobacter pylori exploits human CEACAMs via HopQ for adherence and translocation of CagA. *Nat Microbiol* **2**, 16188 (2016).
70. Q. A. Nguyen, L. Schmitt, R. Mejias-Luque, M. Gerhard, Effects of Helicobacter pylori adhesin HopQ binding to CEACAM receptors in the human stomach. *Front Immunol* **14**, 1113478 (2023).
71. M. H. Feige, O. Sokolova, A. Pickenhahn, G. Maubach, M. Naumann, HopQ impacts the integrin alpha5beta1-independent NF-kappaB activation by Helicobacter pylori in CEACAM expressing cells. *Int J Med Microbiol* **308**, 527-533 (2018).
72. K. Taxauer *et al.*, Engagement of CEACAM1 by Helicobacterpylori HopQ Is Important for the Activation of Non-Canonical NF-kappaB in Gastric Epithelial Cells. *Microorganisms* **9** (2021).
73. K. S. Johnson, K. M. Ottemann, Colonization, localization, and inflammation: the roles of H. pylori chemotaxis in vivo. *Curr Opin Microbiol* **41**, 51-57 (2018).
74. C. Josenhans, A. Labigne, S. Suerbaum, Comparative ultrastructural and functional studies of Helicobacter pylori and Helicobacter mustelae flagellin mutants: both flagellin subunits, FlaA and FlaB, are necessary for full motility in Helicobacter species. *J Bacteriol* **177**, 3010-3020 (1995).

75. S. Schreiber *et al.*, The spatial orientation of Helicobacter pylori in the gastric mucus. *Proc Natl Acad Sci U S A* **101**, 5024-5029 (2004).
76. E. Andersen-Nissen *et al.*, Evasion of Toll-like receptor 5 by flagellated bacteria. *Proc Natl Acad Sci U S A* **102**, 9247-9252 (2005).
77. S. K. Lee *et al.*, Helicobacter pylori flagellins have very low intrinsic activity to stimulate human gastric epithelial cells via TLR5. *Microbes Infect* **5**, 1345-1356 (2003).
78. K. A. Eaton, C. L. Brooks, D. R. Morgan, S. Krakowka, Essential role of urease in pathogenesis of gastritis induced by Helicobacter pylori in gnotobiotic piglets. *Infect Immun* **59**, 2470-2475 (1991).
79. D. L. Weeks, S. Eskandari, D. R. Scott, G. Sachs, A H+-gated urea channel: the link between Helicobacter pylori urease and gastric colonization. *Science* **287**, 482-485 (2000).
80. E. M. El-Omar *et al.*, Increased risk of noncardia gastric cancer associated with proinflammatory cytokine gene polymorphisms. *Gastroenterology* **124**, 1193-1201 (2003).
81. E. M. El-Omar *et al.*, Interleukin-1 polymorphisms associated with increased risk of gastric cancer. *Nature* **404**, 398-402 (2000).
82. M. Takashima, T. Furuta, H. Hanai, H. Sugimura, E. Kaneko, Effects of Helicobacter pylori infection on gastric acid secretion and serum gastrin levels in Mongolian gerbils. *Gut* **48**, 765-773 (2001).
83. I. R. Hwang *et al.*, Effect of interleukin 1 polymorphisms on gastric mucosal interleukin 1beta production in Helicobacter pylori infection. *Gastroenterology* **123**, 1793-1803 (2002).
84. J. E. Crabtree, T. M. Shallcross, R. V. Heatley, J. I. Wyatt, Mucosal tumour necrosis factor alpha and interleukin-6 in patients with Helicobacter pylori associated gastritis. *Gut* **32**, 1473-1477 (1991).
85. K. Oguma *et al.*, Activated macrophages promote Wnt signalling through tumour necrosis factor-alpha in gastric tumour cells. *EMBO J* **27**, 1671-1681 (2008).
86. S. Tsugane, Salt, salted food intake, and risk of gastric cancer: epidemiologic evidence. *Cancer Sci* **96**, 1-6 (2005).
87. S. A. Lee *et al.*, Effect of diet and Helicobacter pylori infection to the risk of early gastric cancer. *J Epidemiol* **13**, 162-168 (2003).
88. K. Shikata *et al.*, A prospective study of dietary salt intake and gastric cancer incidence in a defined Japanese population: the Hisayama study. *Int J Cancer* **119**, 196-201 (2006).
89. D. G. Beevers, G. Y. Lip, A. D. Blann, Salt intake and Helicobacter pylori infection. *J Hypertens* **22**, 1475-1477 (2004).
90. A. Gamboa-Dominguez *et al.*, Salt and stress synergize H. pylori-induced gastric lesions, cell proliferation, and p21 expression in Mongolian gerbils. *Dig Dis Sci* **52**, 1517-1526 (2007).
91. S. Kato *et al.*, High salt diets dose-dependently promote gastric chemical carcinogenesis in Helicobacter pylori-infected Mongolian gerbils associated with a shift in mucin production from glandular to surface mucous cells. *Int J Cancer* **119**, 1558-1566 (2006).

92. K. Nozaki *et al.*, Synergistic promoting effects of Helicobacter pylori infection and high-salt diet on gastric carcinogenesis in Mongolian gerbils. *Jpn J Cancer Res* **93**, 1083-1089 (2002).
93. S. Zhang *et al.*, Hyperosmotic stress enhances interleukin-1beta expression in Helicobacter pylori-infected murine gastric epithelial cells in vitro. *J Gastroenterol Hepatol* **21**, 759-766 (2006).
94. J. Sun *et al.*, Effect of NaCl and Helicobacter pylori vacuolating cytotoxin on cytokine expression and viability. *World J Gastroenterol* **12**, 2174-2180 (2006).
95. J. T. Loh, V. J. Torres, T. L. Cover, Regulation of Helicobacter pylori cagA expression in response to salt. *Cancer Res* **67**, 4709-4715 (2007).
96. H. Gancz, K. R. Jones, D. S. Merrell, Sodium chloride affects Helicobacter pylori growth and gene expression. *J Bacteriol* **190**, 4100-4105 (2008).
97. E. Grohmann, P. J. Christie, G. Waksman, S. Backert, Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol Microbiol* **107**, 455-471 (2018).
98. E. Cascales, P. J. Christie, The versatile bacterial type IV secretion systems. *Nature reviews. Microbiology* **1**, 137-149 (2003).
99. M. Bhatty, J. A. Laverde Gomez, P. J. Christie, The expanding bacterial type IV secretion lexicon. *Res Microbiol* **164**, 620-639 (2013).
100. P. J. Christie, K. Atmakuri, V. Krishnamoorthy, S. Jakubowski, E. Cascales, Biogenesis, Architecture, and Function of Bacterial Type IV Secretion Systems. *Annual review of microbiology* **59**, 10.1146/annurev.micro.1158.030603.123630 (2005).
101. T. R. D. Costa *et al.*, Type IV secretion systems: Advances in structure, function, and activation. *Mol Microbiol* **115**, 436-452 (2021).
102. E. Cascales, P. J. Christie, Definition of a Bacterial Type IV Secretion Pathway for a DNA Substrate. *Science* **304**, 1170-1173 (2004).
103. J. Lederberg, E. L. Tatum, Gene recombination in Escherichia coli. *Nature* **158**, 558 (1946).
104. K. Wallden, A. Rivera-Calzada, G. Waksman, Type IV secretion systems: versatility and diversity in function. *Cell Microbiol* **12**, 1203-1212 (2010).
105. C. E. Alvarez-Martinez, P. J. Christie, Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* **73**, 775-808 (2009).
106. G. Waksman, R. Fronzes, Molecular architecture of bacterial type IV secretion systems. *Trends in biochemical sciences* **35**, 691-698 (2010).
107. G. Waksman, From conjugation to T4S systems in Gram-negative bacteria: a mechanistic biology perspective. *EMBO Rep* 10.15252/embr.201847012 (2019).
108. E. Boudaher, C. L. Shaffer, Inhibiting bacterial secretion systems in the fight against antibiotic resistance. *MedChemComm* **10**, 682-692 (2019).
109. M. Bauer *et al.*, The ALPK1/TIFA/NF-kappaB axis links a bacterial carcinogen to R-loop-induced replication stress. *Nat Commun* **11**, 5117 (2020).
110. W. Fischer, N. Tegtmeyer, K. Stingl, S. Backert, Four Chromosomal Type IV Secretion Systems in Helicobacter pylori: Composition, Structure and Function. *Front Microbiol* **11**, 1592 (2020).
111. G. G. Sgro *et al.*, Cryo-EM structure of the bacteria-killing type IV secretion system core complex from Xanthomonas citri. *Nat Microbiol* **3**, 1429-1440 (2018).

