Exploring the Type III Secretion System in *Shigella flexneri*: from Biotin Ligase Optimization to IpgC Proxisome Profiling

Nathaline Haidar Ahmad

A Thesis Submitted to the University of Ottawa

In Partial Fulfillment of the Requirements for the

Doctorate in Philosophy degree in Chemistry
With a Concentration in Biomolecular Sciences

Department of Chemistry and Biomolecular Sciences

Faculty of Sciences

University of Ottawa

Supervisor:	François-Xavier Campbell-Valois, PhD, Chemistry and Biomolecular Sciences, uOttawa
External Examiner:	David Cisneros, PhD, The Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast
Examining Board:	Christopher N. Boddy, PhD, Chemistry and Biomolecular Sciences, uOttawa
	Corrie daCosta, PhD, Chemistry and Biomolecular Sciences, uOttawa
	Natalie Goto, PhD, Chemistry and Biomolecular Sciences, uOttawa

December 2025

Abstract

Shigella flexneri remains a major global health concern, causing tens of millions of infections and hundreds of thousands of deaths annually. Central to the pathogenesis of *S. flexneri* is the Type III Secretion System (T3SS), a nanomachine that translocates effector proteins into host cells to invade them. The T3SS undergoes major structural and functional rearrangements upon activation, yet many of the protein interactions and changes associated with this system remain poorly defined due to their transient and dynamic nature. To address this, we adapted the biotin ligase TurboID proximity labeling system for use in *S. flexneri*, optimizing it to function efficiently in the bacterial cytoplasm. TurboID exhibited rapid biotinylation kinetics and minimal cross-reactivity with native BirA binding partners, making it well suited for mapping protein proxisomes and interactomes in *Shigella*. We applied this system to two key proteins of the T3SS transcriptional cascade: IpgC, a dual-function chaperone and co-activator, and MxiE, the AraC-family transcription factor that controls late effector gene expression such as the IpaHs. Using IpgC and MxiE as baits, we profiled their proxisomes in both T3SS on and off-states. Our results confirmed established MxiE-IpgC interactions and identified novel candidate partners, including transcription factors and metabolic enzymes, suggesting a broader regulatory network. Notably, we found that IpgC's proxisome extends beyond the T3SS, pointing to unexplored roles. For example, our data also suggests that IpgC may play a role in modulating acid stress responses.

These results provide new insights into the regulatory complexity of the T3SS and serve as a platform for uncovering novel protein interactions, paving the way for future studies on T3SS regulation and antivirulence therapies. Our findings also underscore the utility of biotin ligases for studying protein complexes in *Shigella* and other bacterial models.

Résumé

Shigella flexneri demeure une menace mondiale pour la santé publique, qui est responsable de dizaines de millions d'infections et de centaines de milliers de décès chaque année. Au cœur de la pathogénèse de *S. flexneri* se trouve le système de sécrétion de type III (T3SS), une nanomachine qui injecte des substrats dans les cellules hôtes afin de les envahir. Le T3SS subit d'importantes réorganisations structurales et fonctionnelles lors de son activation, mais de nombreuses interactions protéiques et changements associées à ce système restent mal caractérisées en raison de leur nature transitoire et dynamique. Pour surmonter ces limitations, nous avons adapté le système d'étiquetage de proximité TurboID pour l'utiliser dans *S. flexneri*, en l'optimisant pour qu'il fonctionne efficacement dans le cytoplasme bactérien. TurboID a démontré une cinétique de biotinylation rapide et une réactivité réduite avec les partenaires de liaison natifs de BirA, le rendant particulièrement adapté au profilage des proxisomes et des interactomes protéiques chez *Shigella*. Nous avons appliqué ce système à deux protéines clés de la cascade transcriptionnelle du T3SS: IpgC, une chaperonne et co-activateur à double fonction, et MxiE, le facteur de transcription de la famille AraC qui contrôle l'expression des effecteurs tardifs du T3SS tel que les IpaHs. En utilisant IpgC et MxiE comme cibles, nous avons profilé leurs proxisomes dans les états actif et inactif du T3SS. Nos résultats ont confirmé les interactions connues de MxiE-IpgC et ont identifié de nouveaux partenaires candidats, incluant des facteurs de transcription et des enzymes métaboliques, suggérant l'existence d'un réseau régulateur plus étendu. Notamment, nous avons observé que le proxisome d'IpgC s'étend au-delà du T3SS, ce qui pourrait indiquer des fonctions encore inexplorées. Par exemple, nos données suggèrent qu'IpgC pourrait jouer un rôle dans la modulation de la réponse au stress acide.

Ces résultats offrent de nouvelles perspectives sur la complexité de la régulation du T3SS et constituent une base pour l'exploration de nouvelles interactions protéiques, ouvrant la voie à des études futures sur la régulation du T3SS et le développement de thérapies antivirulence. Nos conclusions soulignent également l'utilité des ligases à biotine pour l'étude des complexes protéiques chez *Shigella* et dans d'autres modèles bactériens.

Dedication

To my parents, whose love, sacrifices and constant support inspire me to aim high.
My achievements are as much yours as they are mine, and I hope they bring you pride and joy.

Acknowledgements

I am deeply grateful to my supervisor, François-Xavier Campbell-Valois, for his patience, guidance, and steady support over the years. His mentorship has made a big difference in both my academic journey and personal growth.

I am also thankful to my Thesis Committee Members, Christopher Boddy, Corrie daCosta, and Natalie Goto, for their encouragement and for generously sharing their time and expertise throughout this process.

Shout-out to my colleagues at the Host-Microbes Interactions group for their scientific insights and for being such a supportive presence throughout this experience. Special thanks go to Kyle Tomaro and France Manigat for their direct contributions to my work and for offering thoughtful perspectives.

I would also like to acknowledge the Technomise-Create program for their generous scholarship support, which helped me to pursue and complete this research.

And finally, to my family, thank you for always being there cheering me on. I couldn't have done this without you.

Table of Contents

List of Figures

List of Tables

Supplemental tables

List of Abbreviations

AR: acid resistance

BccP: biotin carboxyl carrier protein

BID: BioID

BID2: BioID2

BPL (or BL): biotin protein ligase

DTT: dithiothreitol

EMSA: electrophoretic mobility shift assay

EV: empty vector

FC: fold change

FDR: false discovery rate

GDAR (GAD): glutamate-dependent acid resistance

GFP: green fluorescent protein

HRP: horseradish peroxidase

IS: insertion sequence

LFQ: label-free quantification

miT: miniTurbo

pINV: invasion plasmid

pWR100: the virulence plasmid of *Shigella flexneri* strain M90T

ROI: region of interest

SAINT: Significance Analysis of INTeractome

SDS: sodium dodecyl sulfate

Spp.: species

T3SA: Type III Secretion Apparatus

T3SS: Type III Secretion System

TID: TurboID

TSA: tryptic soy agar

TSB: tryptic soy broth

Chapter 1 : Unraveling *Shigella*: Evolution, Pathogenesis, and Investigative Tools

1.1 Preface

Shigella flexneri is known for causing bacillary dysentery, a severe diarrheal disease with significant global health impact. *Shigella* pathogenicity depends on invasion and replication within intestinal epithelial cells via a sophisticated type III secretion system (T3SS) that coordinates virulence through regulatory proteins including the chaperone IpgC. Understanding *Shigella* pathogenesis requires detailed knowledge of T3SS components interactions within the confined bacterial cytoplasm. However, the dynamic nature of protein interactions has challenged traditional biochemical approaches. Proximity-dependent labeling technologies, particularly engineered biotin ligases, have opened new avenues for mapping protein neighborhoods and investigating bacterial secretion systems spatial organization.

This introduction chapter provides comprehensive background for this research, beginning with the historical context and clinical significance of *Shigella* infections, followed by current epidemiological concerns and therapeutic challenges. The chapter then examines *Shigella*'s evolutionary origins, genomic adaptations, and pathogenic mechanisms. A detailed exploration of the T3SS structure, assembly, and regulation follows. Finally, proximity labeling methodologies are reviewed, with emphasis on biotin ligase-based approaches for proteomics in bacterial systems. Together, these sections establish the rationale for employing proximity-dependent biotinylation to investigate the molecular environment surrounding the T3SS-associated protein IpgC and advance our understanding of the T3SS in *Shigella flexneri.*

1.2 Historical Overview of *Shigella*

Shigella is a genus of Gram-negative, non-motile, non-spore-forming bacilli that thrives in both aerobic and anaerobic conditions. It was named after Kiyoshi Shiga, the Japanese microbiologist who first isolated *Shigella dysenteriae* type 1 in 1897 during a severe dysentery outbreak in Japan (Lampel et al., 2018; Okui, 2024; Trofa et al., 1999). At the time, dysentery was so lethal that it was described as “more fatal to armies than powder and shot” underscoring

its public health impact (Kotloff et al., 2018; Osler, 1892). Following Shiga's discovery, the genus was expanded to include three additional species, *S. flexneri*, *S. boydii*, and *S. sonnei*. These species are classified into four serogroups (A-D) based on their O-antigen structures, and they display notable differences in global distribution, serological diversity, and disease severity. Although traditionally referred to as distinct species, these four groups are now increasingly recognized as subgroups within the *Escherichia coli* species. This shift reflects their close genetic relationship and acknowledges that comparisons among them should be framed at the subgroup level rather than as separate species (Dif et al., 2025; Hawkey et al., 2020; Pupo et al., 2000; J. Yang et al., 2007).

S*higella flexneri* (serogroup B), first characterized by Simon Flexner in 1900, is the most prevalent subgroup in low- and middle-income countries and remains a leading cause of shigellosis, particularly among young children in sub-Saharan Africa and South Asia (Anderson et al., 2016; Baker & Scott, 2023; Flexner, 1990; Stenhouse et al., 2023). In contrast, *Shigella sonnei* (serogroup D), isolated and characterized by Carl Sonne in 1906, has become the dominant subgroup in high-income countries and is increasingly emerging in transitional regions. This shift might be attributed to improvements in water quality, sanitation, and hygiene infrastructure, which affect the transmission dynamics of the different *Shigella* species. Unlike the other *Shigella* subgroups encompassing multiple serotypes, *S. sonnei* only includes one, yet it is remarkably adaptable and often associated with antimicrobial resistance (AMR), making it a growing concern in global health surveillance (Barceloux, 2008; Scott et al., 2025; Shiga, 1936). *Shigella boydii* (serogroup C), discovered by John Boyd in 1921, is predominantly found in South Asia, with sporadic cases reported outside the region, and makes minimal contribution to the global burden of shigellosis (Barceloux, 2008; The et al., 2016). Meanwhile, *Shigella dysenteriae* (serogroup A), particularly its serotype 1, the only serotype that produces Shiga toxin, has historically been responsible for several high-fatality outbreaks throughout the 20th century. The presence of Shiga toxin contributes to severe clinical outcomes such as hemolytic uremic syndrome and neurologic complications (Johannes & Römer, 2010). Although this toxin is classically linked to *S. dysenteriae* type 1, there have been rare reports of Shiga toxin genes detected in clinical isolates of other *Shigella* subgroups, including *S. flexneri*.

These findings underscore the evolving genetic landscape of *Shigella* and highlight the importance of continued surveillance for emerging virulence traits (Lampel et al., 2018). Altogether, the genus *Shigella* comprises over 50 serotypes, each with varying degrees of virulence, host immune evasion strategies, and antimicrobial resistance profiles. These differences are shaped by the acquisition and loss of mobile genetic elements, which contribute to the organism's rapid evolution and adaptation. (Bennish & Ahmed, 2020; Kotloff et al., 2018; NCCID, 2024; Nyarkoh et al., 2024; The et al., 2016).

1.3 Shigellosis: Clinical and Epidemiological Features

Shigella is a major contributor to the global burden of enteric infections, particularly as the causative agent of bacillary dysentery, or shigellosis. Although genetically related to *Escherichia coli* within the Enterobacteriaceae family, *Shigella* differs significantly in its pathogenic behavior. One of its defining features is an extremely low infectious dose (Kothary & Babu, 2001), as fewer than 10 bacterial cells are sufficient to cause disease, making it highly transmissible, especially in settings with inadequate sanitation and hygiene. Humans are the sole natural reservoir for *Shigella*, and transmission occurs predominantly via the fecal-oral route, either through direct person-to-person contact or through ingestion of contaminated food or water. (NCCID, 2024; The et al., 2016).Shigellosis is an acute enteric infection caused by *Shigella* spp, typically presenting with the abrupt onset of abdominal pain, fever, and diarrhea, which may range from watery to bloody and mucoid. These symptoms are often accompanied by nausea and general malaise. As the illness progresses, patients may develop anorexia, tenesmus, rectal discomfort, and frequent, small-volume bowel movements. In immunocompetent individuals, the disease course is generally self-limiting, resolving within a week or two. However, in some severe cases, complications such as toxic megacolon, seizures, kidney injury, and intestinal perforation may occur. In young children, elderly, and other vulnerable individuals, the infection is often more severe and prolonged, with an increased risk of morbidity and mortality (Akhondi et al., 2025; Aslam et al., 2025; DuPont et al., 1989; Government of Canada, 2024).

1.4 *Shigella* Today: A Continuing Public Health Concern

Estimates suggest that *Shigella* causes between 80 and 190 million cases of disease annually, resulting in 200,000 to 600,000 deaths worldwide. The majority of these cases (~90%) are attributed to *S. flexneri* and *S. sonnei*, with *S. flexneri* being the most frequently identified subgroup responsible for shigellosis among children under five years of age in low- and middle-income countries (Aslam et al., 2025; Nemhauser & Centers for Disease Control (U.S.), 2023; Scott et al., 2025; Troeger et al., 2018). Although *Shigella* infections have long been associated with low-income countries, recent outbreaks demonstrate their increasing prevalence in wealthier nations. Between 2020 and 2022, various developed countries in Europe reported numerous cases of shigellosis caused by *Shigella sonnei*, prompting the WHO Regional Office for Europe to issue an alert in March 2022. The alert was triggered by a rise in extensively drug-resistant (XDR) *S. sonnei* infections, particularly among men who have sex with men (PAHO WHO, 2022; Public Health Agency of Canada, 2024; WHO, 2022).

In North America, the situation is similarly concerning. In Canada, shigellosis is a nationally notifiable disease with approximately 880 cases reported annually, and rising trends have been observed (NCCID, 2024). During the COVID-19 pandemic, an outbreak of multidrug-resistant *S. flexneri* in downtown Vancouver led to over 100 confirmed cases. Strikingly, around 90% of affected individuals were experiencing homelessness and found in unsanitary street conditions; 88% had substance use disorders and required emergency care. One of the main facilitators of this outbreak was the lack of access to sanitation during lockdowns, as public washrooms in restaurants and other establishments were closed (Leung et al., 2025; Public Health Agency of Canada, 2024). Edmonton, Alberta, also faced a prolonged outbreak starting in September 2022 until March 2025, resulting in 448 confirmed cases and hospitalization rates of nearly 70% (Alberta Health Services, 2025). In the United States, a parallel outbreak took place in Seattle and its area, among people experiencing homelessness resulting in 117 cases of *S. flexneri*, with over half of them hospitalized (Leung et al., 2025).

These outbreaks emphasize that *Shigella* is an emerging public health concern even in high-income countries, where vulnerable populations, particularly people experiencing homelessness, face elevated risk due to inadequate access to sanitation and hygiene. Preventing

future outbreaks will require tackling underlying structural challenges, alongside long-term efforts to address broader societal issues such as the housing crisis, poverty, and substance use.

1.5 Management of Shigellosis: Therapeutic Strategies and Vaccines Development

The cornerstone of shigellosis management remains supportive care, with oral or intravenous rehydration depending on disease severity. Antimicrobial therapy is generally reserved for patients with severe symptoms, those who are immunocompromised, or individuals at risk of transmitting the infection, such as food handlers or caregivers. While antimicrobials can shorten illness duration and reduce transmission, their use must be carefully considered due to rising resistance patterns globally (Aslam et al., 2025; NCCID, 2024). Initially, antibiotics such as sulfonamides, tetracycline, and chloramphenicol were commonly used to manage *Shigella* infections, followed later by ampicillin and co-trimoxazole, a combination of sulfamethoxazole and trimethoprim. However, the rapid emergence of resistance to these drugs necessitated a transition to nalidixic acid and later to fluoroquinolones. Resistance to fluoroquinolones soon became widespread as well. In response, the World Health Organization recommended alternative agents, including ceftriaxone, azithromycin, and pivmecillinam. Nevertheless, since the early 2000s, increasing resistance to these alternatives has been documented, prompting consideration of carbapenems despite concerns about preserving these critical last-line antibiotics. (Aslam et al., 2025; Leow et al., 2024; Raso et al., 2023; Williams & Berkley, 2018).

The rapid and stepwise development of resistance to nearly all major classes of antimicrobials has significantly narrowed the therapeutic options for shigellosis. This challenge is further intensified by the global rise in multidrug-resistant (MDR) and extensively drug-resistant (XDR) *Shigella* strains, which complicate treatment decisions and limit the effectiveness of standard therapies (Leung et al., 2025). Considering these persistent limitations, vaccines have long been viewed as a promising long-term strategy in the fight against *Shigella*. Earlier vaccine efforts, which focused on whole-cell or O-antigen components, were hindered by limited cross-protection and high serotype variability. However, recent advances in genomics and immunology have enabled the development of more refined candidates, including live-attenuated, protein subunit, outer membrane vesicle based, and glycoconjugate vaccines (Baker

& Scott, 2023; Leow et al., 2024). Several candidates, such as InvaplexAR, Shigella4V, altSonflex 1-2-3, SF2a-TT15, and ZF0901, are currently undergoing clinical evaluation, with early trials indicating favorable safety and immunogenicity profiles (Braun et al., 2021; Leow et al., 2024; Medicine NLo, 2020, 2021; Mo et al., 2021; Turbyfill et al., 2023; Walker, 2025). Given the escalating threat of antimicrobial resistance, the World Health Organization has identified *Shigella* as a priority pathogen for research and development (Kaminski et al., 2025). Given the limited treatment options and the growing threat of resistance, a deeper understanding of *Shigella*'s virulence mechanisms, such as those involving the type III secretion system, could be key to identifying novel pharmacological targets and guiding future therapeutic development.

1.6 Evolutionary Origins and Genomic Adaptation of *Shigella*

The taxonomic status of *Shigella* has undergone significant re-evaluation due to advances in molecular phylogenetics. Although historically treated as a separate genus due to its clinical relevance and distinct pathogenic profile, genomic analyses have revealed that *Shigella* species are not monophyletic but instead represent multiple lineages embedded within the broader *Escherichia coli* species complex (Abram et al., 2021; Dif et al., 2025; Lan & Reeves, 2002; Pupo et al., 2000). Molecular clock estimates suggest that these lineages diverged from ancestral *E. coli* between 35,000 and 270,000 years ago, indicating a relatively recent evolutionary emergence. This phylogenetic positioning has prompted calls to reclassify *Shigella* as pathogenic subgroups of *E. coli*, reflecting its evolutionary origin more accurately (Dif et al., 2025; Hawkey et al., 2020; Pupo et al., 2000; J. Yang et al., 2007).

The emergence of *Shigella* as a human-adapted pathogen is a product of convergent evolution, where distinct *E. coli* lineages independently acquired similar traits that enable them to invade and replicate within human intestinal epithelial cells. This convergence is not the result of a single evolutionary event but rather multiple, parallel adaptations that led to the formation of the four recognized *Shigella* subgroups (Hawkey et al., 2020; Pupo et al., 2000; The et al., 2016; J. Yang et al., 2007). These adaptations include the acquisition of the 220 kbp virulence plasmid pINV, which encodes the T3SS and the actin-polymerizing protein IcsA, essential for the invasion of the host cell and cytoplasmic motility, respectively (Bernardini et al., 1989; Sansonetti et al., 1982). Additionally, genomic islands such as *Shigella* island 1 (SHI-1) and

Shigella island 2 (SHI-2), which encode toxins and iron acquisition systems, further enhanced virulence (Hawkey et al., 2020; Moss et al., 1999; Rajakumar et al., 1997; The et al., 2016).

The evolutionary origin of the pINV plasmid in *Shigella* remains a subject of debate, with two primary hypotheses proposed. The first suggests that pINV was acquired independently by multiple *E. coli* lineages, each giving rise to a distinct *Shigella* clade. This model is supported by phylogenomic evidence showing that *Shigella* strains are distributed across several *E. coli* phylogroups and exhibit lineage-specific adaptations (Pupo et al., 2000). However, a major challenge to this theory is that pINV is not conjugative, hence it lacks the machinery for self-transfer, making widespread horizontal transmission between lineages difficult to explain (Sansonetti et al., 1982). The alternative hypothesis supports a single ancestral acquisition of pINV by a common progenitor of *Shigella*, followed by differential retention in some lineages and loss in others (Escobar-Paramo et al., 2003). This model aligns with the non-transmissibility of pINV but raises questions about how such an unstable plasmid could be maintained over long evolutionary timescales. Supporting this view, *Shigella* spp. have evolved distinct plasmid maintenance systems. For instance, *S. flexneri* and *S. sonnei* differ in their toxin-antitoxin modules and partitioning systems, which influence pINV stability and, consequently, virulence gene expression (Martyn et al., 2022; McVicker et al., 2019).