112. D. P. Souza *et al.*, Bacterial killing via a type IV secretion system. *Nat Commun* **6**, 6453 (2015).
113. D. L. Burns, Secretion of Pertussis Toxin from Bordetella pertussis. *Toxins (Basel)* **13** (2021).
114. M. M. Callaghan, J. H. Heilers, C. van der Does, J. P. Dillard, Secretion of Chromosomal DNA by the Neisseria gonorrhoeae Type IV Secretion System. *Curr Top Microbiol Immunol* **413**, 323-345 (2017).
115. H. L. Hamilton, N. M. Dominguez, K. J. Schwartz, K. T. Hackett, J. P. Dillard, Neisseria gonorrhoeae secretes chromosomal DNA via a novel type IV secretion system. *Mol Microbiol* **55**, 1704-1721 (2005).
116. D. A. Baltrus, M. J. Blaser, K. Guillemin, Helicobacter pylori Genome Plasticity. *Genome dynamics* **6**, 75-90 (2009).
117. E. Fernandez-Gonzalez, S. Backert, DNA transfer in the gastric pathogen Helicobacter pylori. *J Gastroenterol* **49**, 594-604 (2014).
118. D. A. Baltrus, K. Guillemin, P. C. Phillips, Natural transformation increases the rate of adaptation in the human pathogen Helicobacter pylori. *Evolution; international journal of organic evolution* **62**, 39-49 (2008).
119. F. de la Cruz, L. S. Frost, R. J. Meyer, E. L. Zechner, Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev* **34**, 18-40 (2010).
120. J. Lu *et al.*, Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol Microbiol* **70**, 89-99 (2008).
121. G. Moncalián, F. de la Cruz, DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. *Biochim Biophys Acta* **1701**, 15-23 (2004).
122. J. J. Wong, J. Lu, R. A. Edwards, L. S. Frost, J. N. Glover, Structural basis of cooperative DNA recognition by the plasmid conjugation factor, TraM. *Nucleic Acids Res* **39**, 6775-6788 (2011).
123. J. J. Wong, J. Lu, J. N. Glover, Relaxosome function and conjugation regulation in F-like plasmids - a structural biology perspective. *Mol Microbiol* **85**, 602-617 (2012).
124. E. L. Zechner, S. Lang, J. F. Schildbach, Assembly and mechanisms of bacterial type IV secretion machines. *Philos Trans R Soc Lond B Biol Sci* **367**, 1073-1087 (2012).
125. Y. G. Li, B. Hu, P. J. Christie, Biological and Structural Diversity of Type IV Secretion Systems. *Microbiol Spectr* **7** (2019).
126. M. P. Garcillan-Barcia, M. V. Francia, F. de la Cruz, The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev* **33**, 657-687 (2009).
127. J. Guglielmini, L. Quintais, M. P. Garcillán-Barcia, F. de la Cruz, E. P. Rocha, The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. *PLoS Genet* **7**, e1002222 (2011).
128. E. L. Zechner, G. Moncalián, F. de la Cruz, Relaxases and Plasmid Transfer in Gram-Negative Bacteria. *Curr Top Microbiol Immunol* **413**, 93-113 (2017).
129. S. Datta, C. Larkin, J. F. Schildbach, Structural insights into single-stranded DNA binding and cleavage by F factor TraI. *Structure* **11**, 1369-1379 (2003).

130. L. Dostál, J. F. Schildbach, Single-stranded DNA binding by F TraI relaxase and helicase domains is coordinately regulated. *J Bacteriol* **192**, 3620-3628 (2010).
131. S. W. Matson, H. Ragonese, The F-plasmid TraI protein contains three functional domains required for conjugative DNA strand transfer. *J Bacteriol* **187**, 697-706 (2005).
132. L. M. Guogas, S. A. Kennedy, J. H. Lee, M. R. Redinbo, A novel fold in the TraI relaxase-helicase c-terminal domain is essential for conjugative DNA transfer. *J Mol Biol* **386**, 554-568 (2009).
133. A. Ilangovan *et al.*, Cryo-EM Structure of a Relaxase Reveals the Molecular Basis of DNA Unwinding during Bacterial Conjugation. *Cell* **169**, 708-721.e712 (2017).
134. A. Redzej *et al.*, Structure of a translocation signal domain mediating conjugative transfer by type IV secretion systems. *Mol Microbiol* **89**, 324-333 (2013).
135. A. Alperi *et al.*, A translocation motif in relaxase TrwC specifically affects recruitment by its conjugative type IV secretion system. *J Bacteriol* **195**, 4999-5006 (2013).
136. B. Burton, D. Dubnau, Membrane-associated DNA transport machines. *Cold Spring Harb Perspect Biol* **2**, a000406 (2010).
137. F. X. Gomis-Rüth, M. Coll, Structure of TrwB, a gatekeeper in bacterial conjugation. *Int J Biochem Cell Biol* **33**, 839-843 (2001).
138. F. X. Gomis-Rüth *et al.*, The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase. *Nature* **409**, 637-641 (2001).
139. N. Whitaker *et al.*, The All-Alpha Domains of Coupling Proteins from the Agrobacterium tumefaciens VirB/VirD4 and Enterococcus faecalis pCF10-Encoded Type IV Secretion Systems Confer Specificity to Binding of Cognate DNA Substrates. *J Bacteriol* **197**, 2335-2349 (2015).
140. H. D. de Paz *et al.*, Functional dissection of the conjugative coupling protein TrwB. *J Bacteriol* **192**, 2655-2669 (2010).
141. M. Llosa, I. Alkorta, Coupling Proteins in Type IV Secretion. *Curr Top Microbiol Immunol* **413**, 143-168 (2017).
142. M. J. Kwak *et al.*, Architecture of the type IV coupling protein complex of Legionella pneumophila. *Nat Microbiol* **2**, 17114 (2017).
143. N. Whitaker *et al.*, Chimeric Coupling Proteins Mediate Transfer of Heterologous Type IV Effectors through the Escherichia coli pKM101-Encoded Conjugation Machine. *J Bacteriol* **198**, 2701-2718 (2016).
144. Y. G. Li, P. J. Christie, The TraK accessory factor activates substrate transfer through the pKM101 type IV secretion system independently of its role in relaxosome assembly. *Mol Microbiol* **114**, 214-229 (2020).
145. C. J. Gruber *et al.*, Conjugative DNA Transfer Is Enhanced by Plasmid R1 Partitioning Proteins. *Front Mol Biosci* **3**, 32 (2016).
146. K. Atmakuri, E. Cascales, O. T. Burton, L. M. Banta, P. J. Christie, *Agrobacterium ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions* (2007), vol. 26, pp. 2540-2551.
147. M. Trokter, G. Waksman, Translocation through the Conjugative Type IV Secretion System Requires Unfolding of Its Protein Substrate. *J Bacteriol* **200** (2018).

148. G. Grandoso *et al.*, Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. *J Mol Biol* **295**, 1163-1172 (2000).
149. C. E. César, C. Machón, F. de la Cruz, M. Llosa, A new domain of conjugative relaxase TrwC responsible for efficient oriT-specific recombination on minimal target sequences. *Mol Microbiol* **62**, 984-996 (2006).
150. R. Fronzes *et al.*, Structure of a type IV secretion system core complex. *Science* **323**, 266-268 (2009).
151. A. Rivera-Calzada *et al.*, Structure of a bacterial type IV secretion core complex at subnanometre resolution. *EMBO J* **32**, 1195-1204 (2013).
152. V. Chandran *et al.*, Structure of the outer membrane complex of a type IV secretion system. *Nature* **462**, 1011-1015 (2009).
153. H. H. Low *et al.*, Structure of a type IV secretion system. *Nature* **508**, 550-553 (2014).
154. K. Macé *et al.*, Cryo-EM structure of a type IV secretion system. *Nature* **607**, 191-196 (2022).
155. S. J. Jakubowski *et al.*, Agrobacterium VirB10 domain requirements for type IV secretion and T pilus biogenesis. *Molecular Microbiology* **71**, 779-794 (2009).
156. V. Chandran Darbari, G. Waksman, Structural Biology of Bacterial Type IV Secretion Systems. *Annu Rev Biochem* **84**, 603-629 (2015).
157. M. Clarke, L. Maddera, R. L. Harris, P. M. Silverman, F-pili dynamics by live-cell imaging. *Proc Natl Acad Sci U S A* **105**, 17978-17981 (2008).
158. B. Hu, P. Khara, P. J. Christie, Structural bases for F plasmid conjugation and F pilus biogenesis in Escherichia coli. *Proc Natl Acad Sci U S A* **116**, 14222-14227 (2019).
159. T. R. D. Costa *et al.*, Structure of the Bacterial Sex F Pilus Reveals an Assembly of a Stoichiometric Protein-Phospholipid Complex. *Cell* **166**, 1436-1444 e1410 (2016).
160. A. Babic, A. B. Lindner, M. Vulic, E. J. Stewart, M. Radman, Direct visualization of horizontal gene transfer. *Science* **319**, 1533-1536 (2008).
161. J. Amro *et al.*, Cryo-EM structure of the Agrobacterium tumefaciens T-pilus reveals the importance of positive charges in the lumen. *Structure* **31**, 375-384.e374 (2023).
162. S. Kreida *et al.*, Cryo-EM structure of the Agrobacterium tumefaciens T4SS-associated T-pilus reveals stoichiometric protein-phospholipid assembly. *Structure* **31**, 385-394.e384 (2023).
163. J. E. Gordon *et al.*, Use of chimeric type IV secretion systems to define contributions of outer membrane subassemblies for contact-dependent translocation. *Molecular Microbiology* **105**, 273-293 (2017).
164. E. Sagulenko, V. Sagulenko, J. Chen, P. J. Christie, Role of Agrobacterium VirB11 ATPase in T-pilus assembly and substrate selection. *J Bacteriol* **183**, 5813-5825 (2001).
165. J. Guglielmini, F. de la Cruz, E. P. Rocha, Evolution of conjugation and type IV secretion systems. *Mol Biol Evol* **30**, 315-331 (2013).
166. N. Goessweiner-Mohr, K. Arends, W. Keller, E. Grohmann, Conjugation in Gram-Positive Bacteria. *Microbiol Spectr* **2**, PLAS-0004-2013 (2014).

167. B. K. Kozlowicz *et al.*, Molecular basis for control of conjugation by bacterial pheromone and inhibitor peptides. *Mol Microbiol* **62**, 958-969 (2006).
168. M. V. Francia *et al.*, A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiol Rev* **28**, 79-100 (2004).
169. J. A. Laverde Gomez, M. Bhatty, P. J. Christie, PrgK, a multidomain peptidoglycan hydrolase, is essential for conjugative transfer of the pheromone-responsive plasmid pCF10. *J Bacteriol* **196**, 527-539 (2014).
170. E. Grohmann, G. Muth, M. Espinosa, Conjugative plasmid transfer in gram-positive bacteria. *Microbiol Mol Biol Rev* **67**, 277-301, table of contents (2003).
171. J. P. Dillard, H. S. Seifert, A variable genetic island specific for Neisseria gonorrhoeae is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. *Mol Microbiol* **41**, 263-277 (2001).
172. M. Zweig *et al.*, Secreted single-stranded DNA is involved in the initial phase of biofilm formation by Neisseria gonorrhoeae. *Environ Microbiol* **16**, 1040-1052 (2014).
173. H. L. Hamilton, J. P. Dillard, Natural transformation of Neisseria gonorrhoeae: from DNA donation to homologous recombination. *Mol Microbiol* **59**, 376-385 (2006).
174. P. Dutka *et al.*, Structure and Function of the Dot/Icm T4SS. *bioRxiv* 10.1101/2023.03.22.533729 (2023).
175. L. A. S. Snyder, S. A. Jarvis, N. J. Saunders, Complete and variant forms of the 'gonococcal genetic island' in Neisseria meningitidis. *Microbiology (Reading)* **151**, 4005-4013 (2005).
176. K. L. Woodhams, Z. L. Benet, S. E. Blonsky, K. T. Hackett, J. P. Dillard, Prevalence and detailed mapping of the gonococcal genetic island in Neisseria meningitidis. *J Bacteriol* **194**, 2275-2285 (2012).
177. W. Salgado-Pabon, S. Jain, N. Turner, C. van der Does, J. P. Dillard, A novel relaxase homologue is involved in chromosomal DNA processing for type IV secretion in Neisseria gonorrhoeae. *Mol Microbiol* **66**, 930-947 (2007).
178. W. Salgado-Pabón *et al.*, Increased expression of the type IV secretion system in piliated Neisseria gonorrhoeae variants. *J Bacteriol* **192**, 1912-1920 (2010).
179. H. L. Hamilton, K. J. Schwartz, J. P. Dillard, Insertion-duplication mutagenesis of neisseria: use in characterization of DNA transfer genes in the gonococcal genetic island. *J Bacteriol* **183**, 4718-4726 (2001).
180. E. Pachulec *et al.*, Functional analysis of the Gonococcal Genetic Island of Neisseria gonorrhoeae. *PLoS One* **9**, e109613 (2014).
181. B. Koch *et al.*, Protein interactions within and between two F-type type IV secretion systems. *Mol Microbiol* **114**, 823-838 (2020).
182. M. E. Ramsey *et al.*, TraK and TraB are conserved outer membrane proteins of the Neisseria gonorrhoeae Type IV secretion system and are expressed at low levels in wild-type cells. *J Bacteriol* **196**, 2954-2968 (2014).
183. M. E. Ramsey, K. L. Woodhams, J. P. Dillard, The Gonococcal Genetic Island and Type IV Secretion in the Pathogenic Neisseria. *Front Microbiol* **2**, 61 (2011).
184. W. Salgado-Pabón, S. Jain, N. Turner, C. van der Does, J. P. Dillard, A novel relaxase homologue is involved in chromosomal DNA processing for type IV secretion in Neisseria gonorrhoeae. *Molecular microbiology* **66**, 930-947 (2007).