In parallel with gene acquisition, *Shigella* evolution is marked by extensive gene loss and genome reduction, a trend commonly observed in pathogenic bacteria (Maurelli et al., 1998; Maurelli, 2007). Several metabolic and structural genes present in *E. coli* are either inactivated or deleted in *Shigella*. For instance, genes involved in motility (e.g., flagellar operons) and adhesion (e.g., fimbriae) are frequently disrupted, likely as a strategy to evade host immune detection (Bravo et al., 2015; Feng et al., 2011; The et al., 2016; F. Yang, 2005). Moreover, key metabolic genes such as *nadAB* (involved in NAD biosynthesis), *cadA* (lysine decarboxylation), *speG* (polyamine detoxification), and *ompT* (outer membrane protease) are consistently lost across *Shigella* species (Barbagallo et al., 2011; Maurelli et al., 1998; Prunier et al., 2007; The et al., 2016; F. Yang, 2005). Restoration of many of these pathways has been shown to attenuate virulence, underscoring their role in host adaptation (Barbagallo et al., 2011; Prosseda et al., 2012; Prunier et al., 2007). The main driving force behind genome reduction in *Shigella* is the expansion of insertion sequences (annotated as IS in genomes) mobile genetic elements that

disrupt coding regions and promote genomic rearrangements (The et al., 2016). The impact of insertion sequences has been further elucidated through genome-scale metabolic modeling. These analyses demonstrate that *Shigella* spp. have lost numerous metabolic functions compared to *E. coli*, especially those involved in carbohydrate metabolism. Among the subgroups, *S. dysenteriae* shows the most extensive metabolic reduction, followed by *S. flexneri* and *S. sonnei* (Hawkey et al., 2020; Pilla et al., 2017). Interestingly, while the same metabolic functions are often lost across species, the underlying genetic disruptions differ, indicating convergent evolution at the functional level but divergence in the mutational mechanisms (Hawkey et al., 2020).

In summary, *Shigella* evolved from commensal *E. coli* through a combination of gene acquisition (plasmid pINV and chromosomal such as SHI-1), insertion sequence expansion, and targeted gene loss. These changes produced a group of pathoadapted lineages that, despite genetic diversity, share a unified pathogenic strategy, making *Shigella* a valuable model for exploring bacterial adaptation to intracellular pathogenic lifestyles.

1.7 *Shigella* Pathogenesis and Survival Mechanisms

While the four *Shigella* subgroups; *S. flexneri*, *S. sonnei*, *S. dysenteriae*, and *S. boydii*, utilize a core set of virulence factors, primarily encoded on a shared invasion plasmid, to execute similar infection strategies, variations in their regulation and efficacy lead to distinct clinical outcomes and epidemiological patterns. *S. flexneri*, accounting for approximately 60% of shigellosis cases in developing regions and distinguished by its potent virulence, serves as the primary model for studying *Shigella* pathogenesis due to its prevalence and experimental tractability (Anderson et al., 2016; Mattock & Blocker, 2017; The et al., 2016). The following sections focus on *S. flexneri* to elucidate its multifaceted pathogenesis and survival strategies, including the T3SS, intracellular lifestyle and cell-to-cell spread, and stress responses, which collectively enable its persistence and disease severity in the host.

1.7.1 *Shigella*'s Intercellular Lifecycle

Shigella initiates infection by exploiting M cells in Peyer's patches, which serve as portals through the intestinal epithelial barrier (Jensen et al., 1998). The bacterium employs a T3SS to coordinate invasion, with IpaD at the needle tip detecting host cell contact and triggering the release of effector proteins. IpaB and IpaC form translocation pores in the host membrane, enabling bacterial entry. IpaC also alters the host's actin cytoskeleton, inducing membrane ruffles and macropinocytosis to facilitate uptake into vacuoles (Cossart & Sansonetti, 2004; High et al., 1992; Picking & Picking, 2016). Additional effectors like IpgD, a phosphatidylinositol 4,5-bisphosphate phosphatase, further enhance invasion by reshaping membrane dynamics and promoting actin rearrangement (Mellouk et al., 2014). Once inside, *Shigella* rapidly escapes the vacuole, a process primarily facilitated by the formation of a translocation pore by IpaB and IpaC (Campbell-Valois et al., 2014; Duncan-Lowey et al., 2020; High et al., 1992; Page et al., 1999).

After entering the host cell cytoplasm, *Shigella* replicates and moves using actin-based motility. This process is facilitated by IcsA, which recruits host proteins such as the Neural Wiskott-Aldrich Syndrome Protein (N-WASP) and the Arp2/3 complex, triggering the formation of actin-rich structures known as comet tails. These structures enable the bacterium to travel within the cell at speeds ranging from 1-50 μm/min (Bernardini et al., 1989; Egile et al., 1999; Valenzuela-Valderas et al., 2023). Directional movement is maintained by IcsP, a protease that restricts IcsA activity to one pole (Hensley et al., 2011; Shere et al., 1997). After replicating in the cytoplasm, *Shigella* spreads laterally between epithelial cells, evading immune surveillance and promoting colonization of the mucosa. The T3SS is dynamically regulated; highly active during invasion and vacuole escape, but suppressed in the cytoplasm to facilitate efficient intercellular spread (Campbell-Valois et al., 2014). IcsA, localized at one pole of the bacterium, recruits N-WASP and activates the Arp2/3 complex to initiate actin comet tail formation. At cell junctions, *Shigella* induces actin-driven membrane protrusions, a process supported by host factors such as formins and myosins, which facilitate the formation and extension of these structures (Agaisse, 2016). These protrusions are then engulfed by adjacent cells, forming double-membrane vacuoles. IpaB and IpaC disrupt these vacuoles, while other effectors including IcsB, VirA, and IpaH9.8 interfere with host defenses such as membrane trafficking and repair, autophagy and innate immunity (Agaisse, 2016; Bajunaid et al., 2020; Ogawa et al.,

2005). Together, these mechanisms allow *Shigella* to efficiently colonize the intestinal epithelium by avoiding extracellular immune responses and sustaining intracellular spread.

Among enteric pathogens, *Shigella*'s mechanisms of host invasion and intracellular survival are distinctive yet share similarities with other pathogens. Both *Shigella* and *Listeria monocytogenes* colonize the host cell cytoplasm, yet they employ distinct strategies for invasion and intercellular spread. While *Listeria* uses listeriolysin O and ActA, *Shigella* relies on IpaB, IpaC, and IcsA (Rey et al., 2020). Enteroinvasive *E. coli* (EIEC), a close relative of *Shigella*, shares many T3SS effectors with the latter, but shows reduced efficiency in cell-to-cell spread (Ramos Moreno et al., 2009). In contrast, *Salmonella* uses its SPI-1 T3SS to invade host cells, but once internalized, it hijacks the host endocytic system to form a specialized compartment known as the *Salmonella*-containing vacuole (SCV) that supports intracellular survival and systemic dissemination (Stévenin et al., 2019). Comparative studies of these pathogens not only highlight the diversity of strategies used to colonize intracellular niches but also underscore the evolutionary flexibility bacteria employ to thrive within host cells.

1.7.2 Adapting Under Pressure: *Shigella*'s Stress Response Mechanisms

As *Shigella* traverses the host's gastrointestinal tract, it encounters a range of hostile conditions, including oxidative stress from immune defenses, limited nutrient availability, temperature changes, and extreme pH variations. To adapt and survive, *Shigella* activates a suite of regulatory systems that coordinate its stress response and virulence, with the stringent response pathway, driven by the nucleotide alarmones (p)ppGpp, playing a central role. Originally characterized in *E. coli*, the stringent response is a core bacterial adaptation mechanism triggered by harsh environmental conditions like nutrient deprivation and chemical stressors (Cashel & Gallant, 1969; Kago et al., 2024). In *S. flexneri*, the synthesis of (p)ppGpp has been recently shown to be also crucial for pathogenicity. Disruption of the *relA* gene, which encodes the primary enzyme responsible for (p)ppGpp production, leads to a marked decrease in the expression of key virulence genes. Mutants lacking (p)ppGpp show diminished levels of invasion plasmid antigens and the intercellular spread protein IcsA, resulting in impaired epithelial cell invasion and reduced intercellular dissemination (Kago et al., 2024). Downstream of (p)ppGpp, the sigma factor RpoS acts as a master regulator of stress-related gene expression.

Its activity increases under stress conditions, further enhancing bacterial resilience (Small et al., 1994; Traxler et al., 2011).

To counteract oxidative damage from reactive oxygen species (ROS) produced by host immune cells, *S. flexneri* is proposed to utilize regulatory systems like OxyR and SoxRS to activate antioxidant gene expression, such as superoxide dismutase, enhancing bacterial persistence during infection (Daugherty et al., 2012). Additionally, efflux systems such as AcrAB-TolC contribute to multidrug resistance by expelling harmful compounds like antibiotics and bile salts (Coluccia et al., 2023), while the EmrKY pump, part of the major facilitator superfamily (MFS), supports survival within macrophages (Pasqua et al., 2019). Environmental sensing is further refined through two-component systems like PhoP/PhoQ, which detect and respond to changes in magnesium levels, antimicrobial peptides, and pH (Lin et al., 2018).

1.7.3 Molecular Mechanisms of Acid Resistance in *Shigella*

Acid resistance is crucial for *Shigella*'s survival in highly acidic environments, particularly enabling passage through stomach acid (pH < 3) to reach the large intestine (pH 5-8) where infection is established (Chiang et al., 2021; Small et al., 1994). The primary mechanism involves the glutamate-dependent acid resistance (GDAR or GAD) system, which utilizes glutamate decarboxylase enzymes, converting glutamate to γ-aminobutyric acid (GABA) while consuming cytoplasmic protons and raising intracellular pH. The GadC antiporter then exchanges external glutamate for internal GABA, maintaining the acid resistance cycle (De Biase & Pennacchietti, 2012; Richard & Foster, 2004). *S. flexneri* possesses two distinct isoforms of glutamate decarboxylase (GadA and GadB) that are differentially regulated within complex regulatory networks. The transcriptional regulator SlyA, belonging to the MarR family, plays a pivotal role in this acid resistance system by directly binding to the *gadA* promoter and activating its expression (Weatherspoon-Griffin & Wing, 2016; Zhang et al., 2018).

Shigella also employs hydrogenase systems, particularly the Hya enzyme, as an alternative acid resistance mechanism. Under anaerobic fermentative conditions, the Hya enzyme becomes highly active upon acid exposure, with substantial rates of H_2 oxidation providing metabolic support through proton-consuming hydrogen oxidation reactions (McNorton & Maier, 2012). This system provides an additional pathway that complements the glutamate

decarboxylase system. Additionally, the periplasmic chaperone HdeA contributes significantly to acid tolerance by undergoing conformational changes at low pH, shifting from a dimer to a monomer configuration to prevent aggregation of acid-denatured proteins, thereby maintaining protein integrity during acid exposure (Gajiwala & Burley, 2000).

The integration of *Shigella*'s acid resistance systems with broader stress responses, such as RpoS and (p)ppGpp signaling, highlights its adaptability in surviving passage through the acidic gastric barrier to establish infection. However, the potential crosstalk between acid resistance mechanisms and virulence pathways like the T3SS remains poorly understood. Whether these systems influence one another's activation or regulation during infection is an open question. Clarifying this relationship is not only important for understanding *Shigella*'s adaptive strategies, but may also reveal points of vulnerability. For example, if acid-response pathways contribute to modulating or regulating virulence, then disrupting this crosstalk could represent a promising avenue for future anti-virulence therapeutic strategies.

1.7.4 The Type III Secretion System

The pathogenic potential of *Shigella* is primarily dependent on its T3SS. The T3SS encodes the Type III Secretion Apparatus (T3SA), which is also called the syringe, the injectisome or simply the T3SS, along with associated regulators, translocated effector proteins, and their cognate molecular chaperones. This system operates through precisely coordinated mechanisms to facilitate the direct translocation of bacterial virulence effectors into host cells, thus enabling *Shigella* to successfully invade, survive, disseminate between cells, and exert disease pathogenesis (Bajunaid et al., 2020; Haidar-Ahmad et al., 2023). The structure of the injectisome exhibits remarkable conservation across diverse gram-negative bacterial pathogens, including *Shigella*, *Salmonella*, pathogenic *Escherichia coli*, *Pseudomonas*, and *Yersinia* species (Izoré et al., 2011; S. Johnson et al., 2019). Over the past decade, intensive research efforts have been directed toward studying the *Shigella* T3SS, with numerous investigations to decipher its complex architecture, functional mechanisms, and regulatory systems (Deng et al., 2017; Haidar-Ahmad et al., 2023).

1.8 A Deep Dive into The Type III Secretion System of *Shigella*

Given its essential role in *Shigella* virulence, the T3SS is a central focus of this PhD project. This section provides a comprehensive overview of the T3SS, with emphasis on three key aspects: the structural architecture of the injectisome nanomachine, the sequential assembly process that governs its formation, and the complex regulatory mechanisms that fine-tune its activity in response to environmental and host-derived signals.

Much of the content presented here is developed in detail in two recent reviews published by our lab, to which I contributed as a co-author, offering a thorough review of current knowledge and emerging insights into the *Shigella* T3SS (Bajunaid et al., 2020; Haidar-Ahmad et al., 2023).

1.8.1 Structural Architecture of the *Shigella* T3SA

The 3.5 MDa *Shigella* T3SA is encoded by approximately 20 genes within a ~30 kbp segment of the virulence plasmid, termed the entry region, encompassing the *mxi/spa* and *ipa* operons (Buchrieser et al., 2000; Venkatesan et al., 2001). Under permissive conditions, such as temperatures above 32°C, these operons express the proteins that assemble into the functional T3SA (K. A. Kane & Dorman, 2012). This syringe-like structure traverses the bacterial inner membrane, periplasm, outer membrane, and extends into the extracellular space, facilitating effector protein delivery into host cells (Wagner et al., 2018). Since 2020, cutting-edge techniques like cryo-electron microscopy (cryo-EM) and cryo-electron tomography (cryo-ET), have significantly advanced our understanding of the T3SA's architecture, revealing intricate details of its molecular organization and assembly (B. Hu et al., 2015; J. Hu et al., 2018, 2020; Lunelli et al., 2020; Majewski et al., 2019).

The needle complex was the earliest T3SS component studied, with pioneering electron microscopy studies in the late 1990s and early 2000s (Blocker et al., 2001; Kubori et al., 1998), revealing its syringe-like architecture composed of approximately 100 MxiH subunits extending around 50 nm from the outer membrane (Figure 1.1). Recent studies have refined the needle's structure, confirming that MxiH adopts a helical, α-hairpin structure, forming a narrow 1.5-2.5 nm channel to facilitate the passage of unfolded effectors (Habenstein et al., 2019). The needle is capped by a tip complex, either an IpaD homopentamer or a heteropentamer of four IpaD and

one IpaB subunits, with IpaB's N-terminus stabilizing a closed state to prevent premature secretion (Barta et al., 2018; Cheung et al., 2015; Dickenson et al., 2013).

The basal body of T3SA serves as the structural basis of the injectisome (Figure 1.1), anchoring this machinery across the bacterial inner and outer membranes. It comprises two main components: the outer ring (OR), formed by the secretin protein MxiD, and the inner rings (IR), composed of MxiG and MxiJ (Lunelli et al., 2020). The basal body acts as a base anchoring the needle and integrating it with the inner rod, creating a continuous channel from the bacterial cytoplasm to the extracellular environment (Torres-Vargas et al., 2019).

The export apparatus comprises Spa24, Spa9, Spa29, and Spa40, with a stoichiometry of 5:4:1:1, forming an inverted conical structure beneath the inner rings (S. Johnson et al., 2019; Kuhlen et al., 2020). MxiA is positioned at the secretion channel's entrance, serving as the gate of the export apparatus, with the sorting platform located directly beneath it. This sorting platform, critical for substrate selection and secretion hierarchy, adopts a hexagonal cage-like structure with six pods, each comprising SpaO (AKA Spa33) complexed with MxiK and MxiN (J. Hu et al., 2020). These pods are connected via MxiN spokes to the central core consisting of a homohexameric ATPase SpaL (AKA Spa47), which is stabilized by the stalk protein Spa13 (Cherradi et al., 2014).

1.8.2 *Shigella* T3SA Secretion and Assembly

Although the early stages of *Shigella*'s T3SA assembly are not fully characterized, its structural and functional resemblance to the *Salmonella* SPI-1 and *E. coli* systems supports an "inside-out" assembly model. This begins with the Sec-dependent integration of inner membrane-embedded core export apparatus components, around which the inner ring is formed. Subsequently, the outer ring assembles, completing the basal body structure (Deng et al., 2017; Wagner et al., 2018). Following basal body formation, the cytosolic sorting platform attaches, and the ATPase SpaL energizes substrate processing, initiating the secretion of substrates. The *Shigella* T3SA secretome comprises roughly 35 proteins, most encoded on the virulence plasmid within T3SS-related operons. However, effectors including some members of the *ipaH* family and substrates like *icaR*, *icaT*, and *yccE*, are chromosomally encoded (Hall et al., 2022; Pinaud et al., 2017; Silué & Campbell-Valois, 2022).

Secretion begins with early substrates like MxiI and MxiH, which polymerize to form the inner rod and the extracellular needle, respectively. These structures create a continuous conduit with the basal body for substrate delivery. Upon reaching optimal needle length, the regulator Spa32 is secreted, lifting Spa40-mediated inhibition and enabling the transition to middle substrate secretion (Deng et al., 2017; Journet et al., 2003; Wagner et al., 2018). T3SA activation occurs upon host cell contact or chemical induction (e.g., Congo red), shifting the system from an inactive to an active state (Ménard, Sansonetti, & Parsot, 1994; Parsot et al., 1995). The tip complex, composed of IpaB and IpaD, detects host interaction and undergoes conformational changes, allowing IpaB's transmembrane domain to insert into the host membrane. IpaC is then secreted and integrates into the host membrane, forming the translocon pore, a direct passage between the bacterial cytoplasm and the host cell (Figure 1.1). Although the activation signal is transmitted through the needle to the basal export apparatus and sorting platform, the precise molecular mechanisms remain unclear (Barta et al., 2018; Olive et al., 2007; Veenendaal et al., 2007).

Substrate secretion follows a tightly regulated hierarchy. SpaL plays a central role by separating substrates from their protein chaperone escort, and unfolding them for secretion (Akeda & Galán, 2005). Initially, middle substrates are prioritized. Once depleted, the gatekeeper protein MxiC, previously inhibiting late substrate secretion, is itself secreted, permitting the processing of late substrates (Botteaux et al., 2009; Roehrich et al., 2017; D.-K. Shen & Blocker, 2016). These late substrates are divided into first-wave effectors, pre-stored in the cytoplasm with specific chaperones, and second-wave effectors, which are unchaperoned and secreted immediately upon synthesis (Ernst et al., 2018; Lara-Tejero et al., 2011; Portaliou et al., 2016). This temporal hierarchy, modulated by chaperone diversity and MxiC, ensures efficient and strategic delivery of effectors to optimize infection.

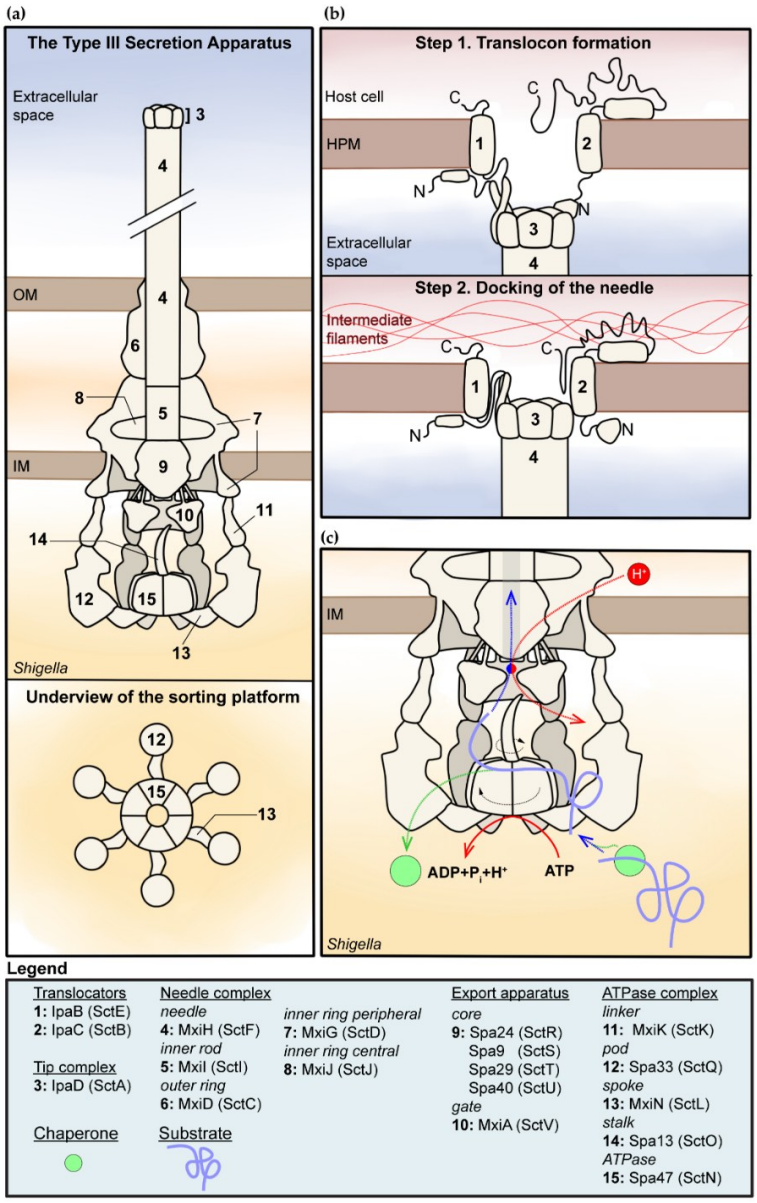


Figure 1.1. Overview of the Type III Secretion Apparatus in *Shigella* spp.
(a) Schematic view of the T3SA in its inactive state, showing the tip complex and cytosolic components in 3D, with the apparatus body in cross-section. The bottom panel illustrates the sorting platform with a bottom-up perspective. (b) Model of translocon formation and the docking of the needle on the host plasma membrane (HPM) in association with intermediate filaments. (c) Model of substrate secretion through the T3SA. Chaperones deliver substrates (purple) to the secretion machinery, where ATP hydrolysis and proton motive force drive the unfolding and export of the substrates through the needle. Arrows indicate movement (dashed line) and rotation (curved line); the red-blue circle represents coupling between substrate export and proton flow. Component names are shown for *Shigella* and in standardized nomenclature. As a reminder Spa33 and Spa47 will be presented elsewhere in the thesis with their alternative name SpaO and SpaL, respectively. This figure is adapted from (Bajunaid et al., 2020) under a Creative Commons Attribution (CC BY 4.0) license.