185. J. H. Heilers *et al.*, DNA processing by the MOBH family relaxase TraI encoded within the gonococcal genetic island. *Nucleic Acids Res* **47**, 8136-8153 (2019).
186. L. Aravind, E. V. Koonin, The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends in biochemical sciences* **23**, 469-472 (1998).
187. C. Bignell, C. M. Thomas, The bacterial ParA-ParB partitioning proteins. *J Biotechnol* **91**, 1-34 (2001).
188. K. Atmakuri, E. Cascales, O. T. Burton, L. M. Banta, P. J. Christie, Agrobacterium ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions. *Embo j* **26**, 2540-2551 (2007).
189. M. M. Callaghan *et al.*, Expression, Localization, and Protein Interactions of the Partitioning Proteins in the Gonococcal Type IV Secretion System. *Front Microbiol* **12**, 784483 (2021).
190. P. L. Kohler, H. L. Hamilton, K. Cloud-Hansen, J. P. Dillard, AtlA functions as a peptidoglycan lytic transglycosylase in the Neisseria gonorrhoeae type IV secretion system. *J Bacteriol* **189**, 5421-5428 (2007).
191. P. L. Kohler *et al.*, Mating pair formation homologue TraG is a variable membrane protein essential for contact-independent type IV secretion of chromosomal DNA by Neisseria gonorrhoeae. *J Bacteriol* **195**, 1666-1679 (2013).
192. C. Johnston, B. Martin, G. Fichant, P. Polard, J.-P. Claverys, Bacterial transformation: distribution, shared mechanisms and divergent control. *Nature Reviews Microbiology* **12**, 181-196 (2014).
193. D. Dubnau, M. Blokesch, Mechanisms of DNA Uptake by Naturally Competent Bacteria. *Annu Rev Genet* **53**, 217-237 (2019).
194. R. Provvedi, I. Chen, D. Dubnau, NucA is required for DNA cleavage during transformation of Bacillus subtilis. *Mol Microbiol* **40**, 634-644 (2001).
195. A. Silale, S. M. Lea, B. C. Berks, The DNA transporter ComEC has metal-dependent nuclease activity that is important for natural transformation. *Mol Microbiol* **116**, 416-426 (2021).
196. S. Quevillon-Cheruel *et al.*, Structure-function analysis of pneumococcal DprA protein reveals that dimerization is crucial for loading RecA recombinase onto DNA during transformation. *Proc Natl Acad Sci U S A* **109**, E2466-2475 (2012).
197. P. P. Damke *et al.*, ComFC mediates transport and handling of single-stranded DNA during natural transformation. *Nat Commun* **13**, 1961 (2022).
198. Z. T. Pimentel, Y. Zhang, Evolution of the Natural Transformation Protein, ComEC, in Bacteria. *Front Microbiol* **9**, 2980 (2018).
199. C. K. Ellison *et al.*, Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in Vibrio cholerae. *Nat Microbiol* **3**, 773-780 (2018).
200. H. Gangel *et al.*, Concerted spatio-temporal dynamics of imported DNA and ComE DNA uptake protein during gonococcal transformation. *PLoS Pathog* **10**, e1004043 (2014).
201. P. Seitz, M. Blokesch, DNA-uptake machinery of naturally competent Vibrio cholerae. *Proc Natl Acad Sci U S A* **110**, 17987-17992 (2013).
202. N. Mirouze, C. Ferret, C. Cornilleau, R. Carballido-López, Antibiotic sensitivity reveals that wall teichoic acids mediate DNA binding during competence in Bacillus subtilis. *Nat Commun* **9**, 5072 (2018).

203. D. Hofreuter, S. Odenbreit, R. Haas, Natural transformation competence in Helicobacter pylori is mediated by the basic components of a type IV secretion system. *Mol Microbiol* **41**, 379-391 (2001).
204. A. Karnholz *et al.*, Functional and topological characterization of novel components of the comB DNA transformation competence system in Helicobacter pylori. *J Bacteriol* **188**, 882-893 (2006).
205. N. J. Krüger, K. Stingl, Two steps away from novelty--principles of bacterial DNA uptake. *Mol Microbiol* **80**, 860-867 (2011).
206. K. Stingl, S. Müller, G. Scheidgen-Kleyboldt, M. Clausen, B. Maier, Composite system mediates two-step DNA uptake into Helicobacter pylori. *Proc Natl Acad Sci U S A* **107**, 1184-1189 (2010).
207. P. P. Damke *et al.*, Identification of the periplasmic DNA receptor for natural transformation of Helicobacter pylori. *Nat Commun* **10**, 5357 (2019).
208. N. J. Krüger, M. T. Knüver, A. Zawilak-Pawlik, B. Appel, K. Stingl, Genetic Diversity as Consequence of a Microaerobic and Neutrophilic Lifestyle. *PLoS Pathog* **12**, e1005626 (2016).
209. C. Corbinais, A. Mathieu, T. Kortulewski, J. P. Radicella, S. Marsin, Following transforming DNA in Helicobacter pylori from uptake to expression. *Mol Microbiol* **101**, 1039-1053 (2016).
210. C. Hepp, B. Maier, Kinetics of DNA uptake during transformation provide evidence for a translocation ratchet mechanism. *Proc Natl Acad Sci U S A* **113**, 12467-12472 (2016).
211. T. J. Treangen, O. H. Ambur, T. Tonjum, E. P. C. Rocha, The impact of the neisserial DNA uptake sequences on genome evolution and stability. *Genome Biology* **9**, R60 (2008).
212. H. Kavermann *et al.*, Identification and characterization of Helicobacter pylori genes essential for gastric colonization. *The Journal of experimental medicine* **197**, 813-822 (2003).
213. M. S. Dorer, I. E. Cohen, T. H. Sessler, J. Fero, N. R. Salama, Natural competence promotes Helicobacter pylori chronic infection. *Infect Immun* **81**, 209-215 (2013).
214. J. Haase, R. Lurz, A. M. Grahn, D. H. Bamford, E. Lanka, Bacterial conjugation mediated by plasmid RP4: RSF1010 mobilization, donor-specific phage propagation, and pilus production require the same Tra2 core components of a proposed DNA transport complex. *J Bacteriol* **177**, 4779-4791 (1995).
215. F. I. R. M. Zoolkefli *et al.*, Isolation and Analysis of Donor Chromosomal Genes Whose Deficiency Is Responsible for Accelerating Bacterial and Trans-Kingdom Conjugations by IncP1 T4SS Machinery. *Frontiers in Microbiology* **12** (2021).
216. J. A. Dodsworth *et al.*, Interdomain conjugal transfer of DNA from bacteria to archaea. *Appl Environ Microbiol* **76**, 5644-5647 (2010).
217. K. Moriguchi *et al.*, Transkingdom genetic transfer from Escherichia coli to Saccharomyces cerevisiae as a simple gene introduction tool. *Appl Environ Microbiol* **79**, 4393-4400 (2013).
218. B. J. Karas *et al.*, Designer diatom episomes delivered by bacterial conjugation. *Nat Commun* **6**, 6925 (2015).

219. G. L. Regnard, R. P. Halley-Stott, F. L. Tanzer, Hitzeroth, II, E. P. Rybicki, High level protein expression in plants through the use of a novel autonomously replicating geminivirus shuttle vector. *Plant Biotechnol J* **8**, 38-46 (2010).
220. V. L. Waters, Conjugation between bacterial and mammalian cells. *Nat Genet* **29**, 375-376 (2001).
221. K. Moriguchi, S. Yamamoto, Y. Ohmine, K. Suzuki, A Fast and Practical Yeast Transformation Method Mediated by Escherichia coli Based on a Trans-Kingdom Conjugal Transfer System: Just Mix Two Cultures and Wait One Hour. *PLoS One* **11**, e0148989 (2016).
222. J. A. Heinemann, G. F. Sprague, Jr., Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* **340**, 205-209 (1989).
223. K. Inomata, M. Nishikawa, K. Yoshida, The yeast Saccharomyces kluyveri as a recipient eukaryote in transkingdom conjugation: behavior of transmitted plasmids in transconjugants. *J Bacteriol* **176**, 4770-4773 (1994).
224. G. T. Hayman, P. L. Bolen, Movement of shuttle plasmids from Escherichia coli into yeasts other than Saccharomyces cerevisiae using trans-kingdom conjugation. *Plasmid* **30**, 251-257 (1993).
225. M. J. de Groot, P. Bundock, P. J. Hooykaas, A. G. Beijersbergen, Agrobacterium tumefaciens-mediated transformation of filamentous fungi. *Nat Biotechnol* **16**, 839-842 (1998).
226. C. B. Michielse, P. J. Hooykaas, C. A. van den Hondel, A. F. Ram, Agrobacterium-mediated transformation as a tool for functional genomics in fungi. *Curr Genet* **48**, 1-17 (2005).
227. E. Machado-Ferreira *et al.*, Transgene expression in tick cells using Agrobacterium tumefaciens. *Exp Appl Acarol* **67**, 269-287 (2015).
228. T. Kunik *et al.*, Genetic transformation of HeLa cells by Agrobacterium. *Proc Natl Acad Sci U S A* **98**, 1871-1876 (2001).
229. F. Jasper, C. Koncz, J. Schell, H. H. Steinbiss, Agrobacterium T-strand production in vitro: sequence-specific cleavage and 5' protection of single-stranded DNA templates by purified VirD2 protein. *Proc Natl Acad Sci U S A* **91**, 694-698 (1994).
230. W. Pansegrau, F. Schoumacher, B. Hohn, E. Lanka, Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of Agrobacterium tumefaciens Ti plasmids: analogy to bacterial conjugation. *Proc Natl Acad Sci U S A* **90**, 11538-11542 (1993).
231. K. Atmakuri, E. Cascales, P. J. Christie, Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. *Mol Microbiol* **54**, 1199-1211 (2004).
232. K. Atmakuri, Z. Ding, P. J. Christie, VirE2, a Type IV secretion substrate, interacts with the VirD4 transfer protein at cell poles of Agrobacterium tumefaciens. *Molecular Microbiology* **49**, 1699-1713 (2003).
233. V. Citovsky *et al.*, Protein interactions involved in nuclear import of the Agrobacterium VirE2 protein in vivo and in vitro. *J Biol Chem* **279**, 29528-29533 (2004).
234. B. Lacroix, M. Vaidya, T. Tzfira, V. Citovsky, The VirE3 protein of Agrobacterium mimics a host cell function required for plant genetic transformation. *Embo j* **24**, 428-437 (2005).

235. B. Lacroix, V. Citovsky, A Functional Bacterium-to-Plant DNA Transfer Machinery of Rhizobium etli. *PLoS Pathog* **12**, e1005502 (2016).
236. G. Schroder, R. Schuelein, M. Quebatte, C. Dehio, Conjugative DNA transfer into human cells by the VirB/VirD4 type IV secretion system of the bacterial pathogen Bartonella henselae. *Proc Natl Acad Sci U S A* **108**, 14643-14648 (2011).
237. D. L. Guzmán-Herrador *et al.*, DNA Delivery and Genomic Integration into Mammalian Target Cells through Type IV A and B Secretion Systems of Human Pathogens. *Front Microbiol* **8**, 1503 (2017).
238. C. González-Prieto, R. Gabriel, C. Dehio, M. Schmidt, M. Llosa, The Conjugative Relaxase TrwC Promotes Integration of Foreign DNA in the Human Genome. *Appl Environ Microbiol* **83** (2017).
239. H. Nagai, T. Kubori, Type IVB Secretion Systems of Legionella and Other Gram-Negative Bacteria. *Front Microbiol* **2**, 136 (2011).
240. J. P. Vogel, H. L. Andrews, S. K. Wong, R. R. Isberg, Conjugative transfer by the virulence system of Legionella pneumophila. *Science* **279**, 873-876 (1998).
241. E. Fernandez-Gonzalez *et al.*, Transfer of R388 derivatives by a pathogenesis-associated type IV secretion system into both bacteria and human cells. *J Bacteriol* **193**, 6257-6265 (2011).
242. R. Rad *et al.*, Extracellular and intracellular pattern recognition receptors cooperate in the recognition of Helicobacter pylori. *Gastroenterology* **136**, 2247-2257 (2009).
243. S. D. R. Dooyema *et al.*, Helicobacter pylori actively suppresses innate immune nucleic acid receptors. *Gut Microbes* **14**, 2105102 (2022).
244. Y. W. Chang, C. L. Shaffer, L. A. Rettberg, D. Ghosal, G. J. Jensen, In Vivo Structures of the Helicobacter pylori cag Type IV Secretion System. *Cell Rep* **23**, 673-681 (2018).
245. D. Ghosal, Y. W. Chang, K. C. Jeong, J. P. Vogel, G. J. Jensen, In situ structure of the Legionella Dot/Icm type IV secretion system by electron cryotomography. *EMBO Rep* **18**, 726-732 (2017).
246. D. Ghosal *et al.*, Molecular architecture, polar targeting and biogenesis of the Legionella Dot/Icm T4SS. *Nat Microbiol* **4**, 1173-1182 (2019).
247. K. C. Jeong, D. Ghosal, Y. W. Chang, G. J. Jensen, J. P. Vogel, Polar delivery of Legionella type IV secretion system substrates is essential for virulence. *Proc Natl Acad Sci U S A* 10.1073/pnas.1621438114 (2017).
248. D. Chetrit, B. Hu, P. J. Christie, C. R. Roy, J. Liu, A unique cytoplasmic ATPase complex defines the Legionella pneumophila type IV secretion channel. *Nat Microbiol* **3**, 678-686 (2018).
249. K. Mace *et al.*, Cryo-EM structure of a type IV secretion system. *Nature* 10.1038/s41586-022-04859-y (2022).
250. B. Hu *et al.*, In Situ Molecular Architecture of the Helicobacter pylori Cag Type IV Secretion System. *mBio* **10** (2019).
251. A. S. Lin *et al.*, Bacterial Energetic Requirements for Helicobacter pylori Cag Type IV Secretion System-Dependent Alterations in Gastric Epithelial Cells. *Infect Immun* **88** (2020).
252. A. Redzej *et al.*, Structure of a VirD4 coupling protein bound to a VirB type IV secretion machinery. *EMBO J* 10.15252/embj.201796629 (2017).