1.8.3 Regulation of the T3SS in *Shigella*

Shigella's T3SS regulation is a multi-layered process responsive to environmental cues, primarily temperature, and the functional state of the secretion apparatus. This system integrates global repressors, transcriptional activators, and feedback mechanisms to ensure precise, host-specific expression of virulence genes. At the core of *Shigella*'s T3SS regulation is a temperature-sensitive cascade mediated by the nucleoid-associated protein H-NS, which acts as a global transcriptional silencer (Picker & Wing, 2016; Riccardi et al., 2019; B. A. Shen et al., 2022). Below 30°C, H-NS binds to AT-rich regions of the virulence plasmid and blocks the access of transcription regulators as well as RNA polymerase to promoters of T3SS genes, including *virF* and *virB* (Beloin & Dorman, 2003). Upon exposure to host-like conditions at 37°C, temperature-induced changes in DNA supercoiling destabilize H-NS binding, enabling expression of *virF*, an AraC-family transcriptional activator (Falconi et al., 1998). VirF senses host temperature and initiates the virulence cascade by binding to the *virB* promoter, counteracting H-NS repression. VirB, a ParB-family antisilencer, amplifies this derepression by displacing H-NS from additional promoter regions, activating the transcription of T3SS related genes (K. A. Kane & Dorman, 2012; McKenna et al., 2022; Weatherspoon-Griffin et al., 2018).

A second regulatory layer, centered on the AraC-family activator MxiE, governs the expression of late effector genes. MxiE activity is tightly coupled to the functional state of the secretion apparatus rather than temperature cues, providing dynamic control over effector secretion. In the absence of T3SS activity, MxiE is sequestered by OspD1 and its chaperone Spa15, which stabilizes MxiE and prevents premature transcription of second-wave late substrates. Similarly, IpgC is sequestered by its cargo effectors, IpaB and IpaC, preventing their aggregation. Upon host cell contact, sensed by the T3SS needle tip, IpaB, IpaC, and the anti-activator OspD1 are secreted, freeing IpgC and releasing MxiE from sequestration (Ferrari et al., 2021; C. D. Kane et al., 2002; Mavris, Page, et al., 2002; Parsot et al., 2005). The MxiE-IpgC complex then binds to MxiE boxes in the promoters of late effector genes, leading to their transcription and secretion (Figure 1.2) (Bongrand et al., 2012; Mavris, Sansonetti, et al., 2002). This complex can also form a negative feedback loop by repressing the transcription of *virB*, thereby balancing the virulence cascade (McKenna et al., 2022).

SlyA, a member of the MarR family of transcriptional regulators previously discussed in the context of acid resistance, has been suggested to play a role in regulating T3SS genes in *Shigella*. Research has shown that when *slyA* is overexpressed in strains lacking *virB*, it can restore the expression of genes encoding injectisome components (Weatherspoon-Griffin & Wing, 2016). One possible explanation is that SlyA may help lift the repression caused by H-NS, a protein that normally silences these genes. This hypothesis is based on findings in *E. coli*, where SlyA competes with H-NS to allow gene expression, suggesting that a similar anti-silencing mechanism might also occur in *Shigella* (Lithgow et al., 2007).

Other environmental modulators, including pH (CpxAR), oxygen levels (FNR), and iron concentrations (RyhB), fine-tune T3SS regulation through the addition of other layers of control (Broach et al., 2012; Haidar-Ahmad et al., 2023; Marteyn et al., 2010; Murphy & Payne, 2007; Nakayama & Watanabe, 1995; Vergara-Irigaray et al., 2014). While these studies highlight how housekeeping and stress response pathways influence T3SS activity, it remains unclear whether the reverse is true; can T3SS activity itself modulate housekeeping functions or broader cellular physiology? This question marks a significant gap in our understanding of the relationship between virulence and core bacterial processes.

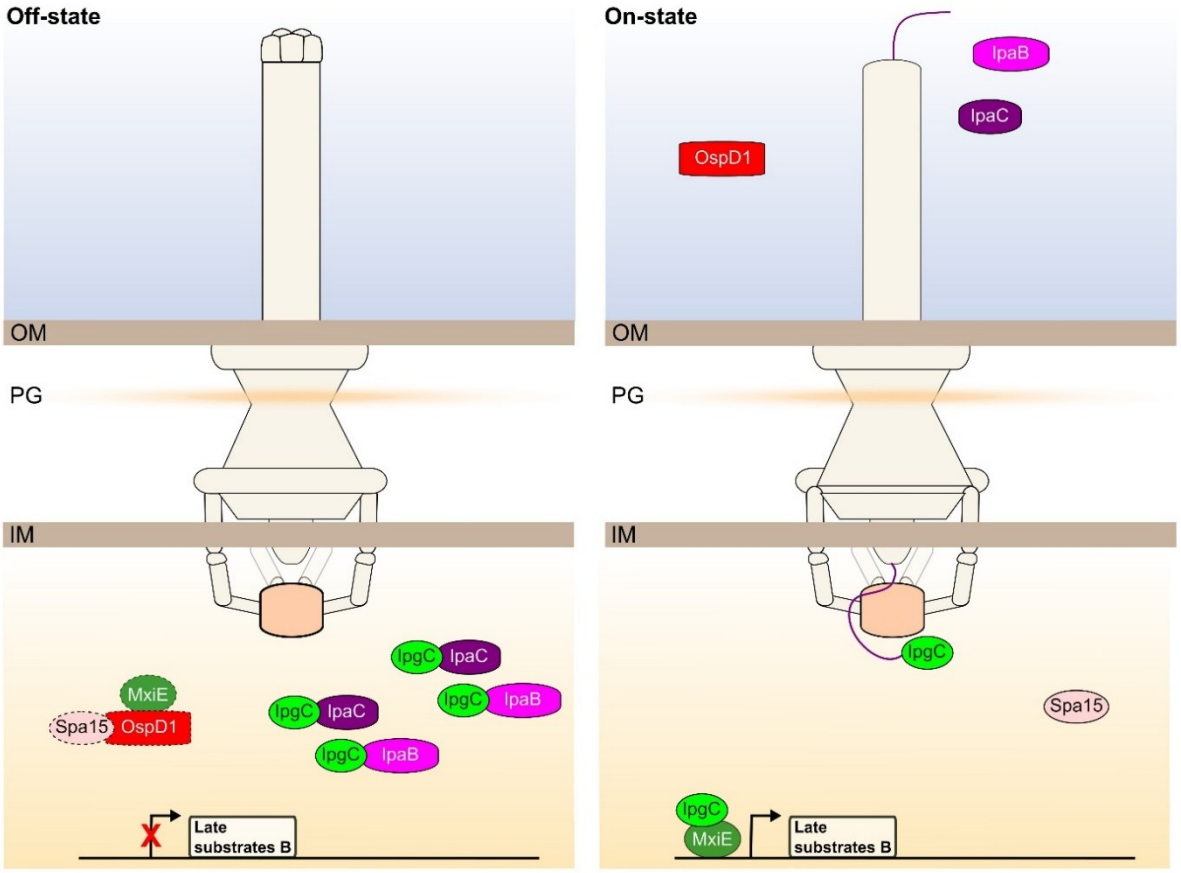


Figure 1.2. Model describing the IpgC-MxiE regulatory system in both T3SS states.
In the T3SS-off state, MxiE and IpgC are sequestered by OspD1-Spa15 and IpaBC respectively, preventing premature transcription of late effectors. Upon T3SS activation, triggered by host cell contact, IpaB, IpaC, and the anti-activator OspD1 are secreted via the T3SS needle complex, freeing IpgC to interact with MxiE. Together, IpgC and MxiE form a transcriptional activation complex that initiates the late substrates transcription and secretion. The T3SS model has been described and adapted from (Bajunaid et al., 2020; Parsot et al., 2005). This schematic is also revised in Chapter 4 based on insights derived from my thesis.
OM: outer membrane; PG: peptidoglycan cell wall; IM: inner membrane.

1.9 Proximity Labeling in Proteomics

While traditional approaches have provided valuable insights into T3SS components and their basic functions, our understanding of the changes that occur during the active and inactive states of this system remains limited. The dynamic nature of the T3SS combined with the small volume of the bacterial cytoplasm, poses unique challenges for mapping proxisomes or protein-protein interactions during different functional states. In this study, proximity labeling, particularly biotin ligase-based approaches, was employed to investigate the T3SS in *Shigella flexneri*. This technique was selected for its high specificity, compatibility with mass spectrometry workflows, and ability to label proteins in their native cellular environments without disrupting complex assembly or function. By capturing transient and spatially restricted interactions, proximity labeling offers a snapshot of the dynamic protein complexes that facilitate T3SS function. These are insights that static structural studies may struggle to provide. This section describes the most widely used proximity labeling techniques with a particular focus on biotin ligase-dependent labeling systems.

1.9.1 Overview of Key Proximity Labeling Approaches

Great advancements in protein-protein interactions and protein localization studies are due to proximity labeling, which enables the identification of proteins that are vicinal to each other in their natural environment within living cells. This technique overcomes critical limitations of traditional methods based on binding, like co-immunoprecipitation, which often miss transient or weak interactions that are not maintained after cellular disruption (Guo et al., 2023). There are three key techniques for proximity labeling, they involve peroxidases, biotin ligases, and pupylation-based interaction tagging. All of these methods involve fusing an engineered enzyme to a bait protein, which then catalyzes the covalent tagging of nearby proteins within a defined radius

Biotin ligase-based proximity labeling uses engineered ligases that activate biotin to form a reactive intermediate (biotinyl-5'-AMP), which selectively reacts with lysine residues on adjacent proteins, forming stable amide bonds. The resulting biotinylated proteins are subsequently enriched using streptavidin affinity purification and identified by mass spectrometry, creating comprehensive maps of protein neighborhoods or "proxisomes", as well

as identifying protein-protein interactions (Choi-Rhee et al., 2004; Cronan, 2024; Wu et al., 2025).

Another strategy employs peroxidase enzymes, such as horseradish peroxidase (HRP), which is extensively used due to its high catalytic activity and effectiveness in biochemical assays. Activated by H_2O_2, HRP generates highly reactive radicals from substrates like biotin-tyramide that covalently modify nearby proteins at electron rich residues, primarily tyrosine, with additional reactivity toward tryptophan, histidine, and cysteine. However, HRP is unsuitable for cytosolic environments, as its proper folding and stability depend on disulfide bonds and calcium-binding sites, which are incompatible with the reducing conditions of the cytosol (Azevedo et al., 2003; Liu et al., 2018; Samavarchi-Tehrani et al., 2020). To address this limitation, researchers turned to plant-derived ascorbate peroxidase (APX), which is more compatible with intracellular environments. An enhanced variant of APX for proximity labeling named APEX was obtained by introducing three mutations (K14D, W41F, E112K) to prevent dimerization, enabling its use in electron microscopy and proximity labeling (Bosch et al., 2021; Martell et al., 2012). APEX catalyzes the oxidation of biotin-phenol to generate phenoxyl radicals that tag nearby proteins in a similar manner to HRP. Although APEX was a major advancement, its sensitivity for proteomics was limited. To address this, a further mutation, A134P, led to the development of APEX2, which offers enhanced catalytic efficiency and enables rapid proximity labeling, typically within 1 minute (Guo et al., 2023; Lam et al., 2015; Martell et al., 2012). Nevertheless, peroxidase-based methods require H_2O_2 and phenol derivatives such as biotin-phenol or biotin-tyramide, which can induce oxidative stress in living cells (Guo et al., 2023).

One of the newest techniques is Pupylation-based interaction tagging (PUP-IT). It was originally developed to identify membrane proteins interactions by utilizing a prokaryotic enzyme (PafA), fused to a bait protein, to covalently attach the Pup protein to neighboring prey proteins. The reaction forms a bond between the C-terminal glutamate residue on Pup and the ε-amino group of lysine on the target protein (Liu et al., 2018). Due to Pup's large size and limited diffusion, PUP-IT offers a highly specific labeling radius. Unlike APEX, it avoids toxic substrates but operates more slowly. However, the large size of PafA (54 kDa) and substrate restricts its use to extracellular environments (Bosch et al., 2021; Liu et al., 2018; Shkel et al.,

2022). PUP-IT2 addresses this problem by fusing the smaller PupE (7 kDa) instead of PafA to the bait, while free PafA mediates the labeling. Though more specific, PUP-IT2 supports only single-turnover tagging, as PupE becomes covalently linked to both bait and prey (Yue et al., 2022).

While peroxidase and pupylation-based techniques offer valuable specificity and spatial resolution, their limitations, such as cytotoxicity, size, and single-turnover constraints, make them less suitable for this project. This has led us to focus on biotin ligases, which provide a more versatile and high-throughput platform for mapping protein-protein interactions.

1.9.2 Biotin Ligases in Proximity Labeling

The foundation of biotin ligase (BL) proximity labeling stems from engineering *E. coli* BirA, whose natural function is to biotinylate the biotin carboxyl carrier protein (BccP) subunit of acetyl-CoA carboxylase with high specificity (Figure 1.3). BirA uses ATP and Biotin to form the reactive biotinyl-5'-AMP as an intermediate through the attack of biotin's carboxylate group on the α-phosphate of ATP (Figure 1.4). This intermediate is tightly retained in the active site until transfer exclusively to the accepting lysine residue of BccP, forming a stable amide bond (also known as an isopeptide bond) and releasing AMP (Figure 1.3) (Adikaram & Beckett, 2012; Chapman-Smith & Cronan, 1999; Samavarchi-Tehrani et al., 2020).

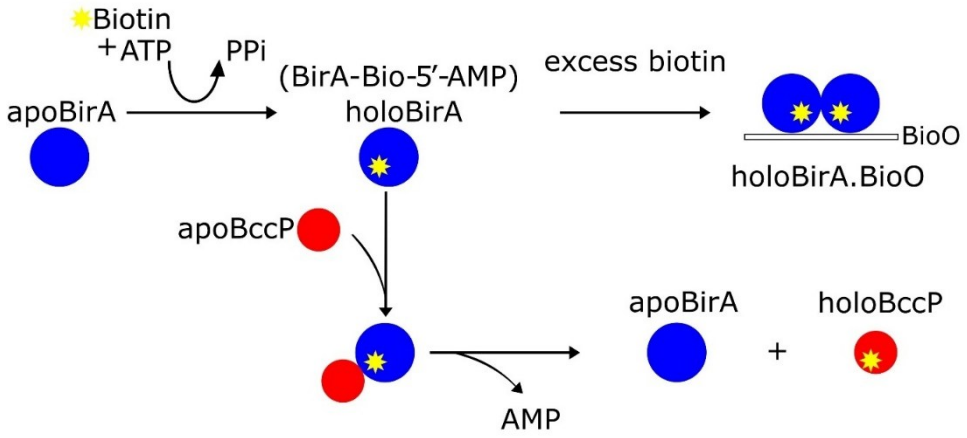


Figure 1.3. The biotin ligase BirA native functions in *E. coli*.
In the presence of biotin and ATP, the *E. coli* biotin ligase BirA synthesizes biotinyl-5'-AMP and transfers it to the biotin carboxyl carrier protein (BccP). However, when biotin is in excess, BirA can repress its biosynthesis by binding the biotin operator under a homodimer form.

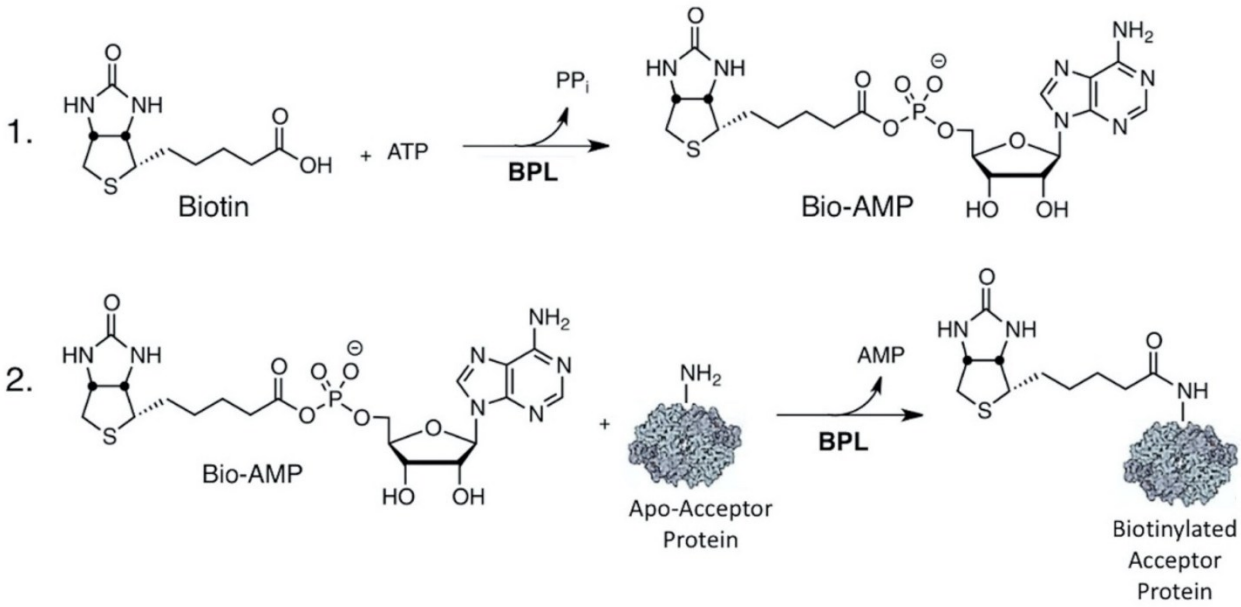


Figure 1.4. Enzymatic biotinylation via biotinyl-5'-AMP
The Biotin Protein Ligase (BPL) catalyzes a two-step reaction mechanism to transfer biotin to the acceptor protein. (1) the carboxyl group of biotin reacts with the α-phosphate of ATP, forming biotinyl-5′-AMP (Bio-AMP), which remains bound to the BPL through hydrogen bonding. (2) the ε-amino group of a lysine residue on the acceptor protein performs a nucleophilic attack on the carbonyl carbon of the biotin moiety in Bio-AMP, resulting in the formation of a stable amide bond and release of AMP. This figure is reproduced from (Cronan, 2024) under a Creative Commons Attribution (CC BY 4.0) license.

The first engineered variant, proximity-dependent biotin identification (BioID), introduced an R118G mutation in the active site of BirA destabilizing its binding to biotinyl-5'-AMP and increasing the dissociation constant by approximately 400 times compared to WT BirA. This weakening of BioID's ability to retain the reactive intermediate allows it to diffuse and react with lysine residues on proximate proteins leading to their biotinylation. Therefore, instead of being exclusive to BccP, BioID biotinylation activity became promiscuous, enabling its use as a proximity labeling tool. Although the potential of the R118G mutation for studying protein interactions was first noted in 2004, the term "BioID" and its formal application as a proximity labeling method were established in 2012 (Choi-Rhee et al., 2004; Roux et al., 2012). While highly effective and widely used especially in mammalian cells, BioID requires 15-18 hours of biotin treatment, limiting its utility for studying dynamic processes (Guo et al., 2023; Kim & Roux, 2016; Roux et al., 2012).

BioID2 was developed as a smaller and more efficient alternative, derived from *Aquifex aeolicus* biotin ligase rather than *E. coli*. It lacks a DNA-binding domain, therefore offering a reduced size (27 kDa vs. 35 kDa), enhanced expression, lower biotin requirements, and more efficiency. Its development was achieved through a specific R40G mutation, which, orthologous to the R118G mutation in BioID, granted it promiscuity (Kim et al., 2016).

TurboID represents a major leap forward, developed using yeast display-based directed evolution of *E. coli* BirA and incorporating 15 mutations, it showed over 20-fold increase in catalytic efficiency compared to BioID and activity at a broader temperature range. This enhancement enables labeling within 10 minutes, making it ideal for capturing transient interactions and dynamic processes. However, its high enzymatic activity might lead to mild cytotoxicity, particularly under longer-than-necessary labeling or in sensitive systems (Branon et al., 2018). A smaller miniTurbo version (28 kDa) was developed in parallel by truncating the N-terminus of TurboID. While slightly less efficient in labeling (1.5-2-fold lower than TurboID), its smaller size is advantageous for minimizing interference with bait protein function. However, miniTurbo has shown lower expression and reduced stability in some systems, likely because of the N-terminal truncation (Branon et al., 2018; Guo et al., 2023).

Recent advances in engineered ligases have significantly expanded the proximity labeling toolkit, with ongoing efforts focused on developing techniques tailored to diverse experimental conditions and biological contexts. However, these newer ligases have not yet achieved the same widespread adoption as BioID, TurboID, or peroxidase-based methods. BASU, derived from *Bacillus subtilis*, was reported to be 1000 times faster than BioID (Ramanathan et al., 2018), but its results were deemed irreproducible by others (Branon et al., 2018). AirID, designed de novo with 82% sequence similarity to BioID, is compact and requires lower ATP and biotin concentrations, with minimal cytotoxicity, especially beneficial in systems where TurboID's high activity may be problematic. However, it still requires hours of labeling time (Kido et al., 2020). UltraID, a truncated version of BioID2 with an L41P mutation, is 25% smaller than BioID2 and 44% smaller than BioID/TurboID yet matches TurboID speed in labeling (10 minutes). While very promising, its applications remain under investigation (Kubitz et al., 2022).

Finally, split ligases like split-TurboID and split-BioID represent a recent adaptation of split reporter systems, a concept that has been in use since the 1990s (Johnsson & Varshavsky, 1994; Pelletier et al., 1998). In these systems, the biotin ligase enzyme is divided into two inactive fragments that reassemble and become functional only when brought into proximity through protein-protein interactions (Cho, Branon, Rajeev, et al., 2020; Cho, Branon, Udeshi, et al., 2020; Ramirez et al., 2021). This design enables selective and controllable protein labeling, activating the enzyme under specific conditions such as when the target protein is in a particular cellular region or near another protein, thus allowing precise temporal regulation and minimizing background interference in dynamic biological studies (Guo et al., 2023).