253. M. J. Sheedlo, M. D. Ohi, D. B. Lacy, T. L. Cover, Molecular architecture of bacterial type IV secretion systems. *PLOS Pathogens* **18**, e1010720 (2022).
254. P. Dutka *et al.*, Structure and Function of the Dot/Icm T4SS. *bioRxiv* 10.1101/2023.03.22.533729, 2023.2003.2022.533729 (2023).
255. M. J. Sheedlo *et al.*, Cryo-EM reveals new species-specific proteins and symmetry elements in the Legionella pneumophila Dot/Icm T4SS. *eLife* **10**, e70427 (2021).
256. C. L. Durie *et al.*, Structural analysis of the Legionella pneumophila Dot/Icm type IV secretion system core complex. *Elife* **9** (2020).
257. J. M. Chung *et al.*, Structure of the Helicobacter pylori Cag type IV secretion system. *Elife* **8** (2019).
258. M. J. Sheedlo *et al.*, Cryo-EM reveals species-specific components within the Helicobacter pylori Cag type IV secretion system core complex. *Elife* **9** (2020).
259. S. Kutter *et al.*, Protein subassemblies of the Helicobacter pylori Cag type IV secretion system revealed by localization and interaction studies. *J Bacteriol* **190**, 2161-2171 (2008).
260. T. L. Cover, D. B. Lacy, M. D. Ohi, The Helicobacter pylori Cag Type IV Secretion System. *Trends Microbiol* **28**, 682-695 (2020).
261. V. J. Busler *et al.*, Protein-protein interactions among Helicobacter pylori cag proteins. *J Bacteriol* **188**, 4787-4800 (2006).
262. J. M. Chung *et al.*, Structure of the Helicobacter pylori Cag type IV secretion system. *Elife* **8**, e47644 (2019).
263. A. E. Frick-Cheng *et al.*, Molecular and Structural Analysis of the Helicobacter pylori cag Type IV Secretion System Core Complex. *mBio* **7**, e02001-02015 (2016).
264. M. J. Sheedlo *et al.*, Cryo-EM reveals species-specific components within the Helicobacter pylori Cag type IV secretion system core complex. *Elife* **9**, e59495 (2020).
265. B. Hu *et al.*, *In Situ* Molecular Architecture of the Helicobacter pylori Cag Type IV Secretion System. *mBio* **10**, e00849-00819 (2019).
266. W. Fischer, Assembly and molecular mode of action of the Helicobacter pylori Cag type IV secretion apparatus. *FEBS J* **278**, 1203-1212 (2011).
267. T. Kubori *et al.*, Native structure of a type IV secretion system core complex essential for Legionella pathogenesis. *Proc Natl Acad Sci U S A* **111**, 11804-11809 (2014).
268. H. Amin, A. Ilangovan, T. R. D. Costa, Architecture of the outer-membrane core complex from a conjugative type IV secretion system. *Nature Communications* **12**, 6834 (2021).
269. X. Liu, P. Khara, M. L. Baker, P. J. Christie, B. Hu, Structure of a type IV secretion system core complex encoded by multi-drug resistance F plasmids. *Nature Communications* **13**, 379 (2022).
270. S. C. Stein *et al.*, Helicobacter pylori modulates host cell responses by CagT4SS-dependent translocation of an intermediate metabolite of LPS inner core heptose biosynthesis. *PLOS Pathogens* **13**, e1006514 (2017).
271. G. Karimova, E. Gauliard, M. Davi, S. P. Ouellette, D. Ladant, Protein-Protein Interaction: Bacterial Two-Hybrid. *Methods Mol Biol* **1615**, 159-176 (2017).

272. C. L. Shaffer *et al.*, Helicobacter pylori Exploits a Unique Repertoire of Type IV Secretion System Components for Pilus Assembly at the Bacteria-Host Cell Interface. *PLOS Pathogens* **7**, e1002237 (2011).
273. E. M. Johnson, J. A. Gaddy, B. J. Voss, E. E. Hennig, T. L. Cover, Genes required for assembly of pili associated with the *Helicobacter pylori cag* type IV secretion system. *Infect Immun* 10.1128/iai.01640-14 (2014).
274. A. E. Frick-Cheng *et al.*, Molecular and Structural Analysis of the Helicobacter pylori cag Type IV Secretion System Core Complex. *MBio* **7** (2016).
275. C. L. Shaffer *et al.*, Helicobacter pylori exploits a unique repertoire of type IV secretion system components for pilus assembly at the bacteria-host cell interface. *PLoS Pathog* **7**, e1002237 (2011).
276. M. G. Varga *et al.*, Pathogenic Helicobacter pylori strains translocate DNA and activate TLR9 via the cancer-associated cag type IV secretion system. *Oncogene* **35**, 6262-6269 (2016).
277. M. G. Varga *et al.*, Immunostimulatory membrane proteins potentiate H. pylori-induced carcinogenesis by enabling CagA translocation. *Gut Microbes* **13**, 1-13 (2021).
278. L. M. Banta *et al.*, An Agrobacterium VirB10 mutation conferring a type IV secretion system gating defect. *J Bacteriol* **193**, 2566-2574 (2011).
279. J. Schindelin *et al.*, Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).
280. S. L. Benoit, S. Agudelo, R. J. Maier, A two-hybrid system reveals previously uncharacterized protein–protein interactions within the Helicobacter pylori NIF iron–sulfur maturation system. *Scientific Reports* **11**, 10794 (2021).
281. S. L. Benoit, A. A. Holland, M. K. Johnson, R. J. Maier, Iron-sulfur protein maturation in Helicobacter pylori: identifying a Nfu-type cluster carrier protein and its iron-sulfur protein targets. *Molecular microbiology* **108**, 379-396 (2018).
282. P. P. Damke *et al.*, ComFC mediates transport and handling of single-stranded DNA during natural transformation. *Nature Communications* **13**, 1961 (2022).
283. M. Mirdita *et al.*, ColabFold - Making protein folding accessible to all. *bioRxiv* 10.1101/2021.08.15.456425, 2021.2008.2015.456425 (2021).
284. E. F. Pettersen *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70-82 (2021).
285. N. Murata-Kamiya, K. Kikuchi, T. Hayashi, H. Higashi, M. Hatakeyama, Helicobacter pylori exploits host membrane phosphatidylserine for delivery, localization, and pathophysiological action of the CagA oncoprotein. *Cell host & microbe* **7**, 399-411 (2010).
286. E. Volkan *et al.*, Molecular basis of usher pore gating in Escherichia coli pilus biogenesis. *Proc Natl Acad Sci U S A* **110**, 20741-20746 (2013).
287. J. C. Mao, M. Putterman, Accumulation in gram-postive and gram-negative bacteria as a mechanism of resistance to erythromycin. *J Bacteriol* **95**, 1111-1117 (1968).
288. J. Andrzejewska *et al.*, Characterization of the pilin ortholog of the Helicobacter pylori type IV cag pathogenicity apparatus, a surface-associated protein expressed during infection. *J Bacteriol* **188**, 5865-5877 (2006).