Selecting the appropriate biotin ligase for a given project requires careful consideration, as each variant presents a unique combination of enzymatic activity and background noise. Highly active ligases can accelerate labeling but may also increase non-specific biotinylation, complicating downstream analysis. Therefore, the optimal choice depends on the specific biological context and experimental goals, with factors such as tag size, labeling radius, toxicity, reaction speed, and overall efficiency playing critical roles in determining suitability. Biotin ligases have been widely utilized across various systems, including mammalian cells, plants, parasites, worms, and viruses (Guo et al., 2023). In contrast, their use in prokaryotic systems has remained largely unexplored, primarily due to challenges posed by the small size of the organism and the rapid kinetics of labeling. An additional concern is that many BLs are derived from bacterial species such as *E. coli*, raising the possibility that the ligase may retain their native enzymatic functions when used in closely related organisms, potentially confounding the interpretation of experimental data. At the time this project was initiated, there were virtually no studies exploring the application of promiscuous BL in prokaryotes. Almost concomitantly with the publication of the main findings of this thesis, two groups reported the use of miniTurbo in bacterial systems. Herfurth et al. utilized miniTurbo in *Myxococcus xanthus* to characterize the interactome of the small GTPase MglA, a key regulator of cell polarity (Herfurth et al., 2023). While Remy et al. applied miniTurbo in *Bdellovibrio bacteriovorus* to investigate the subcellular localization of RomR and DivIVA, the proteins recognized as polar hubs in other bacterial species (Remy et al., 2024).

To our knowledge, this study is among the first to implement proximity labeling in the bacterial cytoplasm using biotin ligases, and notably, has been recognized as the first to apply TurboID in this context (Simoens et al., 2025; Snauwaert et al., 2025; Tesseur et al., 2025). This work addresses a critical gap in the application of proximity labeling in prokaryotic systems, where the small cellular volume and cytoplasmic localization of native biotin ligase substrates present unique challenges. The IpgC-MxiE transcriptional regulatory system was selected as a model because its well-characterized binary protein-protein interactions provide a reliable proof-of-concept for validating cytoplasmic proximity labeling by biotin ligases. Moreover, this system undergoes distinct functional changes between the T3SS “off” and “on” states, making it particularly suitable for probing state-dependent interactomes. Although well studied, the proxisomes of IpgC and MxiE remain incompletely mapped; IpgC is highly abundant, and a substantial fraction is not fully titrated by MxiE during the T3SS “on’’ state, suggesting additional interactions may exist.

By systematically evaluating multiple biotin ligases, we identified TurboID as the most effective tool for cytoplasmic labeling in *S. flexneri*. Its use expanded the IpgC proxisome beyond its established roles, revealing potential interactions with transcriptional regulators and core cellular components that may participate in broader signaling networks. Together, these results demonstrate the feasibility and value of using promiscuous biotin ligases in bacterial systems. Ultimately, as antimicrobial resistance in *Shigella* continues to rise, expanding the mechanistic understanding of the T3SS, its central virulence determinant and the primary driver of host cell invasion, may guide future efforts toward the development of targeted pharmacological strategies to counter *Shigella* infection.

Chapter 2 : Optimizing the Use of Proximity Labeling Tools in the Cytoplasm of *Shigella Flexneri*

2.1 Preface

This chapter includes results and figures (Figures 2.9 and 2.10) adapted and edited from a study published in *mSphere* (Haidar-Ahmad et al., 2024). The experiments and data analysis presented in this chapter were performed by me as part of my thesis work. As an author of the article, I retain the rights to reuse this content in my thesis, in accordance with the journal's policy and the Creative Commons Attribution License (CC BY 4.0), which permits reuse provided the original published version is properly cited. I acknowledge the contributions of Tyler Do in characterizing some of the BioID2 single mutants described in this chapter.

2.2 Introduction

To investigate the molecular organization and interactions that govern key bacterial processes, particularly those involved in pathogenesis, proximity labeling has emerged as a powerful tool for mapping proxisomes within living cells. However, the complexity of the bacterial cytoplasm presents unique challenges, including concerns over the size and activity of biotin ligases, as well as their efficiency in labeling proteins of interest without causing cellular disruption. To address these challenges, we sought to identify the most suitable biotin ligase for proximity labeling in *Shigella flexneri* cytoplasm. In the following sections, we assess the activity of four well-established biotin ligases to determine which one offers the highest specificity and sensitivity for protein labeling in *Shigella*, while minimizing background labeling or cross-reactivity, ultimately guiding our choice of method for downstream studies in *Shigella flexneri* cytoplasm.

2.3 Results

2.3.1 BioID2 mutants showed an increase in activity

BioID2, a biotin ligase (BL) from *Aquifex aeolicus*, was the main ligase that we considered at the beginning of the project due to its compact size that can be advantageous to study large protein complexes in the T3SS. However, due to the lack of studies at the time exploring the use of biotin ligases in the bacterial cytoplasm, limited information was available regarding their behavior in this context, therefore further investigation was needed before applying this technique in bacteria. One key issue was the potential binding of BioID2 to BccP, the natural ligand of its homolog BirA in *E. coli* and *S. flexneri*. This concern stemmed from observations in the original study characterizing the R118G BirA mutant, which later developed into BioID (Choi-Rhee et al., 2004). In that study, the mutant retained a strong preference for binding BccP, its natural ligand in *E. coli*, over other proteins, raising the possibility that similar behavior could occur with BioID2 in the bacterial cytoplasm. Such preferential binding could interfere with the biotinylation of proteins associated with our selected bait, reducing the efficiency of BioID2 in detecting proximal proteins and increasing the likelihood of false positives.

To disrupt BioID2 binding to its natural ligand and thus, increase its bioavailability in the cell, we decided to design mutations at the predicted interface between BioID2 and BccP. Since no existing model for their interaction was available, we used the X-ray crystal structure of a BirA homolog from *Pyrococcus horikoshii* bound to BccP (PDB: 2EJG). Sequence alignment using Clustal Omega revealed a 34.53% sequence identity between the biotin ligases from *Pyrococcus horikoshii* and *Aquifex aeolicus* (PDB: 2EAY) (Figure 2.1A). Furthermore, the 3D structures of both ligases showed a high structural similarity at the BccP binding domain interface, leading us to believe that indeed, the biotin ligase from *Aquifex aeolicus* could be binding to this ligand in *S. flexneri*. The amino acid sequence alignment, along with the available protein structure, enabled us to map previously reported key residues involved in stabilizing the BL-BccP complex in *P. horikoshii* (Figure 2.1B) (Bagautdinov et al., 2008). These residues then guided our identification of the corresponding positions in BioID2 (Figure 2.1C), which were subsequently substituted with alanine. While the primary goal of introducing these mutations was to disrupt BccP binding, two of them, L66A and L67A, were also found to enhance BioID2

activity based on unpublished data from our lab involving the use of a BioID2 fusion with the T3SS protein IpgC, prompting us to select BioID2 L66A L67A as the starting point of this mutagenesis campaign.

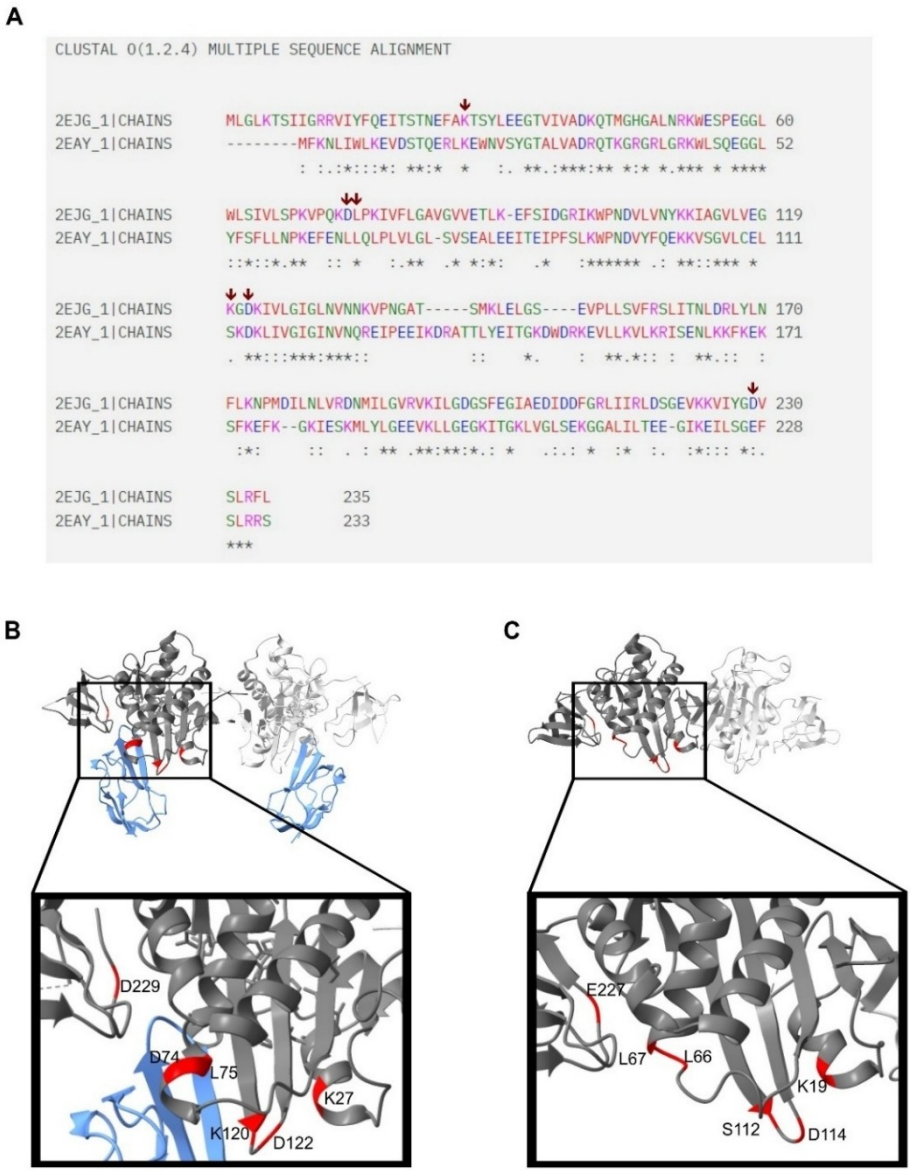


Figure 2.1. The sequence alignment and X-ray crystal structures that were used to select mutations at the BL-BccP interface.
(A) Sequence alignment of the biotin ligase from *Pyrococcus horikoshii* (PDB: 2EJG) and the biotin ligase from *Aquifex aeolicus* (PDB: 2EAY) with key residues marked by an arrow. Predicted structures of (B) 2EJG (235aa) and (C) 2EAY (233aa). The biotin ligase dimer is shown in grey, BccP in blue, and the BL residues predicted to stabilize the BL-BccP complex are highlighted in red.

The selected mutations on the BioID2-BccP predicted binding interface were tested individually in the BioID2 L66A L67A construct to assess their effect on the activity of the enzyme and verify that we did not compromise the activity in our attempt to increase the bioavailability. We expressed the BioID2 mutants in the cytoplasm of *Shigella flexneri* str. M90T and incubated them with biotin for 16 hours, in conditions similar to those used for wild type BioID2 in mammalian cells (Kim et al., 2016). Streptavidin-horseradish peroxidase was used to evaluate activity, while the Myc tag was used to monitor expression (Figure 2.2). The results showed variations in activity levels, while D114A and E227A mutations preserved the high activity of BioID2 L66A L67A compared to the original BioID2, S112A and K19A appeared to sabotage this advantage leading us to discard them.

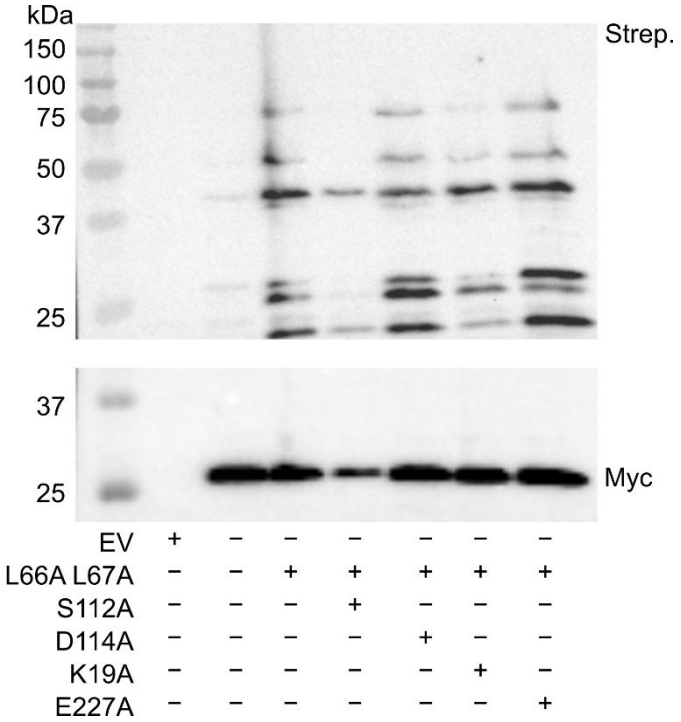


Figure 2.2. BioID2 mutants show different levels of activity.
Biotinylation patterns of the different BioID2 mutants in *S. flexneri* strain M90T after incubation with 50 µM biotin for 16 h. pSU2.1 was used as the backbone vector. Expression of the BioID2 variants was assessed by immunoblotting using a monoclonal antibody against the Myc tag, and biotinylation blots were performed with streptavidin-horseradish peroxidase (Strep.).

Using densitometry on three biological replicates, BioID2 promiscuous variant with alanine mutations at positions 66, 67, 114 and 227 showed approximately 6.4 ± 7.5 times

increase in activity compared to the original BioID2 (Figure 2.3). The L66A L67A mutant alone increased activity by 2.3 ± 3.1 times. The large standard deviation observed in the biotinylation quantification likely reflects variability introduced by signal masking, where intense bands in the BioID2 mutants interfered with accurate detection of faint bands in WT BioID2 (lane 2) across replicates, leading to inconsistent quantification.To demonstrate that this variability is specific to the WT BioID2 signal and not due to general data inconsistency, we compared the BioID2 L66A L67A mutant with the quadruple mutant. This comparison yielded a 2.4 ± 1.5-fold increase in activity, with a notably lower standard deviation. This supports the interpretation that the high variability in WT comparisons arises from signal masking rather than experimental noise and confirms the reproducibility of the mutants' performance.

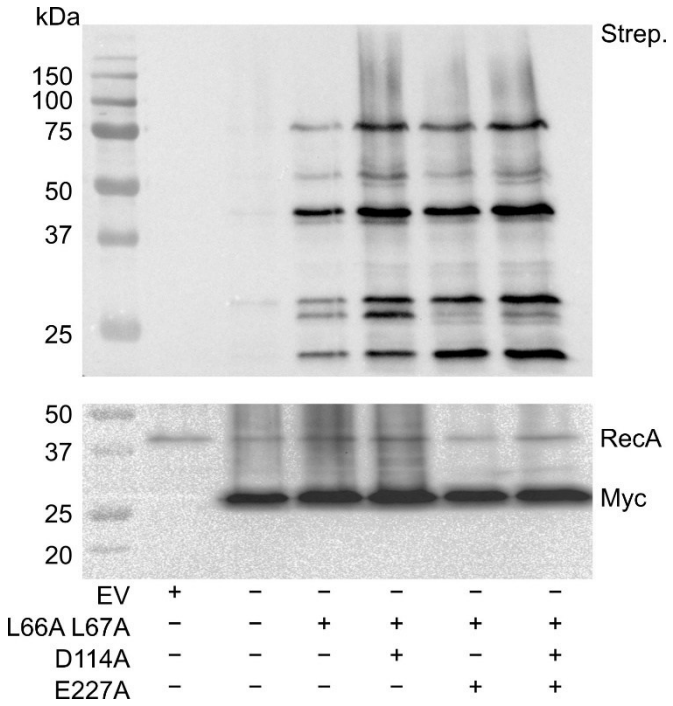


Figure 2.3. BioID2 quadruple mutant maintained a high level of activity.
Biotinylation patterns of the different BioID2 mutants in *S. flexneri* strain M90T after incubation with 50 µM biotin for 16 h. pSU2.1 was used as the backbone vector. Expression of the BioID2 variants was assessed by immunoblotting using a monoclonal antibody against the Myc tag, with RecA as a loading control, and biotinylation blots were performed with streptavidin-horseradish peroxidase.

It is worth noting that when we later tried BioID2 mutants to study the T3SS chaperone IpgC, the L67A mutation was tested separately, and in combination with K115A (one of the single mutants that showed an increase in BioID2 activity based on unpublished data from Tyler Do) or L66A (Figure 2.4), the L66A L67A mutant did not show a visible advantage in biotinylating activity over L67A indicating that L67A is sufficient for increasing the activity. However, we decided to keep both, L66A and L67A, since both mutations are located on the BccP binding interface and thus could contribute to disrupting it. What was unexpected is the decrease of biotinylation activity in the L67A-K115A mutant compared to L67A alone, even though each of these single mutations by themselves have been shown to enhance BioID2 activity.

In conclusion, the quadruple mutant designed to break potential BioID2 binding to the natural ligand BccP not only retained its biotinylation activity but increased it by 6.4-fold. Next, we turned our attention to investigating if these mutations in BioID2 indeed helped to disrupt its binding to BccP.

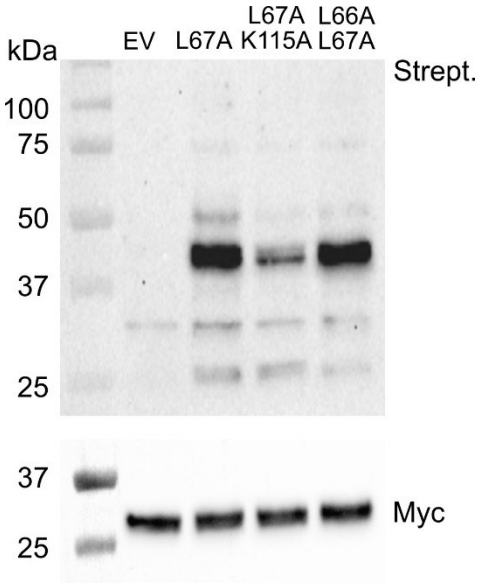


Figure 2.4. L67A mutation is the leading cause of increased BioID2 activity.
Biotinylation patterns of IpgC-BioID2 mutants in *S. flexneri* Δ*ipgC* strain after incubation with 50 µM biotin for 16 h. pSU2.1 was used as the backbone vector. Expression of the IpgC fusions with BioID2 variants was assessed by immunoblotting using a monoclonal antibody against the Myc tag, and biotinylation blots were performed with streptavidin-horseradish peroxidase.

2.3.2 A novel biotin ligase TurboID emerged, overthrowing the BioID2 quadruple mutant

While working on BioID2, a new promiscuous biotin ligase named TurboID was developed (Branon et al., 2018). Like the popular promiscuous biotin ligase BioID, TurboID originated from the *E. coli* BirA biotin ligase. However, TurboID's 15 mutations resulted in a higher activity in mammalian cells, and decreased the biotin incubation time needed to obtain a detectable signal from 16 hours to minutes (Branon et al., 2018). This new finding inspired us to include TurboID in our study and characterize it alongside BioID2 in the bacterial cytoplasm.

To compare BccP binding ability of each of the enzymes *in vitro*, the expression of BioID2, TurboID and its ancestor BioID, and BccP the natural substrate of BirA, was carried out using pQE-80L derived plasmids harboring 6xHis-biotin ligase and *MBP-bccP* introduced in *E. coli* BL21 Rosetta cells. Induction with IPTG successfully yielded high levels of the target proteins, as confirmed by Coomassie gel analysis (Figure 2.5). Importantly, a substantial portion of the expressed proteins were recovered in the soluble fraction, which was essential for downstream purification. This result indicated that the conditions used were effective in promoting soluble protein production although some proteins were detected in the insoluble fraction. Since BccP can be naturally found in a biotinylated form holo BccP in *E. coli* cytoplasm, we tried to remove as much of its biotinylated form from the purified sample to reduce the background biotinylation and we found that growing the cultures in 2xYT media coupled with an additional purification step using streptavidin-agarose beads was able to reduce the amount of biotinylated BccP in the total purified protein sample (Figure 2.6A).

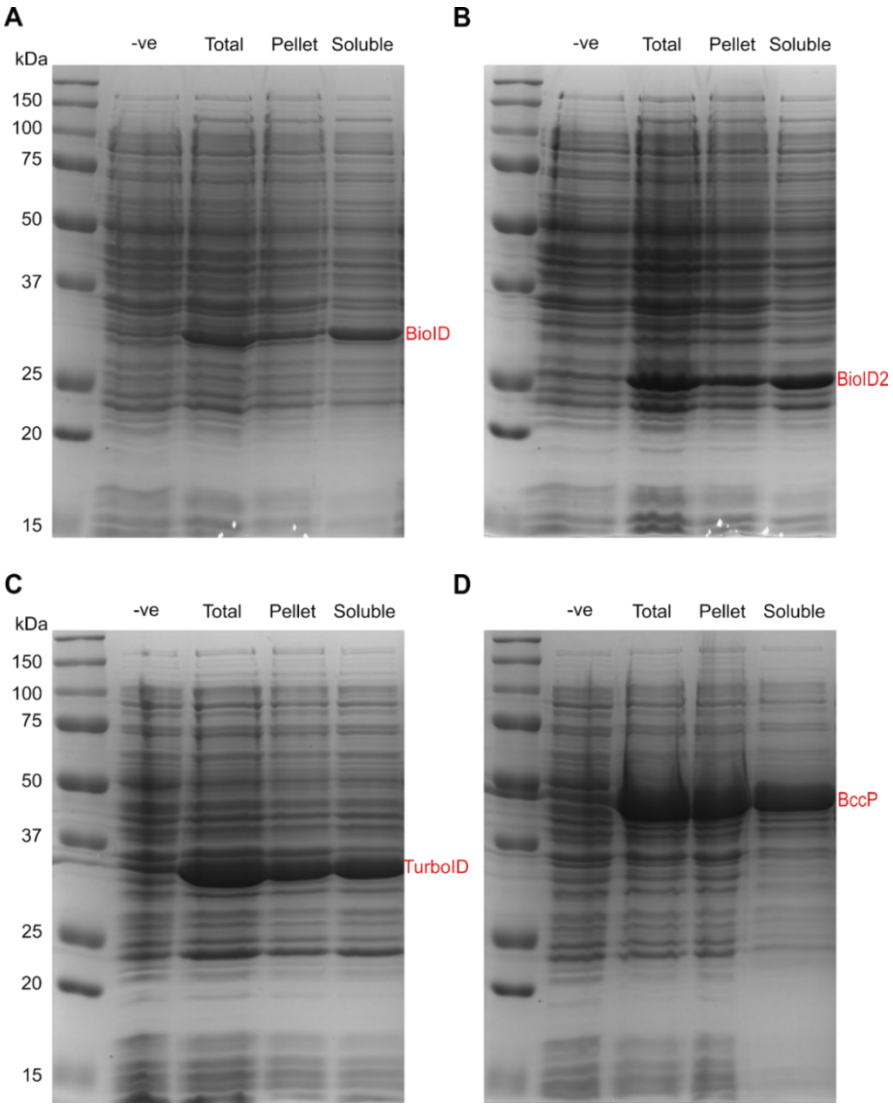


Figure 2.5. Coomassie-stained SDS-PAGE analysis of the expression and solubility of the biotin ligases and BccP.
Coomassie-stained SDS-PAGE gels showing the expression and solubility of the four proteins expressed in *E. coli* BL21 Rosetta cells. BioID (A), BioID2 (B), and TurboID (C) were expressed using pQE-80L plasmid, with a 6xHis tag, while BccP (D) was expressed using MBP tag. For each protein, lanes represent total cell lysate before IPTG induction, total cell lysate after IPTG induction, insoluble fraction after cell lysis and centrifugation, and soluble fraction after cell lysis and centrifugation. The labeled bands represent the protein fused to its tag.

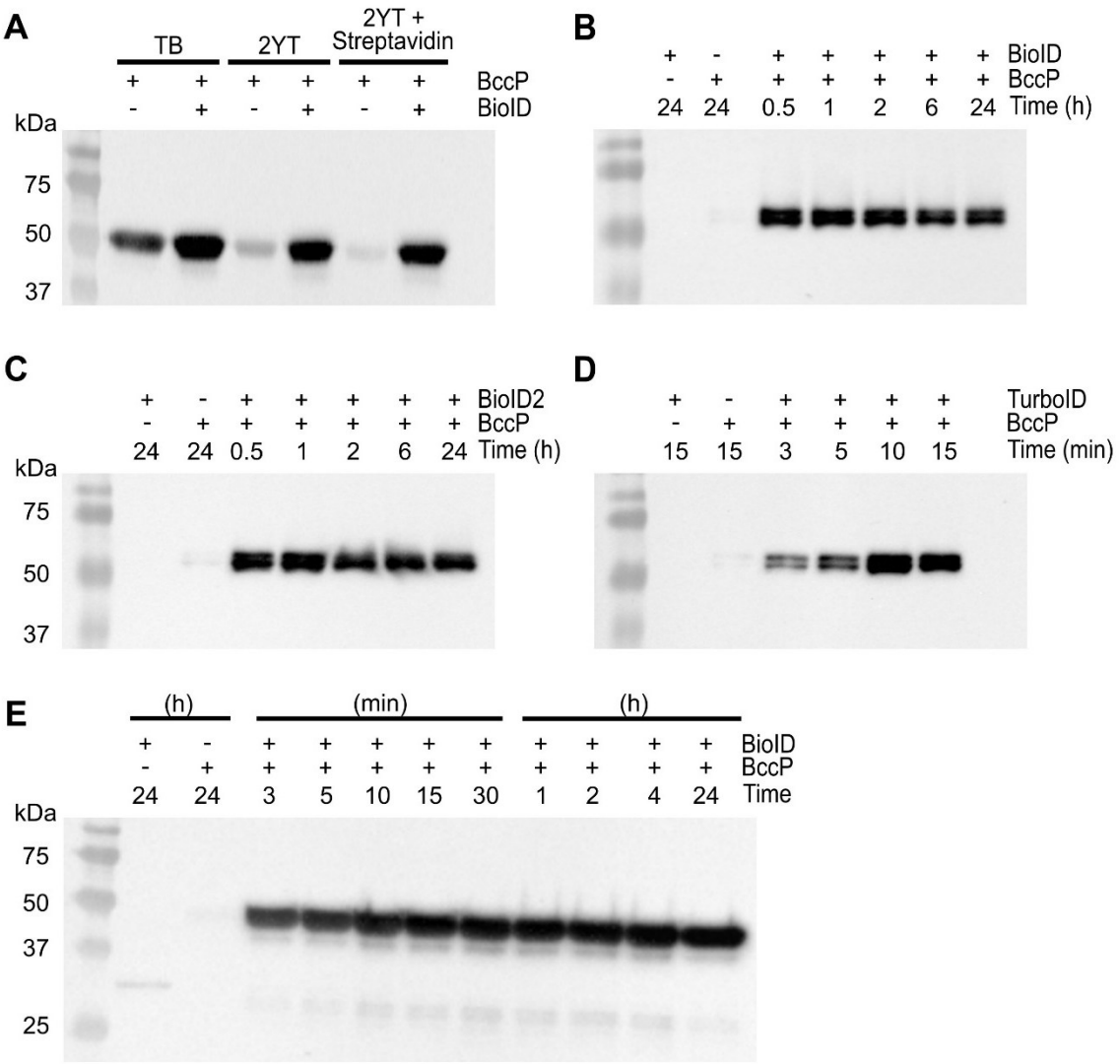


Figure 2.6. Optimization and time-course analysis of MBP-BccP biotinylation by 6×His-tagged biotin ligases.
This figure illustrates the optimization and time-course analysis of MBP-BccP biotinylation using streptavidin conjugated to horseradish peroxidase (HRP) for detection. (A) Optimization of media type and additional incubation step with streptavidin beads to minimize background noise from endogenously biotinylated BccP. (B, C, D) Time-course analysis of MBP-BccP biotinylation catalyzed by the different biotin ligases, with signal intensity corresponding to the extent of biotinylation at each time point. (E) Expanded time-course analysis with shorter time intervals. Controls without enzyme were included in all assays to confirm the absence of nonspecific biotinylation.

After the purification of proteins, an *in vitro* biotinylation optimization assay for each of the biotin ligases was performed with BccP. We chose a time range inspired by the activity time of each ligase when used *in vivo*; over the course of 24 hours for BioID and BioID2, and up to 15 minutes for TurboID. As expected, we observed an increase in biotinylation over time for TurboID (Figure 2.6D). However, BccP biotinylation reached saturation much faster than the recommended 16 hours with the other ligases (Figure 2.6 B-C). This made us aware of the difference between the bystander activity of the ligases, which can take hours, and their specific activity towards BccP, which is much more efficient. Therefore, we expanded the time range to include the time points we had used for TurboID to better resolve biotinylation kinetics (Figure 2.6E).

Following the optimization of the *in vitro* BccP binding assay, we conducted the experiment using BioID2 and its quadruple mutant side-by-side. Based on our hypothesis, the four mutations introduced at the predicted BccP binding interface were expected to hinder binding, resulting in slower BccP biotinylation compared with the WT. Contrary to our expectations, BioID2 quadruple mutant exhibited a similar rate of BccP labeling as the WT (Figure 2.7A). To investigate if this result was related to the overall increased activity of the quadruple mutant as demonstrated before (Figure 2.3) or if it was specific to BccP, the same experiment was repeated using RNase A, a protein unrelated to the BirA-BccP and the biotin biosynthesis pathway. The time frame required to detect biotinylation of the RNase A took hours, thus exceeding what is needed for the biotinylation of BccP. Importantly, no noticeable change was seen when comparing the activity of the two enzymes except a reduction in the autobiotinylation in the BioID2 quadruple mutant (Figure 2.7B).

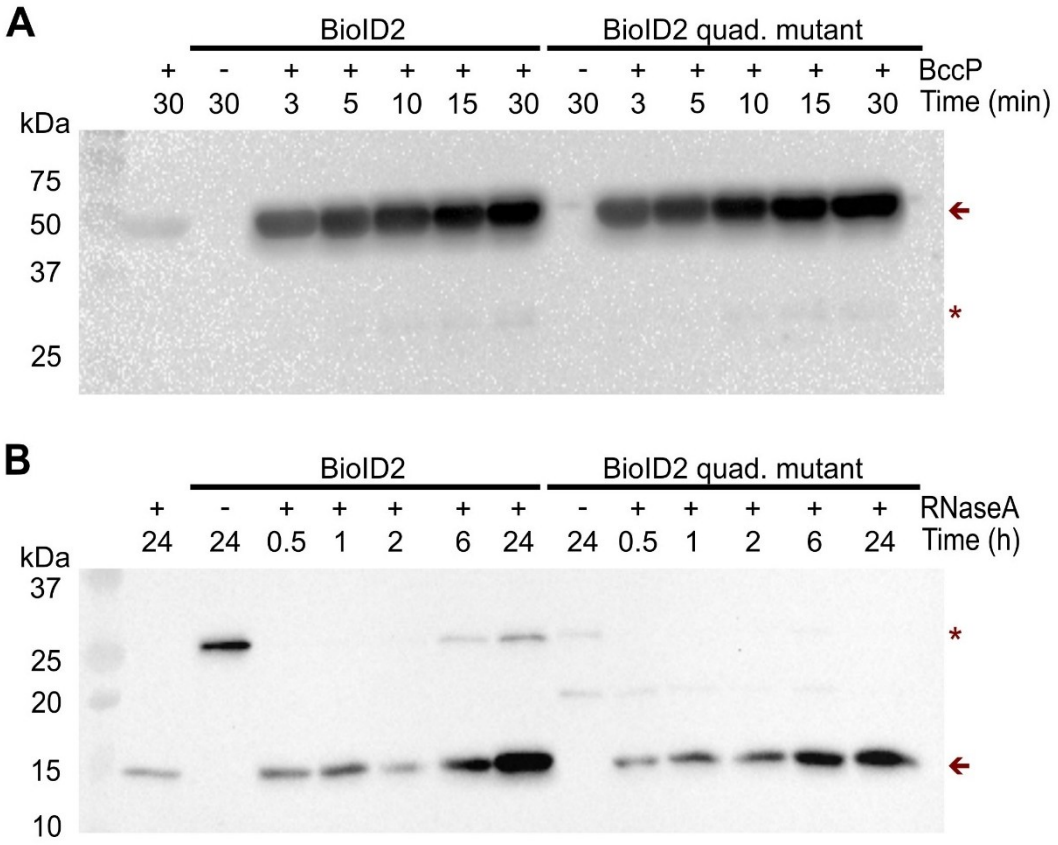


Figure 2.7. BioID2 quadruple mutant does not show any striking difference in biotinylating BccP *in vitro* compared to WT BioID2.
Time course of the biotinylation of MBP-BccP (A) and RNase A (B) by 6×His-tagged WT BioID2 and BioID2 quadruple mutant (L66A L67A D114A E227A) as measured with the streptavidin-horseradish peroxidase. MBP-BccP and RNase A are indicated by the arrow in each of the panels. Autobiotinylation of the ligases is marked with an asterisk, while unlabeled bands represent contaminants.

The *in vitro* results in figure 2.7 led us to question whether the mutations we generated had any advantage other than increasing BioID2 activity (Figure 2.3). However, with TurboID promising a high activity in a shorter time, which is important for studying fast mechanisms that cannot afford the 16 hours labeling time, we had to put BioID2 quadruple mutant to the test and check if the increase in biotinylation activity it has over the original BioID2 enables it to compete with TurboID in the bacterial cytoplasm. The biotinylation assay showed the superiority of TurboID over the BioID2 quadruple mutant at 10 minutes (Figure 2.8) prompting us to pursue TurboID for studying protein complexes in the bacterial cytoplasm.

Although we abandoned our campaign to engineer BioID2 at this point and moved on with

TurboID, BioID2 quadruple mutant has shown an improvement in activity over the original enzyme and is worth pursuing in future studies. Each promiscuous biotin ligase has its own advantages that can make them the enzyme of choice in function of the context and system studied. Nevertheless, a highly promiscuous and efficient enzyme like TurboID seems to be the most appropriate to incorporate into our study.

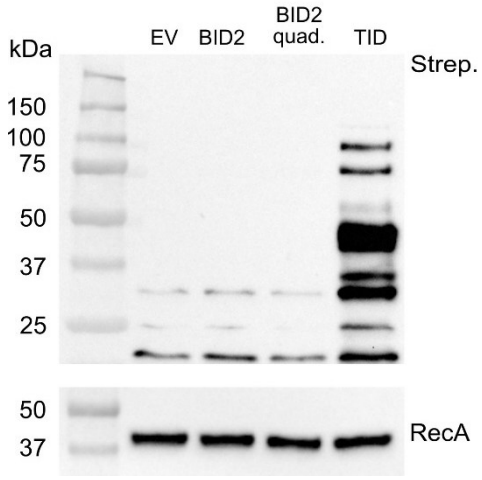


Figure 2.8. TurboID demonstrates superior promiscuous biotinylation activity compared to BioID2 quadruple mutant.
Biotinylation induced by the indicated BioID2 (BID2) variants and TurboID (TID) was measured after 10 min of the addition of 50 µM biotin to cultures of WT *S. flexneri* str. M90T. The quadruple BioID2 mutant is L66A L67A D114A E227A. pSU2.1 was used as the backbone vector. Biotinylation was detected using streptavidin-horseradish peroxidase, and RecA served as a loading control.

2.3.3 TurboID steals the show outperforming other biotin ligases in the cytoplasm of *Shigella flexneri*

The emergence of TurboID prompted us to explore our options and assess its activity in comparison with other widely used biotin ligases, including BioID, BioID2, as well as miniTurbo, a novel BirA-derived ligase developed at the same time as TurboID. This evaluation was essential to identify the most suitable tool for our study. We expressed the promiscuous biotin ligases in the cytoplasm of *Shigella flexneri* and analyzed their biotinylation activity at three-time points; 10 minutes, 1.5 hours, and 16 hours after supplementing the medium with 50 µM biotin. We evaluated the activity by measuring the biotinylation using a streptavidin-

horseradish peroxidase (HRP), and the expression via a Myc tag antibody (Figure 2.9A). TurboID demonstrated an impressively higher biotinylation activity than the other ligases at all time points, followed by BioID and miniTurbo. Although BioID2 showed activity in *S. flexneri* cytoplasm when tested alone as shown before, its signal got masked by that of the other biotin ligases even at 16 h. A densitometric analysis at the 10-minute time point using three independent biological replicates allowed us to compare the activity of each of the ligases by quantifying the overall signal in each lane and showed that TurboID's biotinylation activity is around 20 times higher than the second most active ligase BioID (Figure 2.9B). It was previously reported that the activity of TurboID and miniTurbo is very close, varying around only 1.5-fold. To verify this, we repeated additional experiments with TurboID and miniTurbo at the 10 min time point using three new biological replicates, probing all samples on a single membrane to minimize technical variability (Figure 2.9C) and we concluded using densitometry that TurboID exhibited biotinylation activity 13.2 ± 2.8 times greater than miniTurbo but only because miniTurbo was much less expressed in the cells. Hence, if we take the expression levels into account, TurboID would be 1.6 ± 0.5 times more active than miniTurbo, aligning with previous studies conducted in eukaryotic cells (Branon et al., 2018). The main difference between miniTurbo and the other BirA derived enzymes, BioID and TurboID, is the lack of the conserved N-terminal domain, which might contribute to its reduced expression and poor performance in *S. flexneri* at 37°C. Based on these observations, BioID and TurboID were identified as the most promising candidates for protein studies in the *S. flexneri* cytoplasm.

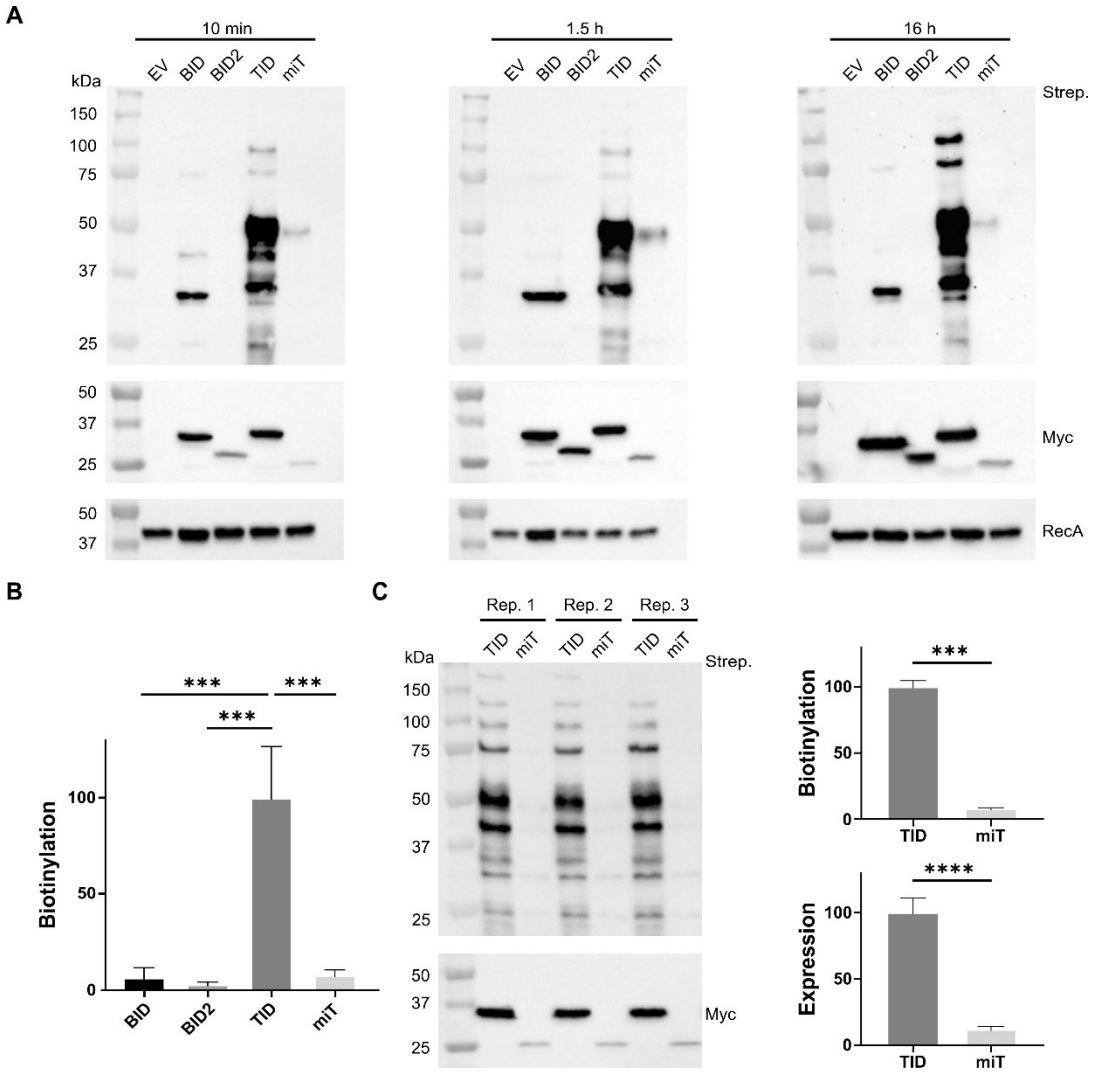


Figure 2.9. TurboID demonstrates superior promiscuous biotinylation activity in the cytoplasm of *Shigella flexneri*.

(A) Biotinylation activity of promiscuous biotin ligases was assessed in *S. flexneri* strain M90T carrying constructs for BioID-Myc (BID), BioID2-Myc (BID2), TurboID-Myc (TID), miniTurbo-Myc (miT), or an empty vector control (EV), all expressed from the pSU2.1 backbone vector. Cultures were supplemented with 50 µM biotin, and samples were collected at 10 min, 1.5 hours, and 16 hours. Biotinylation was detected using streptavidin-horseradish peroxidase, while the expression was measured by immunoblotting with an anti-Myc monoclonal antibody. RecA served as the loading control. (B) Densitometric analysis of biotinylation blots at the 10-minute time point from three independent biological replicates (n = 3) was used to compare the relative activity of the biotin ligases with statistical significance calculated using one-way analysis of variance and Tukey's post hoc test (***, $P < 0.001$). Error bars represent the standard deviation of the mean. (C) Comparative biotinylation activity and expression levels of TurboID and miniTurbo at the 10-minute time point were analyzed in a new set of three biological replicates (n = 3), labeled Rep. 1, 2, and 3. Statistical significance was calculated using Student's t-test for unpaired data with equal variance and a 95% CI (***, $P < 0.001$; ****, $P < 0.0001$). Error bars represent the standard deviation of the mean.

2.3.4 TurboID evolved from its ancestors, losing some of its native functions

TurboID and BioID, derived from the native *E. coli* BirA biotin ligase, are designed to catalyze promiscuous biotinylation of proximal proteins. While this feature makes them powerful tools for mapping protein interactions, their functionality in the bacterial cytoplasm may be influenced by the cellular functions of their ancestral protein. In *S. flexneri*, BirA and BccP are identical to their *E. coli* K-12 orthologs, raising the possibility that BccP could be still interacting with these BirA derivatives, therefore limiting their utility for targeted proxisome studies. Another function BirA has in the cell is the regulation of the biotin synthesis operon by binding the biotin operator (*bioO*), which can also be an issue if TurboID or BioID are attracted to DNA and relocating their bait protein away from its natural binding partners. To address these concerns, we investigated whether TurboID and BioID retained their ability to fulfill native BirA functions of biotinylating the natural substrate BccP and binding the biotin operator.