289. M. Varadi *et al.*, AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research* **50**, D439-D444 (2021).
290. T. R. D. Costa *et al.*, Structure of the Bacterial Sex F Pilus Reveals an Assembly of a Stoichiometric Protein-Phospholipid Complex. *Cell* **166**, 1436-1444.e1410 (2016).
291. R. Eisenbrandt *et al.*, Conjugative pili of IncP plasmids, and the Ti plasmid T pilus are composed of cyclic subunits. *J Biol Chem* **274**, 22548-22555 (1999).
292. D. Sivanesan, M. A. Hancock, A. M. Villamil Giraldo, C. Baron, Quantitative analysis of VirB8-VirB9-VirB10 interactions provides a dynamic model of type IV secretion system core complex assembly. *Biochemistry* **49**, 4483-4493 (2010).
293. J. E. Gunton, M. W. Gilmour, G. Alonso, D. E. Taylor, Subcellular localization and functional domains of the coupling protein, TraG, from IncHI1 plasmid R27. *Microbiology* **151**, 3549-3561 (2005).
294. I. Pattis, E. Weiss, R. Laugks, R. Haas, W. Fischer, The Helicobacter pylori CagF protein is a type IV secretion chaperone-like molecule that binds close to the C-terminal secretion signal of the CagA effector protein. *Microbiology (Reading)* **153**, 2896-2909 (2007).
295. R. M. Barrozo *et al.*, Functional plasticity in the type IV secretion system of Helicobacter pylori. *PLoS Pathog* **9**, e1003189 (2013).
296. T. Kwok *et al.*, Helicobacter exploits integrin for type IV secretion and kinase activation. *Nature* **449**, 862-866 (2007).
297. B. Hu, M. Lara-Tejero, Q. Kong, J. E. Galan, J. Liu, In Situ Molecular Architecture of the Salmonella Type III Secretion Machine. *Cell* **168**, 1065-1074 e1010 (2017).
298. S. Barden *et al.*, A helical RGD motif promoting cell adhesion: crystal structures of the Helicobacter pylori type IV secretion system pilus protein CagL. *Structure* **21**, 1931-1941 (2013).
299. S. Barden *et al.*, Structure of a three-dimensional domain-swapped dimer of the Helicobacter pylori type IV secretion system pilus protein CagL. *Acta crystallographica. Section D, Biological crystallography* **70**, 1391-1400 (2014).
300. S. R. Dix *et al.*, Structural insights into the function of type VI secretion system TssA subunits. *Nature Communications* **9**, 4765 (2018).
301. A. A. Chernyatina, H. H. Low, Core architecture of a bacterial type II secretion system. *Nat Commun* **10**, 5437 (2019).
302. J. Hu *et al.*, Cryo-EM analysis of the T3S injectisome reveals the structure of the needle and open secretin. *Nat Commun* **9**, 3840 (2018).
303. S. Johnson *et al.*, Symmetry mismatch in the MS-ring of the bacterial flagellar rotor explains the structural coordination of secretion and rotation. *Nat Microbiol* **5**, 966-975 (2020).
304. T. R. D. Costa *et al.*, Type IV secretion systems: Advances in structure, function, and activation. *Molecular Microbiology* **115**, 436-452 (2021).
305. Y. G. Li, B. Hu, P. J. Christie, Biological and Structural Diversity of Type IV Secretion Systems. *Microbiol Spectr* **7**, 7.2.30 (2019).
306. G. Waksman, From conjugation to T4S systems in Gram-negative bacteria: a mechanistic biology perspective. *EMBO Rep* **20** (2019).

307. J. I. Grove, M. N. Alandiyjany, R. M. Delahay, Site-specific relaxase activity of a VirD2-like protein encoded within the tfs4 genomic island of Helicobacter pylori. *The Journal of biological chemistry* **288**, 26385-26396 (2013).
308. J. I. Grove, M. N. Alandiyjany, R. M. Delahay, Site-specific Relaxase Activity of a VirD2-like Protein Encoded within the tfs4 Genomic Island of Helicobacter pylori. *Journal of Biological Chemistry* **288**, 26385-26396 (2013).
309. S. J. Jakubowski, E. Cascales, V. Krishnamoorthy, P. J. Christie, Agrobacterium tumefaciens VirB9, an outer-membrane-associated component of a type IV secretion system, regulates substrate selection and T-pilus biogenesis. *J Bacteriol* **187**, 3486-3495 (2005).
310. I. Garza, P. J. Christie, A putative transmembrane leucine zipper of agrobacterium VirB10 is essential for t-pilus biogenesis but not type IV secretion. *J Bacteriol* **195**, 3022-3034 (2013).
311. H. W. Rhee *et al.*, Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* **339**, 1328-1331 (2013).
312. J. D. Martell *et al.*, Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat Biotechnol* **30**, 1143-1148 (2012).
313. S. S. Lam *et al.*, Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat Methods* **12**, 51-54 (2015).
314. L. Trinkle-Mulcahy, Recent advances in proximity-based labeling methods for interactome mapping. *F1000Res* **8**, F1000 Faculty Rev-1135 (2019).
315. B. Singer-Krüger *et al.*, APEX2-mediated proximity labeling resolves protein networks in Saccharomyces cerevisiae cells. *Febs j* **287**, 325-344 (2020).
316. Y. G. Santin *et al.*, In vivo TssA proximity labelling during type VI secretion biogenesis reveals TagA as a protein that stops and holds the sheath. *Nat Microbiol* **3**, 1304-1313 (2018).
317. E. A. Rucks, M. G. Olson, L. M. Jorgenson, R. R. Srinivasan, S. P. Ouellette, Development of a Proximity Labeling System to Map the Chlamydia trachomatis Inclusion Membrane. *Frontiers in Cellular and Infection Microbiology* **7** (2017).
318. E. A. Rucks, M. G. Olson, L. M. Jorgenson, R. R. Srinivasan, S. P. Ouellette, Development of a Proximity Labeling System to Map the Chlamydia trachomatis Inclusion Membrane. *Front Cell Infect Microbiol* **7**, 40 (2017).
319. K. F. Cho *et al.*, Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nat Protoc* **15**, 3971-3999 (2020).
320. A. Jurik *et al.*, The coupling protein Cagbeta and its interaction partner CagZ are required for type IV secretion of the Helicobacter pylori CagA protein. *Infection and immunity* **78**, 5244-5251 (2010).
321. X. Wu *et al.*, Mechanism of regulation of the Helicobacter pylori Cagβ ATPase by CagZ. *Nature Communications* **14**, 479 (2023).
322. I. Pattis, E. Weiss, R. Laugks, R. Haas, W. Fischer, The Helicobacter pylori CagF protein is a type IV secretion chaperone-like molecule that binds close to the C-terminal secretion signal of the CagA effector protein. *Microbiology* **153**, 2896-2909 (2007).
323. M. R. Couturier, E. Tasca, C. Montecucco, M. Stein, Interaction with CagF is required for translocation of CagA into the host via the Helicobacter pylori type IV secretion system. *Infect Immun* **74**, 273-281 (2006).