Biotinylation of BccP is a well-characterized function of BirA and requires the stable association of the intermediate biotin-5′-AMP with the enzyme's active site. BioID and TurboID, optimized for promiscuous biotinylation, are engineered to release biotin-5′-AMP readily, potentially reducing their specificity for native substrates like BccP. To test this hypothesis, we purified BioID, TurboID, and BccP and performed *in vitro* biotinylation assays. The data revealed that TurboID exhibited a significantly slower biotinylation rate for BccP compared to BioID, with 14.6 min and 3.4 min as midpoint of biotinylation (t^{50}) respectively (Figure 2.10A). This reduced efficiency is consistent with TurboID's design as a more powerful promiscuous enzyme, emphasizing its divergence from BirA's substrate-specific activity. In contrast, when RNase A, a protein not associated with BirA, was used as substrate, both ligases showed much slower labeling and yielded similar values of midpoint of biotinylation (t^{50}), with 4.7 hours for TurboID and 3.4 hours for BioID (Figure 2.10B). To determine whether this outcome was specific to RNase A or reflected a more general effect on bystander proteins, we repeated the assay with bovine serum albumin (BSA). Like RNase A, BSA required hours to become labeled by both enzymes, with no noticeable difference between TurboID and BioID (Figure 2.11). Together, these results indicate that BioID retains an advantage over TurboID for its native substrate BccP, whereas both ligases behave similarly when acting on unrelated proteins. This outcome reflects TurboID's mutations, which favor promiscuous biotinylation at the expense of

activity toward its native substrate, thereby limiting interference with endogenous systems in *S. flexneri*.

The second critical function of BirA is its ability to bind the biotin operator under conditions of excess biotin. This activity is mediated by the homodimerization of BirA, forming a stable complex that interacts with DNA. TurboID and BioID were expected to exhibit reduced binding to the biotin operator due to their fast release of the intermediate biotinyl-5′-AMP, particularly in more promiscuous TurboID. We conducted an electrophoretic mobility shift assay (EMSA) to compare the binding efficiencies of BirA, BioID, and TurboID to a biotin operator DNA probe in the presence of excess biotin and ATP (Figure 2.10C). At 50 nM, BirA demonstrated robust binding, shifting 93% of the DNA probe, whereas BioID and TurboID exhibited noticeably lower affinities, shifting 41% and 0.8%, respectively. Even at 250 nM, TurboID was unable to completely shift the probe, highlighting its reduced DNA-binding ability. This reduction in affinity can be attributed to the high promiscuity of TurboID, which affects its ability to form a complex with the biotin operator.

These results collectively indicate that TurboID has undergone significant functional divergence from its BirA ancestor. Its reduced ability to biotinylate native substrates and diminished biotin operator binding demonstrate a loss of specificity for BirA's endogenous roles which is advantageous for experimental applications in *S. flexneri*. Consequently, TurboID emerges as a superior candidate for studying protein complexes in the cytoplasm of *S. flexneri* and other related Enterobacteriaceae, where its promiscuity can be harnessed for precise and efficient biotinylation.

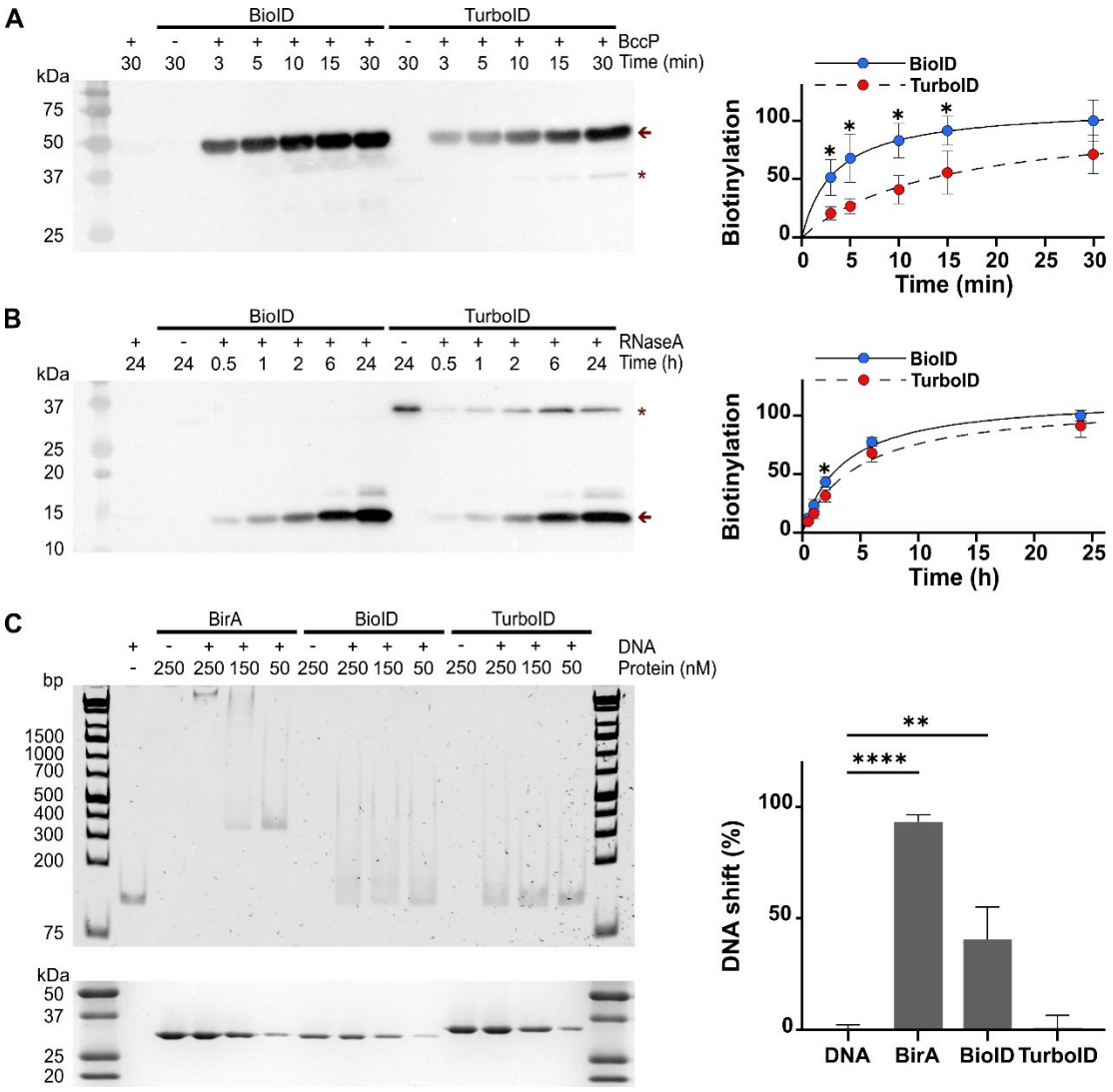


Figure 2.10. TurboID exhibits diminished native functions compared to its ancestor BirA. (A) Time-course analysis of the biotinylation of MBP-BccP by 6×His-tagged BioID and TurboID. MBP-BccP is indicated by the arrow. Autobiotinylation of the ligases is marked with an asterisk (*), while unlabeled bands represent contaminants. Densitometry from three biological replicates (n = 3) of the biotinylation blot was used to quantify the rate of BccP biotinylation, generating the time-course graph (right panel). Data fitting into a rectangular hyperbola revealed the midpoint of biotinylation (t^{50}) for BioID (3.4 min, R = 0.99) compared to TurboID (14.6 min, R = 0.99). Statistical significance was calculated using Student's t-tests for unpaired data with equal variance and a 95% confidence interval (*, $P < 0.05$). Error bars represent the standard deviation of the mean. (B) Time-course analysis of RNase A biotinylation by 6×His-tagged BioID and TurboID. The RNase A protein is marked by the arrow, with the asterisk (*) and unlabeled bands representing autobiotinylation and contaminants, respectively. Densitometry of three biological replicates (n = 3) was used to generate the time-course graph (right panel) and obtain t^{50} values for BioID (3.4 hours, R = 0.99) and TurboID (4.7 hours, R = 0.99). Statistical analysis was performed as in panel A (*, $P < 0.05$). Error bars indicate the standard deviation of the mean. (C) Electrophoretic mobility shift assay (EMSA) assessing the

DNA-binding ability of BirA, BioID, and TurboID to a DNA probe containing the biotin operator. DNA-protein complexes were tested at decreasing concentrations of ligase (250, 150, and 50 nM) and 40 nM of a 112 bp DNA probe in the presence of excess biotin (1 μM) and ATP (1 mM). Coomassie-stained SDS-PAGE gel (bottom panel) confirmed equal protein input. Quantification of DNA probe shifting at 50 nM ligase from three biological replicates (n = 3) showed significantly reduced binding for BioID (41% shift) and TurboID (0.8% shift) compared to BirA (93% shift). Statistical significance was determined using one-way analysis of variance with Dunnett's post hoc test (**, $P < 0.01$; ****, $P < 0.0001$). Error bars represent the standard deviation of the mean

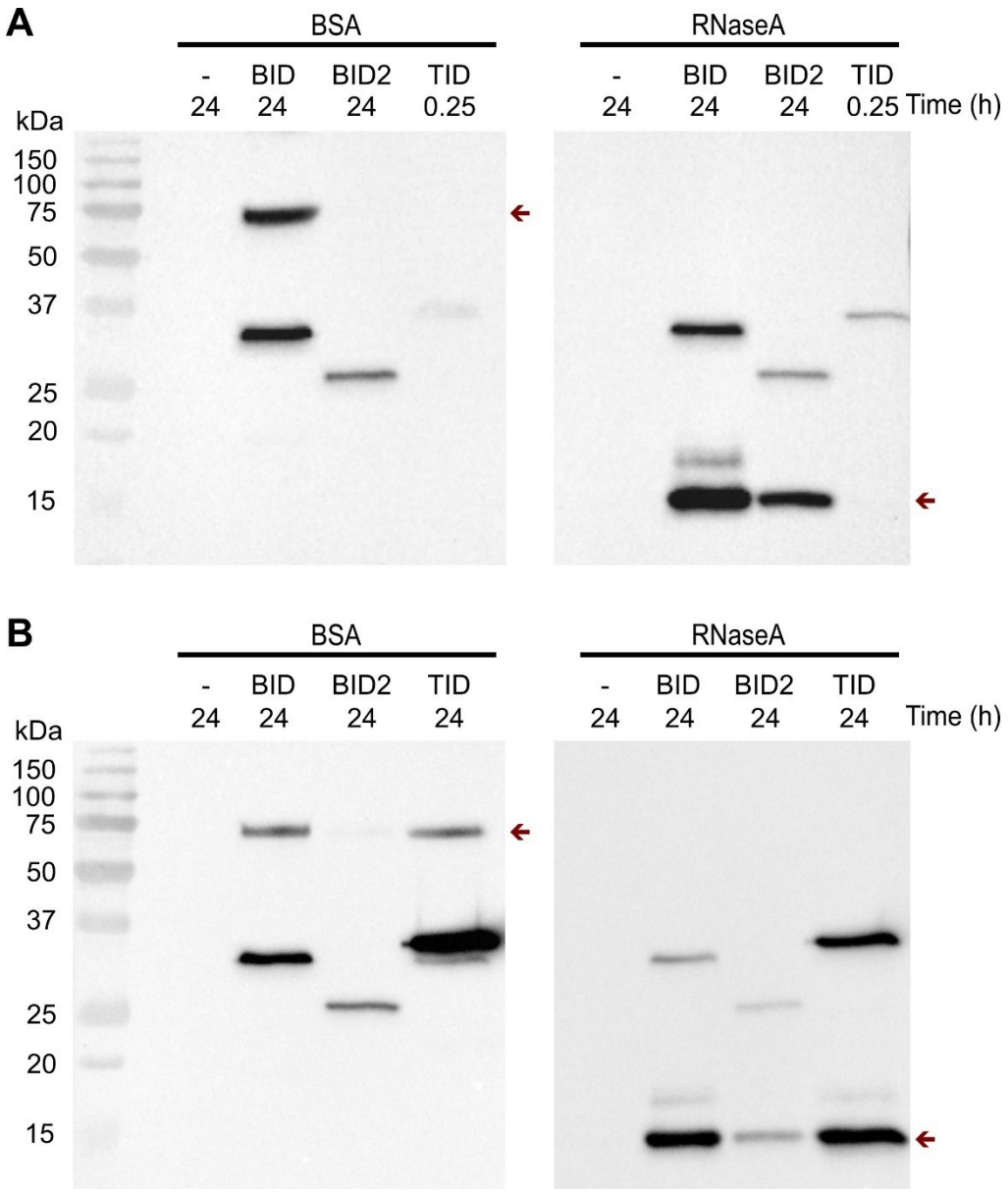


Figure 2.11. BSA and RNase A *in vitro* biotinylation by the different biotin ligases
BSA and RNase A were incubated with the 6×His-tagged ligases at (A) 24 h for BioID (BID) and BioID2 (BID2), and 15 min for TurboID (TID). (B) 24 h for all the biotin ligases. BSA and RNase A are indicated by the arrow. Autobiotinylation of the biotin ligases is visible as bands between 25 and 37 kDa, while the remaining bands represent contaminants.

2.4 Discussion

Proximity-dependent biotinylation has become an essential tool in proteomics and an area of continuous development. Its popularity stems from its ability to detect vicinal proteins and transient events which are valuable especially in host-pathogen studies, while exerting a limited impact on the cell health and function as we have observed in *Shigella*. This method utilises biotin ligases to covalently attach a biotin to proteins near the bait protein and then capture them by affinity chromatography and mass spectrometry, mapping the proxisome and eventually the interactome of the bait protein. In this project, we initially selected BioID2 as our biotin ligase of choice to study the T3SS of *Shigella*. This decision was driven by several key factors, primarily how promising BioID2 seemed at the time as a new ligase offering enhanced proximity labeling as well as its reduced size compared to the original BioID (~25 kDa instead of 35 kDa) which makes it less likely to disrupt the cellular functions or the localization of the bait protein fused to it (Kim et al., 2016). Furthermore, BioID2 originates from *Aquifex aeolicus*, a bacterium that is phylogenetically distant from *E. coli* and *Shigella*, in contrast to BioID.

The implementation of BioID2 or any biotin ligase in the bacterial cytoplasm presents many challenges that would not be problematic in their applications in other systems. First, the bacterial cytoplasm is characterized by a smaller cellular volume, minimal compartmentalization, and protein concentrations that can be up to six times higher than those found in eukaryotic cells (Elowitz et al., 1999; Finka & Goloubinoff, 2013). Such conditions contribute to increased protein crowding and a higher risk of background signal when proximity labeling is used. In addition to that, due to the similarity of this experimental setting to the enzyme's natural environment, there is a possibility of the ligase getting caught up in its native functions in the cell such as binding the biotin carboxyl carrier protein (BccP), the natural ligand of its homolog BirA in *E. coli* and *S. flexneri*, as well as binding DNA. The scarcity of studies in this context raised concerns about how BioID2 would behave in such situations and whether it could be adapted for use in the cytoplasm of *Shigella*.

Since no BioID2-BccP 3D model exists, we had to compare BioID2 to other ligases with an experimentally determined structure in complex with BccP. Sequence alignment of BioID2 with its homolog from *Pyrococcus horikoshii* (PhBPL) using Clustal Omega revealed 34.5% identity. Both proteins showed high structural similarity, particularly in regions associated with

binding BccP (figure 2.1). The structural conservation supports the idea that BioID2 might retain a functional BccP-binding region, which led us to design mutations on that interface to destabilize the binding. After identifying the key residues in PhBPL binding BccP based on the X-ray crystal structure (Bagautdinov et al., 2008), we chose their counterparts in BioID2. This selection aimed to weaken or abolish the interactions between BioID2 and BccP without significantly altering the overall stability of the enzyme (Figure 2.1B), ensuring that the catalytic domain remains intact. The first round of mutations involved two leucines 66 and 67 of BioID2, these residues are positioned in a turn near the predicted docking site of Lys115 of BccP (the conserved biotinylation site where biotin is covalently attached to BccP), forming part of an important binding interface. These two mutations surprisingly increased the overall activity of BioID2 by 2.3-fold. While the location of these residues is not near the biotin and ATP binding site, nor the G40 residue responsible for BioID2 promiscuous activity, they could have caused a change in the protein's global conformation or the region supporting the catalytic domain structure. It was later determined that L67 was the residue responsible for this increase in activity (Figure 2.4). The second round of mutations targeted four additional residues in BioID2 L66A L67A. These residues were selected based on their structural and functional relevance. Asp114, which is conserved in PhBPL and located in the vicinity of the BccP binding site, was one of the key targets. The other mutated residues were Lys19 in α-helix 1, Ser112 in β-sheet 6, and Glu227 in the C-terminal loop. Their homologs in PhBPL have been reported to play critical roles in binding Lys90, Gly105 and Glu119 of BccP (Bagautdinov et al., 2008). We assessed each mutation individually to verify that the promiscuous BioID2 maintained its increased activity in the cytoplasm of *Shigella flexneri* str. M90T (Figure 2.2) and then combined the ones that passed the test into a quadruple BioID2 variant with alanine mutations at positions 66, 67, 114 and 227 (Figure 2.3). This variant had a 6.4-fold higher activity compared to the original BioID2 and has mutations on different sides of the hypothetical BccP binding site to disrupt it.

As the project progressed with BioID2, a new study was published about TurboID, a new biotin ligase developed from *E. coli* BirA, the same ligase used to make BioID (Branon et al., 2018). At first, we were skeptical about trying the new ligase instead of BioID2 because its ancestor BirA originates from *E. coli* (Dif et al., 2025). Therefore, introducing TurboID into a *Shigella* context could subject it to regulatory or environmental influences that modulate its activity, as well as unintended interactions due to the similarity between the two microorganisms,

potentially affecting the function and specificity of the ligase itself. On the other hand, TurboID also has its advantages, particularly in applications where rapid labeling is critical. Its faster biotinylation kinetics enable efficient labeling in as little as 10 minutes, which is crucial when studying transient or weak interactions that require speed for proper labeling before their disruption. BioID2, while more efficient than BioID, still requires a longer time frame than TurboID to achieve similar biotinylation levels. This makes TurboID ideal to detect low-abundance interactions or vicinal proteins while other promiscous biotin ligases may struggle to achieve sufficient labeling (Branon et al., 2018; Kim et al., 2016). In light of the above, we decided to include the two *E. coli* derived enzymes, TurboID and its predecessor BioID, along with BioID2 in our quest to find the best suited ligase to study the T3SS in *Shigella*.

Going back to assessing BccP binding, we faced a challenge in designing an experiment that could measure the dissociation constant (K_D) of the biotin ligases binding to BccP due to their quick release of the intermediate biotinyl-5′-AMP during the catalytic cycle. This characteristic of promiscuous enzymes made it difficult to capture stable enzyme-substrate complexes that are essential for accurate K_D measurements and to use techniques such as surface plasmon resonance (SPR) or isothermal titration calorimetry (ITC). Additionally, the promiscuous nature of the enzyme increases the likelihood of multiple binding events with the substrate, further complicating the interpretation of binding data. Instead, we decided to use an *in vitro* biotinylation assay, where we purified each ligase and BccP and incubated them together to monitor BccP biotinylation over time (Choi-Rhee et al., 2004). The expression of these biotin ligases, alongside BccP, was achieved using the *E. coli* expression system with pQE-80L plasmids, which allowed for high-level expression of the target proteins. As confirmed by Coomassie staining (Figure 2.5), IPTG induction resulted in substantial yields of soluble proteins. Although a portion of the proteins stayed in the insoluble fraction, most expressed proteins were recovered in the soluble fraction, especially for BioID and BioID2, indicating the effectiveness of the expression conditions. As for BccP, a high amount of the purified protein was already biotinylated, even in the absence of added biotin ligase. This is due to the presence of endogenous BirA in the cell. To solve this issue, we used 2xYT media to grow the culture before purifying BccP since this media has reduced levels of biotin compared to TB. We also included an additional purification step using streptavidin-agarose beads to remove the remaining endogenously biotinylated BccP (Figure 2.6A). The experimental setup for the *in vitro*

biotinylation assays was inspired by the recommended time frames for each ligase; up to 24 hours for BioID and BioID2, and 15 minutes for TurboID (Branon et al., 2018; Kim et al., 2016; Roux et al., 2012). As expected, TurboID showed a gradual increase in biotinylation over time (Figure 2.6D), with saturation of biotinylation occurring after 15 minutes. This is in line with previous reports demonstrating TurboID's rapid biotinylation kinetics, which are significantly faster than those of BioID and BioID2. However, an interesting and notable result from this study was that both BioID and BioID2 reached saturation much faster than the 16-hour incubation typically used in the *in vivo* experiments (Figure 2.6 B-C). This finding suggests that the bystander activity that inspired the time range in hours is much slower than the specific activity that these ligases have conserved towards their natural substrate BccP. The rapid saturation of biotinylation for BioID and BioID2 prompted us to expand the time frame to include shorter time points, similar to those used for TurboID, which allowed for a more accurate characterization of the biotinylation kinetics for these ligases (Figure 2.6E), revealing differences in the efficiency of biotinylation between BioID, BioID2, and TurboID under the same experimental conditions.

After optimizing the protein purification and the *in vitro* assay, we examined the kinetics of BccP biotinylation in BioID2 and its quadruple mutant. Our initial hypothesis was that the quadruple mutant of BioID2, which included mutations at the predicted BccP binding interface, would show a decrease in BccP biotinylation efficiency due to impaired interaction. In contrast, this mutant exhibited a similar rate of BccP biotinylation compared to the wild-type (WT) BioID2 enzyme (Figure 2.7A). There are several hypotheses that could explain this unexpected outcome. The mutations may have broken or weakened the targeted binding sites. Alternatively, they may have induced a conformational change that exposed other BioID2 residues to BccP, forming a new binding site in place of the altered one. Or the lower biotinylation due to the BccP binding disruption was masked or compensated by the increased bystander effect of the mutant as observed in figure 2.3. To further investigate these two possibilities, we performed a control experiment using RNase A, a protein nonrelated to the BirA native functions in the cell. Our data showed that the biotinylation efficiency of the BioID2 quadruple mutant on RNase A was also comparable to WT BioID2 (Figure 2.7B), suggesting that the observed result of BccP biotinylation was not due to a general increase in enzyme activity since that would also entail an increase in RNase A biotinylation.

Our *in vitro* results invalidated our hypothesis that the quadruple mutant of BioID2, which harbored mutations at the predicted BccP binding interface, reduced the affinity to BccP. The only advantage this mutant provided was the increased promiscuous activity from the WT BioID2. Nevertheless, when subjected to a side-by-side comparison with TurboID, which is known for its rapid biotinylation kinetics, it became evident that TurboID outperformed the BioID2 quadruple mutant in terms of biotinylation efficiency at short time frames (Figure 2.8). These data made us opt for TurboID for further investigations. In fact, TurboID did not only surpass BioID2, but also the other popular ligases BioID and miniTurbo. Our data revealed that TurboID exhibited superior labeling efficiency and faster biotinylation kinetics compared to the other ligases tested. This aligns with previous reports in eukaryotic systems, which have shown that TurboID provides higher activity and broader application potential in proximity labeling (Branon et al., 2018). In our experiments, TurboID consistently outperformed BioID, miniTurbo, and BioID2 at early and late time points (Figure 2.9), demonstrating its ability to rapidly biotinylate proteins in the bacterial cytoplasm. BioID2 and miniTurbo, despite their promising characteristics such as reduced background labeling and smaller size, exhibited lower expression levels and activity in *S. flexneri*, making them less reliable for our study. A notable result in our experiment is that miniTurbo exhibits 1.6-fold lower activity than TurboID, consistent with observations in mammalian cells (Branon et al., 2018). This reduced performance is likely attributed to its truncated N-terminal domain, which compromises its stability. Furthermore, due to miniTurbo's lower expression levels, its real cellular activity seemed to be reduced even more to 13.2 times compared to TurboID. Although miniTurbo has since been shown to function effectively in other bacterial species (Herfurth et al., 2023; Remy et al., 2024), its limited expression and reduced biotinylation efficiency in *S. flexneri* were significant contributors to its suboptimal performance in our system.

A deeper analysis of TurboID showed how much it evolved from its ancestors, not only in activity, but also in exhibiting reduced binding to BccP and the biotin operator, as we will discuss below. This reduced binding affinity is likely due to the mutations of TurboID, which weaken its ability to form stable complexes with biotin-5′-AMP, a requirement for its native functions, particularly homodimerization and binding to the biotin operator (Streaker & Beckett, 2003). Our *in vitro* biotinylation assays revealed a significantly slower biotinylation rate of BccP with TurboID compared to BioID. This suggests that the affinity of TurboID for BccP is reduced to

the point where its higher overall activity cannot fully compensate for the loss in binding efficiency, preventing it from surpassing BioID in this specific context. In contrast, TurboID and BioID demonstrated comparable biotinylation efficiency when tested with RNase A (Figure 2.10B). Both ligases exhibited similar t^{50} values, highlighting that while TurboID has reduced activity toward BccP, it retains similar promiscuous biotinylation efficiency against bystander proteins. The RNase A result was intriguing since theoretically, the protein biotinylation should be faster with TurboID given its higher overall activity. To determine whether this unexpected outcome was specific to RNase A or reflected a more general phenomenon with non-native substrates, we tested BSA as an additional bystander protein with markedly different biochemical properties. BSA offered a valuable contrast to RNase A, possessing a pI (5.77) much closer to BccP's native pI value (4.96 for MBP-BccP) and a substantially higher lysine content (60 lysines vs 10 in RNase A), more comparable to the native substrate MBP-BccP (~41 lysines). If electrostatic complementarity and lysine availability were primary determinants of biotinylation efficiency, BSA would be expected to show enhanced biotinylation compared to RNase A. However, we observed the same pattern of comparable biotinylation kinetics between TurboID and BioID with BSA (Figure 2.11), indicating that this outcome is not specific to RNase A. Analysis of lysine frequencies reveals that MBP-BccP (9.13%) falls between RNase A (9.62%) and BSA (10.29%), yet despite RNase A having the highest lysine density, MBP-BccP is biotinylated significantly faster by BioID. This reinforces that biotinylation efficiency is not solely determined by lysine abundance but rather reflects evolved structural complementarity and favorable lysine accessibility in the native substrate. Although RNase A and BSA have similar or higher lysine frequencies than MBP-BccP, both bystander proteins exhibited similarly slow biotinylation kinetics despite having dramatically different electrostatic properties in the assay (pH 7.0). At this pH, RNase A (pI 8.64) and the biotin ligases (BioID pI 7.86, TurboID pI 8.85) carry net positive charges, while both BSA (pI 5.77) and MBP-BccP (pI 4.96) are negatively charged. Although BSA and MBP-BccP are both negatively charged, creating potentially favorable electrostatic interactions with the positively charged ligases, BSA still demonstrates poor biotinylation efficiency that is comparable to RNase A. This observation strongly reinforces that electrostatic complementarity alone is insufficient for efficient biotinylation, and that neither bystander protein presents optimal binding interfaces for the biotin ligases, which require evolved structural complementarity beyond simple charge matching.

Furthermore, the unexpected results could be also attributed to the absence of physiological molecular crowding and cellular regulatory processes and systems that normally modulate ligase activity *in vivo*, combined with the biochemical properties of the substrate proteins themselves that play crucial roles in determining biotinylation efficiency. Nevertheless, the lower affinity of TurboID for BccP compared to BioID indicates that the evolved, substrate-specific complementarity between BirA and its native substrate has been reduced in TurboID to favor promiscuity. The analysis of bystander protein biotinylation reveals that this trade-off between specificity and promiscuity becomes most apparent when comparing native versus non-native substrate interactions, while proteins lacking any evolutionary relationship to the biotin ligases show similar, poor biotinylation efficiency regardless of the enzyme variant used.

TurboID also exhibited a significantly reduced ability to bind the biotin operator compared to its ancestor, BirA, and the closely related biotin ligase BioID. The electrophoretic mobility shift assay showed that TurboID exhibited only minimal binding, even at higher concentrations. In fact, TurboID shifted only 0.8% of the DNA probe at 50 nM while BirA and BioID shifted 93% and 41% respectively. These findings suggest that the structural modifications responsible for the biotinylation promiscuity in BioID, and to a greater extend in the more active TurboID, contribute to a reduced affinity for the biotin operator. This is due to the faster release of biotin-5′-AMP needed for the enzyme's stable dimer formation to efficiently bind the operon and regulate biotin synthesis (Streaker & Beckett, 2003). Taken together, the decreased affinity for both BccP and the biotin operator indicates that the evolution from BirA to BioID and then TurboID progressively minimized interactions with native biotin-regulated pathways, thereby reducing interference with the cell's endogenous biotin synthesis system.

In conclusion, TurboID was identified as the optimal enzyme among the four biotin ligases examined for proximity labeling studies in *S. flexneri*, due to its rapid and efficient labeling kinetics, as well as its diminished specificity for its native binding partners in the bacterial cytoplasm. While miniTurbo and BioID2 may offer certain advantages in terms of smaller size and reduced non-specific labeling, their lower activity and expression levels that we observed render them less desirable than TurboID. However, despite the decision to proceed with TurboID in this study, it is important to note that the BioID2 quadruple mutant we designed still exhibited enhanced activity over the original BioID2. This improvement, although not enough to

rival TurboID in speed, suggests that BioID2 quadruple mutant still holds potential and could still be a valuable tool for certain experimental settings where a longer labeling window is acceptable, or where more subtle activity is desirable. After deciding on TurboID as the ligase of choice, the next chapter will delve into implementing TurboID to study *S. flexneri*'s T3SS, the experimental methodologies and challenges faced during its application, and the exciting results we obtained.

Chapter 3 : Using TurboID to Study the Proxisome of the T3SS Chaperone IpgC in *Shigella Flexneri*

3.1 Preface

This chapter includes results and figures (Figures 3.3-3.5, and 3.7) adapted and edited from a study published in *mSphere* (Haidar-Ahmad et al., 2024). The experiments and data analysis presented in this chapter were performed by me as part of my thesis work. I acknowledge the contributions of Kyle Tomaro, one of the study's authors, for his assistance with bioinformatics analyses, as well as Mila Gorodnichy and Stephanie Partheniou for their help in constructing the MxiE-TurboID plasmids. The mass spectrometry data are available at the Mass Spectrometry Interactive Virtual Environment (MassIVE) at ftp://massive.ucsd.edu/v08/MSV000095473/. Raw mass spectrometry data and Significance Analysis of INTeractome (SAINT) analyses for IpgC and MxiE, as well as the fold changes of proteins identified in their proxisomes, are provided in the online supplementary material (Haidar-Ahmad et al., 2024). As an author of the article, I retain the rights to reuse this content in my thesis, in accordance with the journal's policy and the Creative Commons Attribution License (CC BY 4.0), which permits reuse provided the original published version is properly cited.

3.2 Introduction

Having established the effectiveness of TurboID as a biotin ligase for proximity labeling in *Shigella flexneri*, we next applied it to investigate the functional dynamics of key components of the T3SS, specifically the regulatory proteins IpgC and MxiE, which play crucial roles in the transcriptional activation of this system (Figure 1.2). By fusing TurboID to IpgC and MxiE, we aimed to evaluate its adaptability in the bacterial cytoplasm and to capture and characterize the proxisomes of our proteins of interest in both the active and inactive states of the T3SS. In the following sections, we describe the results of using TurboID fusions to explore the protein networks in the vicinity of IpgC and MxiE in *Shigella flexneri.*

3.3 Results

3.3.1 Optimization of IpgC-TurboID expression and functional complementation in *Shigella flexneri*

To evaluate the functionality of IpgC-TurboID fusion, we first examined its expression levels and ability to complement an *ipgC* knockout (Δ*ipgC*) in strains of an otherwise WT genetic background (inactive T3SS), or in a Δ*ipaD* genetic background (active T3SS). Given that BioID was a more established biotin ligase, we included BioID-IpgC construct as a control in the preliminary experiments to compare its performance with TurboID. However, after optimizing expression, we focused solely on TurboID, the primary candidate for our study. Initially, expression from the pSU2.1-based constructs was barely detectable by immunoblotting with a polyclonal antibody against IpgC (Figure 3.1A), both for TurboID and BioID protein fusions in Δ*ipgC*. Likewise, since IpgC plays a role in activating *ipaH* expression, we used IpaH production as a readout for complementation in the *ipgC* knockout active strain (Δ*ipaD* Δ*ipgC*). If IpgC fusions function normally, they should restore *ipaH* expression. Our results showed only minimal IpaH production, substantially lower than the levels observed in the active T3SS strain (Δ*ipaD*) control (Figure 3.1B). These results suggest that expression levels of the IpgC fusions were insufficient for full complementation. To address this issue, we first considered the possibility that the biotin ligases or the Myc tag might be interfering with IpgC expression or stability. To test this, we extended the linker between IpgC and TurboID/BioID from 5 to 10 amino acids to provide additional spatial separation between the fused proteins. However, this modification did not enhance expression (Figure 3.1C), suggesting that neither the biotin ligases nor the Myc tag were directly affecting IpgC.

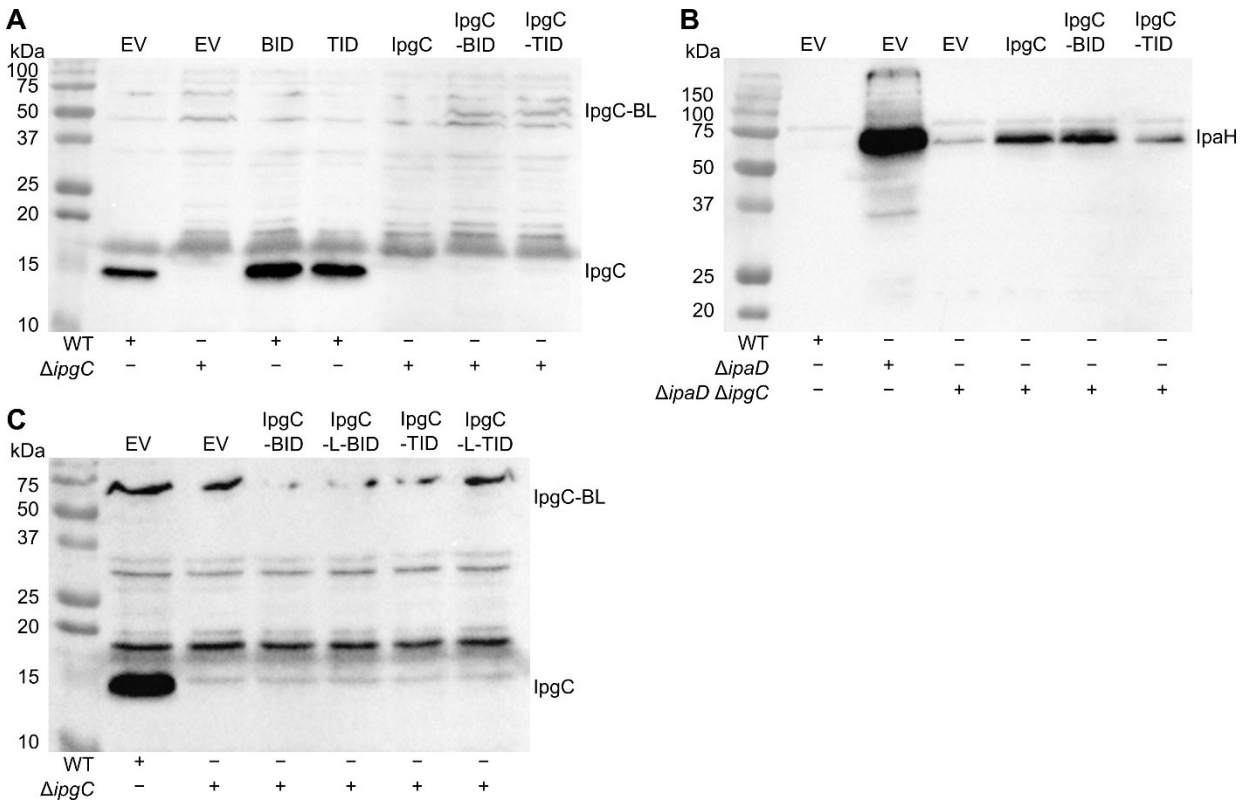


Figure 3.1. TurboID and BioID fusions with IpgC exhibit weak complementation. Experiments were conducted using the indicated *S. flexneri* strains. (A) Immunoblot analysis of IpgC fusions expressed from pSU2.1 plasmid backbone (B) Assessment of IpgC-dependent complementation based on *ipaH* expression. (C) Immunoblot analysis of IpgC fusion constructs incorporating a 10-amino acid linker (L) between IpgC and the biotin ligase. In panels A and C, IpgC was detected using a polyclonal anti-IpgC antibody. Bands corresponding to free IpgC and biotin-ligase fusions (BL) are indicated; unlabeled bands represent contaminants. In panel B, IpaH was detected using a polyclonal anti-IpaH9.8 antibody.

Next, we tested a different plasmid backbone. All preceding experiments in *Shigella* were conducted using the pSU2.1 vector, which carries a *lac* promoter that is constitutively active due to the absence of *lacI* on the plasmid. While pSU2.1 has been successfully used for other applications, it is possible that the strength or regulation of its promoter was suboptimal for expressing the IpgC-TurboID fusion at functional levels. Additionally, pSU2.1 contains a lower-copy-number origin of replication, which may have limited overall plasmid abundance and thus protein expression. To improve expression, we tested two high-copy-number pUC18 derivatives with revised multiple cloning sites; pUC18.1, carrying a *lac* promoter, and pUC18.1rp, carrying the strong constitutive *rpsM* promoter. Both constructs increased IpgC expression relative to pSU2.1, with pUC18.1 yielding slightly lower expression than pUC18.1rp. pUC18.1rp restored IpgC expression to levels comparable to the native expression in WT *Shigella*, as confirmed by immunoblotting (Figure 3.2). pUC18.1 was included as a backup to monitor potential overexpression.

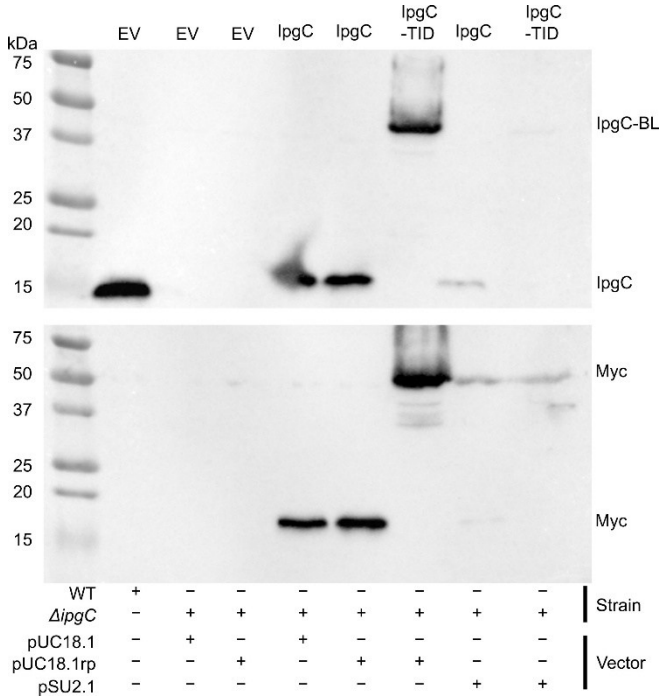


Figure 3.2. pUC18.1rp plasmid backbone rescued IpgC-TurboID expression.
Immunoblot analysis comparing the expression of IpgC and IpgC-TurboID from various plasmid backbones in *Shigella flexneri* Δ*ipgC*. IpgC expression was detected using a polyclonal anti-IpgC antibody (top panel) and an anti-Myc antibody (bottom panel). Bands corresponding to free IpgC and biotin-ligase fusions (BL) are indicated; unlabeled bands represent contaminants.

3.3.2 TurboID fusions with IpgC and MxiE are expressed and functional

Following the optimization of IpgC-TurboID fusions and the selection of pUC18.1rp as the backbone vector, we next investigated the functional compatibility of TurboID fused to either the N- or C-terminus of the bait protein. To broaden our analysis, we included both MxiE and IpgC fusions and assessed their expression and complementation in the corresponding deletion strains. Constructs expressing IpgC-TurboID-Myc and TurboID-Myc-IpgC (IpgC-TurboID and TurboID-IpgC thereafter), as well as MxiE-TurboID-Myc and TurboID-Myc-MxiE (MxiE-TurboID and TurboID-MxiE thereafter), were generated. Expression was evaluated by immunoblotting using a polyclonal antibody against IpgC for the IpgC fusions, and a monoclonal antibody against Myc tag for MxiE fusions due to the lack of a specific MxiE antibody. Complementation assays were performed to determine the ability of these fusions to rescue *ipaH* expression in Δ*ipgC* and Δ*mxiE* mutants in both the off-state (WT genetic background) and the on-state conditions (Δ*ipaD* background for IpgC constructs and *ipaB4* background for MxiE constructs). Both IpgC fusion constructs were expressed at levels comparable to endogenous IpgC (Figure 3.3A, top panel). In the TurboID-IpgC lanes, we observed additional faint bands of lower molecular weight, likely representing degradation products. Despite this, the constructs were functional, as both IpgC-TurboID and TurboID-IpgC successfully rescued *ipaH* expression in the Δ*ipaD* Δ*ipgC* strain. As expected, IpaH levels were higher in the on-state (Δ*ipaD* Δ*ipgC*) compared to the off-state (Δ*ipgC*), mirroring the expression pattern seen with endogenous IpgC in the Δ*ipaD* background. Interestingly, *ipaH* expression in the Δ*ipgC* was slightly higher than in the WT strain, suggesting a potential gain-of-function effect. This may reflect a partial reduction in the anti-co-activator activity of IpaB and IpaC against IpgC. If the TurboID fusion somehow subtly altered IpgC's interaction with IpaB and IpaC, it could have facilitated the formation of IpgC-MxiE complexes, thereby enhancing *ipaH* transcription even under off-state conditions. Such effects are not uncommon with protein fusions, which can influence protein stability, interactions, or regulatory dynamics. Nonetheless, both IpgC-TurboID constructs were functional, as evidenced by their ability to complement *ipgC* in the knockout strains.

In contrast with IpgC, due to the lack of a specific antibody against MxiE, the expression of MxiE-TurboID fusion constructs was assessed using an antibody against the Myc tag. Immunoblotting revealed that both MxiE fusion constructs were expressed, albeit to different degrees. MxiE-TurboID showed signs of significant degradation and failed to effectively

complement *ipaH* expression in the Δ*mxiE ipaB4* (on-state) strain (Figure 3.3B, bottom panel). In contrast, TurboID-MxiE migrated as a single band and successfully restored IpaH production in the same mutant background. However, *ipaH* expression driven by TurboID-MxiE was also elevated in the Δ*mxiE* (off-state) strain compared to the WT. This suggests that increased availability of MxiE may have led to transcriptional upregulation of *ipaH* genes, potentially due to overproduction or a reduced ability of the anti-activator OspD1 to bind and inhibit MxiE. A caveat of this analysis is that, due to the absence of a MxiE-specific antibody, we were unable to directly compare the expression levels of the fusion constructs to native MxiE in the WT strain and therefore could not determine whether the apparent overcomplementation by TurboID-MxiE was due to overexpression. Together, these results indicate that all TurboID fusion constructs retained partial to full functionality, with IpgC-TurboID and TurboID-MxiE showing the most effective complementation relative to the T3SS active and inactive control strains.