324. S. S. Lam *et al.*, Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat Methods* **12**, 51-54 (2015).
325. J. Andrzejewska *et al.*, Characterization of the Pilin Ortholog of the *Helicobacter pylori* Type IV *cag* Pathogenicity Apparatus, a Surface-Associated Protein Expressed during Infection. *Journal of Bacteriology* **188**, 5865-5877 (2006).
326. S. Makishima, K. Komoriya, S. Yamaguchi, S. I. Aizawa, Length of the flagellar hook and the capacity of the type III export apparatus. *Science* **291**, 2411-2413 (2001).
327. K. Muramoto, S. Makishima, S.-I. Aizawa, R. M. Macnab, Effect of Hook Subunit Concentration on Assembly and Control of Length of the Flagellar Hook of *Salmonella*. *Journal of Bacteriology* **181**, 5808-5813 (1999).
328. H. J. Yeo, S. N. Savvides, A. B. Herr, E. Lanka, G. Waksman, Crystal structure of the hexameric traffic ATPase of the Helicobacter pylori type IV secretion system. *Mol Cell* **6**, 1461-1472 (2000).
329. V. Dumrongprechachan *et al.*, Cell-type and subcellular compartment-specific APEX2 proximity labeling reveals activity-dependent nuclear proteome dynamics in the striatum. *Nature Communications* **12**, 4855 (2021).
330. M. Perez Verdaguer *et al.*, Time-resolved proximity labeling of protein networks associated with ligand-activated EGFR. *Cell Rep* **39**, 110950 (2022).
331. M. G. Olson *et al.*, Proximity Labeling To Map Host-Pathogen Interactions at the Membrane of a Bacterium-Containing Vacuole in Chlamydia trachomatis-Infected Human Cells. *Infection and Immunity* **87**, e00537-00519 (2019).
332. M. G. Olson, L. M. Jorgenson, R. E. Widner, E. A. Rucks, Proximity Labeling of the Chlamydia trachomatis Inclusion Membrane. *Methods Mol Biol* **2042**, 245-278 (2019).
333. G. Keb, J. Ferrell, K. R. Scanlon, T. J. Jewett, K. A. Fields, Chlamydia trachomatis TmeA Directly Activates N-WASP To Promote Actin Polymerization and Functions Synergistically with TarP during Invasion. *mBio* **12**, e02861-02820 (2021).
334. M. S. Dickinson *et al.*, Proximity-dependent proteomics of the Chlamydia trachomatis inclusion membrane reveals functional interactions with endoplasmic reticulum exit sites. *PLOS Pathogens* **15**, e1007698 (2019).
335. U. S. Ganapathy *et al.*, Compartment-Specific Labeling of Bacterial Periplasmic Proteins by Peroxidase-Mediated Biotinylation. *ACS Infect Dis* **4**, 918-925 (2018).
336. S. Backert, M. Selbach, Role of type IV secretion in Helicobacter pylori pathogenesis. *Cellular Microbiology* **10**, 1573-1581 (2008).
337. E. Grohmann, P. J. Christie, G. Waksman, S. Backert, Type IV secretion in Gram-negative and Gram-positive bacteria. *Molecular Microbiology* **107**, 455-471 (2018).
338. E. Cascales, K. Atmakuri, M. K. Sarkar, P. J. Christie, DNA substrate-induced activation of the Agrobacterium VirB/VirD4 type IV secretion system. *J Bacteriol* **195**, 2691-2704 (2013).
339. S. P. Hancock, D. Cascio, R. C. Johnson, Cooperative DNA binding by proteins through DNA shape complementarity. *Nucleic Acids Research* **47**, 8874-8887 (2019).

340. D. Panne, T. Maniatis, S. C. Harrison, An atomic model of the interferon-beta enhanceosome. *Cell* **129**, 1111-1123 (2007).
341. Y. M. Chang, C. K. Chen, M. H. Hou, Conformational changes in DNA upon ligand binding monitored by circular dichroism. *Int J Mol Sci* **13**, 3394-3413 (2012).
342. S. J. LeBlanc *et al.*, Coordinated protein and DNA conformational changes govern mismatch repair initiation by MutS. *Nucleic Acids Research* **46**, 10782-10795 (2018).
343. L. Schärfen, M. Schlierf, Real-time monitoring of protein-induced DNA conformational changes using single-molecule FRET. *Methods* **169**, 11-20 (2019).
344. L. L. Kirkemo *et al.*, Cell-surface tethered promiscuous biotinylators enable comparative small-scale surface proteomic analysis of human extracellular vesicles and cells. *eLife* **11**, e73982 (2022).
345. J. A. Taylor *et al.*, Distinct cytoskeletal proteins define zones of enhanced cell wall synthesis in Helicobacter pylori. *eLife* **9**, e52482 (2020).

VITA

Mackenzie Ryan

Education

2018 – 2023: University of Kentucky – Lexington, KY

2014 – 2018: University of Dayton – Dayton, OH

Bachelor of Science, Biology

Professional Positions

2018 – 2023: Research Assistant, Shaffer Lab.

University of Kentucky – Lexington, KY

2017: Plant Lab Intern, Center for Conservations and Research of Endangered Wildlife.

Cincinnati Zoo – Cincinnati, OH

2016 – 2018: Research Assistant, Prather Lab

University of Dayton – Dayton, OH

Publications

Varga, M. G., Wood, C. R., Butt, J., **Ryan, M. E.**, You, W. C., Pan, K., Waterboer, T., Epplein, M., & Shaffer, C. L. (2021). Immunostimulatory membrane proteins potentiate *H. pylori*-induced carcinogenesis by enabling CagA translocation. *Gut Microbes*, 13(1), 1–13.

Ryan, M. E., Damke P. P., Shaffer, C. L. (2023). DNA transport through the dynamic type IV secretion system. *Infection and Immunity*, 91, e00436-00422.

Damke, P. P., Wood, C. R., Wood, **Ryan, M. E.**, Bryant, C., Fischer, W., Haas, R., Shaffer, C. L. *Helicobacter pylori* type IV secretion system crosstalk enables trans-kingdom DNA conjugation. (In revision)

Ryan, M. E., Damke, P. P., Bryant, C., Sheedlo, M. J. & Shaffer, C. L. (2023) Architectural asymmetry enables DNA transport through the *Helicobacter pylori cag* type IV secretion system. (In revision)