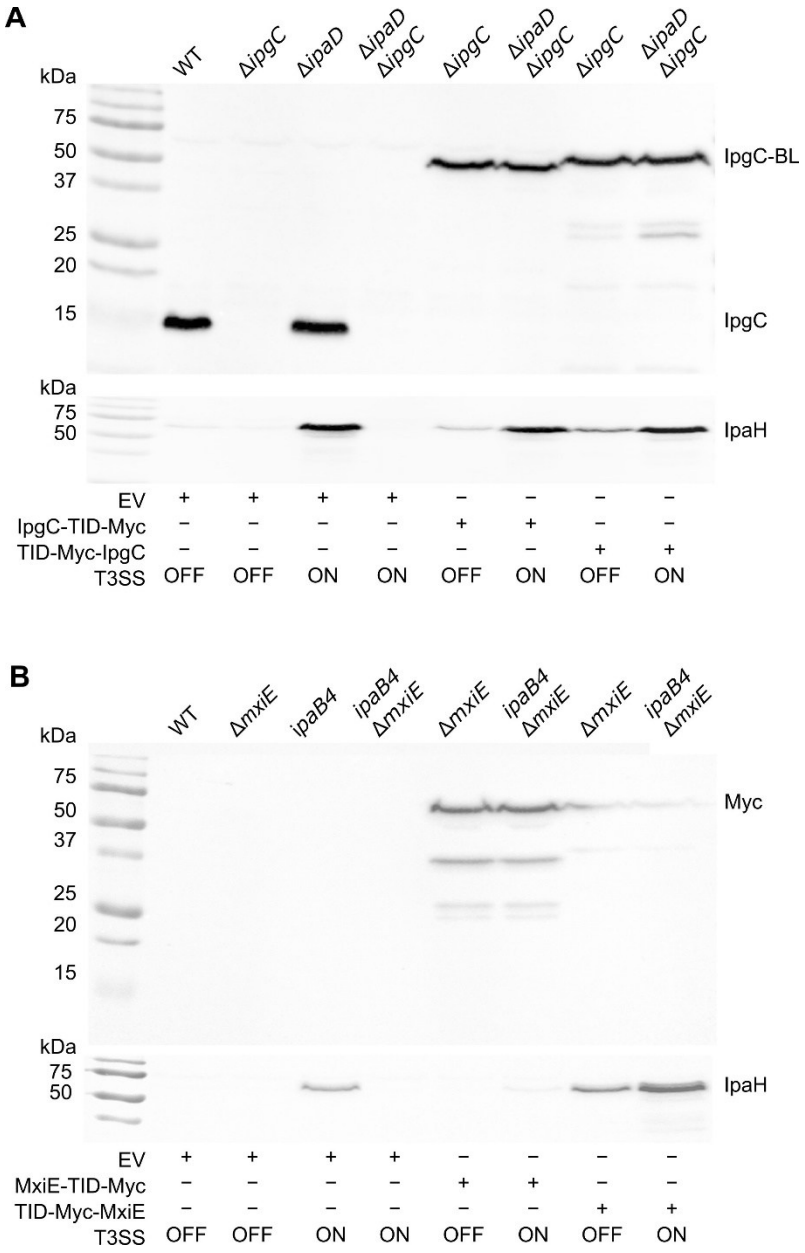


Figure 3.3. IpgC-TurboID fusions fully rescue *ipaH* expression, while MxiE fusions exhibit partial complementation in *Shigella flexneri*.
(A) Expression (top panel) and complementation (bottom panel) of IpgC-TurboID fusions were assessed in strains with the T3SS in the off-state (Δ*ipgC*) and on-state (Δ*ipaD* Δ*ipgC*). IpgC fusions, expressed from the pUC18.1rp vector, were detected by immunoblotting with a polyclonal anti-IpgC antibody, while IpaH production was evaluated using a polyclonal anti-IpaH9.8 antibody. (B) Expression (top panel) and complementation (bottom panel) of MxiE-TurboID fusions were assessed in strains with the T3SS in the off-state (Δ*mxiE*) and on-state (*ipaB4* Δ*mxiE*). Expression of MxiE fusions was detected with a monoclonal anti-Myc antibody, and IpaH production was measured by immunoblotting with anti-IpaH9.8. Data are representative of three independent biological replicates (n = 3).

3.3.3 TurboID fusions with IpgC exhibit distinct biotinylation patterns reflecting T3SS states

We next examined the biotinylation patterns of TurboID fusions with IpgC and MxiE and compared them to that of the free TurboID to assess whether fusion orientation influenced biotinylation efficiency and specificity (Figure 3.4 A-B). One of the first observations was that TurboID exhibited substantial biotinylation even before the exogenous biotin was added to the culture medium. This is consistent with previous reports in HEK 293T cells, yeast, and flies (Branon et al., 2018), where TurboID was shown to utilize basal levels of biotin present in the medium to catalyze protein biotinylation. To better understand the kinetics of TurboID-mediated biotinylation, we supplemented cultures with 50 μM biotin and measured protein biotinylation at multiple time points. Biotinylation levels increased rapidly within the first 10 minutes of supplementation and continued to rise more gradually at 20 minutes. As an example, densitometric analysis of streptavidin blots revealed that in Δ*ipgC* strains expressing IpgC-TurboID and TurboID-IpgC, approximately 61% ± 17% and 68% ± 24% of total biotinylation signal, respectively, was detected at 10 minutes following biotin supplementation, relative to the signal observed prior to biotin addition (no biotin control) (Figure 3.4A). Although TurboID enables continuous labeling of proximal proteins even under conditions of limited biotin availability, our data indicates that the majority of biotinylation occurs after biotin supplementation. This highlights the importance of including appropriate controls, especially in downstream experiments, to effectively differentiate true biotinylation signals from background noise.

Interestingly, the biotinylation pattern of IpgC-TurboID fusions differed from that of free TurboID, suggesting that these fusions engage in specific protein interactions that are distinct from those of freely diffusing TurboID. In the off-state, several unique biotinylated bands were detected between 37 and 75 kDa (Figure 3.4A). These bands were less intense in strains where the T3SS was in the on-state (Figure 3.4B), suggesting that the corresponding proteins may be T3SS substrates. Their reduced biotinylation in the on-state could reflect active secretion or changes that decrease their proximity to IpgC-TurboID. In contrast, MxiE-TurboID fusions exhibited biotinylation patterns that were indistinguishable from those of free TurboID (Figure 3.5 A-B). This suggests that, unlike IpgC, MxiE-TurboID fusions may not interact with proteins

detectable by streptavidin blotting. Instead, high background biotinylation of abundant cellular proteins may have masked any prey-specific labeling events.

Given these observations, we proceeded to use mass spectrometry-based proteomics to investigate the proxisomes of IpgC and MxiE under conditions mimicking the off- and on-states of the T3SS during *Shigella* infection, allowing us to map potential interactors and gain deeper insights into the molecular networks regulating T3SS activity.

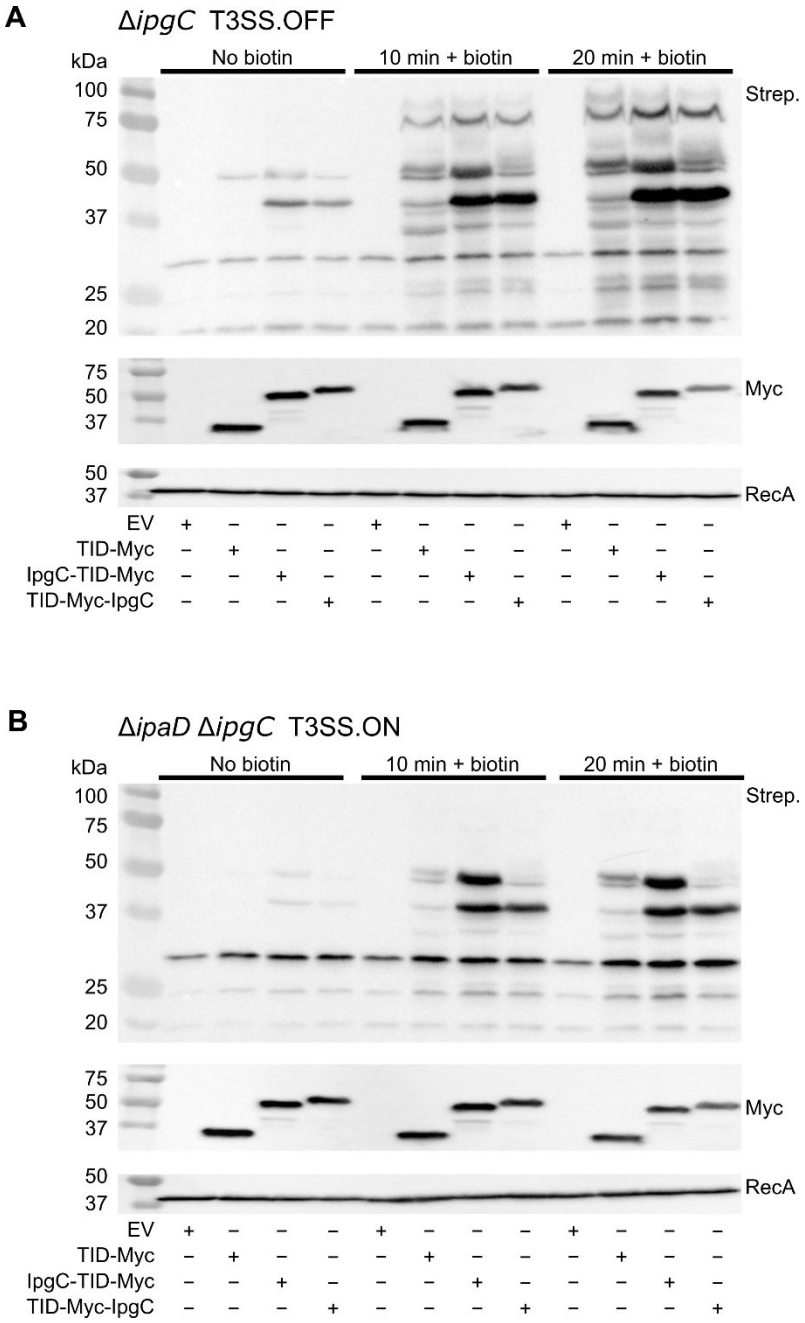


Figure 3.4. Distinct T3SS-dependent biotinylation patterns of IpgC-TurboID fusions. (A) Immunoblot analysis of biotinylation in *S. flexneri* Δ*ipgC* (T3SS off-state) expressing IpgC-TurboID-Myc and TurboID-Myc-IpgC, compared with free TurboID-Myc and an empty vector control. (B) Corresponding biotinylation profiles in the *S. flexneri* Δ*ipaD* Δ*ipgC* strain (T3SS on-state) Constructs were expressed using the pUC18.1rp vector. Protein expression was verified using a monoclonal anti-Myc antibody, with RecA serving as a loading control, while biotinylation was detected with streptavidin-horseradish peroxidase. Cultures were supplemented with 50 μM biotin, and samples were collected at 10- and 20-minutes post-supplementation.

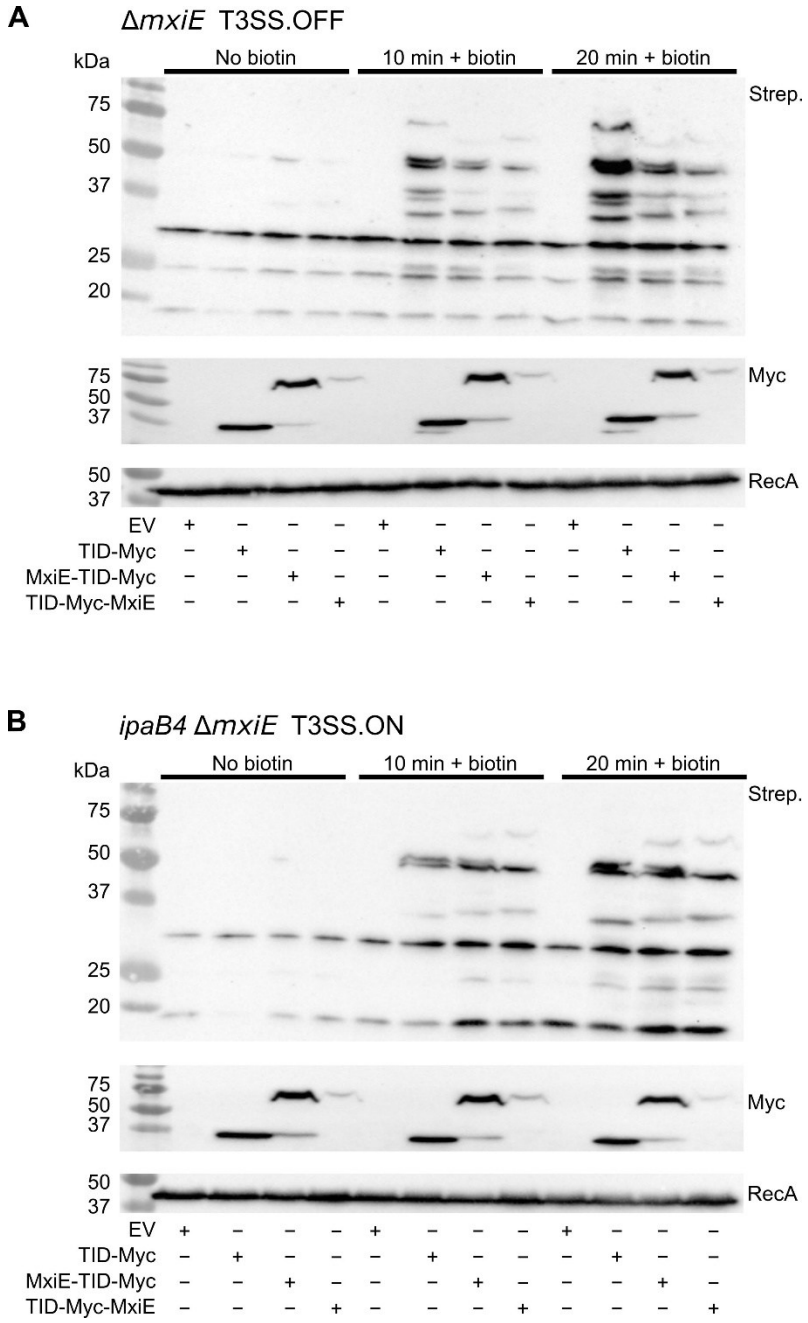


Figure 3.5. MxiE-TurboID fusions yield biotinylation profiles comparable to free TurboID across T3SS states.
(A) Immunoblot showing biotinylation in *S. flexneri* Δ*mxiE* (T3SS off-state) expressing MxiE-TurboID-Myc and TurboID-Myc-MxiE, alongside free TurboID-Myc and an empty vector control. (B) Biotinylation profiles in the *ipaB4* Δ*mxiE* strain (T3SS on-state). All constructs were expressed from the pUC18.1rp vector. Expression levels were determined by immunoblotting with a Myc-specific monoclonal antibody, using RecA as a loading control, while biotinylation was assessed with streptavidin-horseradish peroxidase. Cultures were treated with 50 μM biotin, and samples were collected at 10- and 20-minutes post-supplementation.

3.3.4 Clustering patterns of IpgC and MxiE-TurboID fusions reveal differential influences of T3SS state and fusion orientation

To identify proteins interacting with IpgC and MxiE in the active state of the T3SS, we performed proximity-based proteomics using mass spectrometry. Strains expressing IpgC-TurboID and MxiE-TurboID fusions were grown in the presence of biotin, and the biotinylated proteins were isolated using streptavidin resin and identified using LC-MS/MS. To analyze the reproducibility and structural influences of TurboID fusions, we generated heatmaps of the biotinylated proteomes from different replicates of IpgC and MxiE fusions. The goal was to determine whether the primary clustering pattern reflected the T3SS activation state (on vs. off) or the orientation of the TurboID fusion (N-terminal vs. C-terminal placement). Interestingly, IpgC replicates grouped consistently based on the T3SS state, regardless of fusion orientation (Figure 3.6, top panel). This suggests that IpgC remains functionally engaged in T3SS-related functions in both fusion orientations, and that the captured proxisome is strongly dictated by the T3SS regulatory state. In contrast, the clustering pattern for MxiE replicates was less consistent, with samples sometimes grouped by T3SS state and other times by fusion orientation (Figure 3.6, bottom panel). This variability is consistent with our earlier observation that MxiE-TurboID fusions did not show striking differences in biotinylation patterns between the on- and off-states or between N- and C-terminal fusion orientations compared to the free TurboID (Figure 3.5). These findings suggest that the biotinylation profile of MxiE-TurboID may be influenced by factors such as fusion-induced conformational changes or limited accessibility of interactors, rather than being strongly dictated by T3SS status. This highlights the importance of analyzing multiple fusion configurations and biological conditions to more reliably capture the full spectrum of a protein’s network.

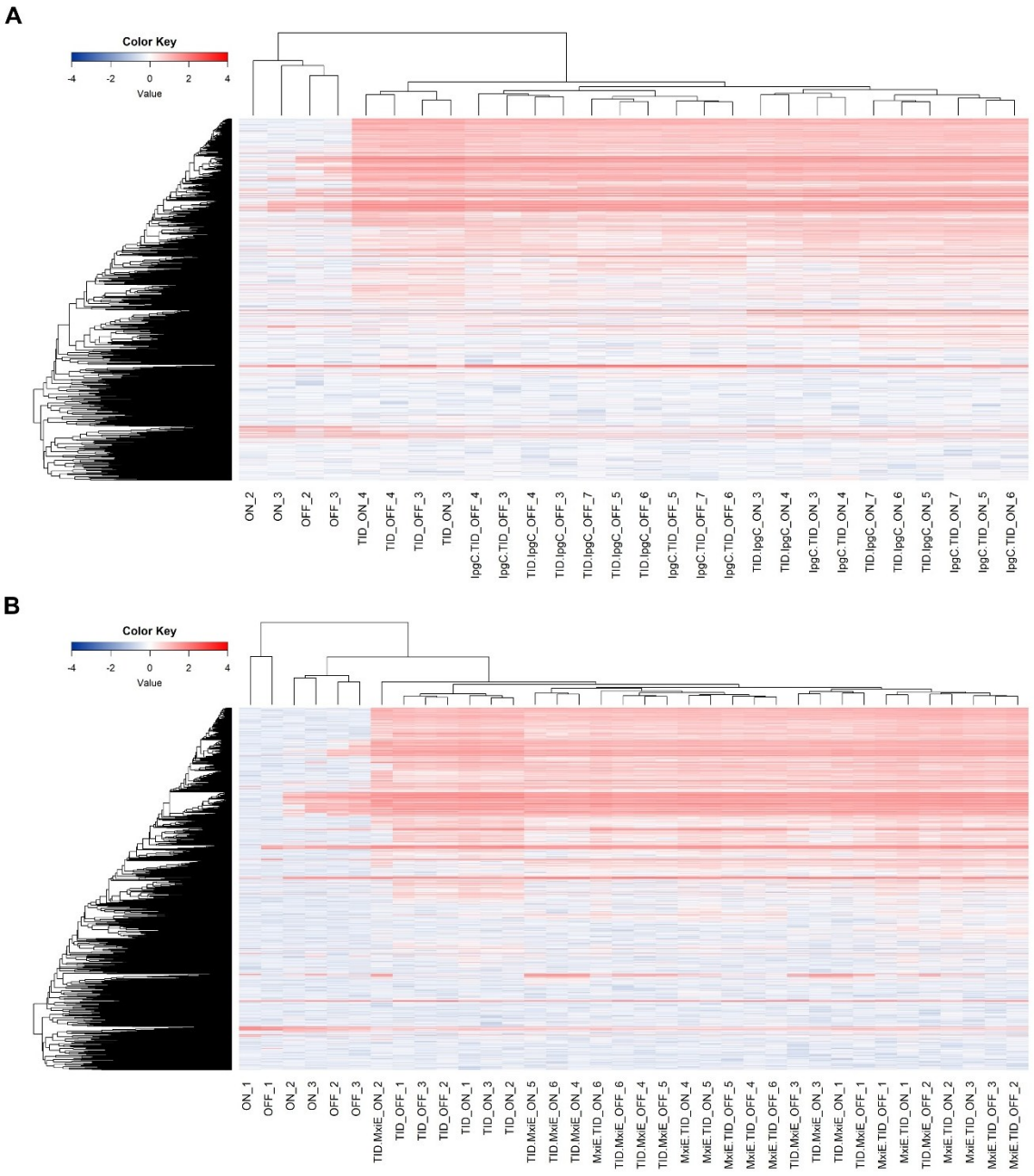


Figure 3.6. Heatmaps using Label-Free Quantification intensities for IpgC and MxiE-TurboID fusions.

(A) Heatmap of IpgC-TurboID fusions. “ON” denotes the active T3SS state in the Δ*ipaD* Δ*ipgC* strain, while “OFF” represents the inactive state in the Δ*ipgC* strain. Replicates cluster distinctly by T3SS state, indicating that IpgC’s vicinal protein composition is strongly modulated by secretion activity. (B) Heatmap of MxiE TurboID fusions. “ON” corresponds to the active state in the *ipaB4* Δ*mxiE* strain, and “OFF” to the inactive state in the Δ*mxiE* strain. Here, clustering patterns are mixed, with some replicates grouping by T3SS state and others by fusion orientation, reflecting the influence of both biological state and TurboID placement on the detected proxisome.

3.3.5 Refining the IpgC-MxiE proxisome: addressing contaminants and false positives

Mass spectrometry raw data revealed more than 1600 protein detected with IpgC-TurboID fusions and more than 1880 with MxiE-TurboID. The complete raw data are available via the link provided in the chapter preface, and supplementary datasets are accessible in the published article (Haidar-Ahmad et al., 2024). However, not all detected proteins represent true interactors, and not all data points are biologically significant. To ensure the validity of our findings, we evaluated protein relevance based on reproducibility across replicates, enrichment relative to controls, and the identification of background signals and potential false positives. Using spectral counts and the Significance Analysis of INTeractome (SAINT) algorithm (Choi et al., 2011), we filtered for significantly enriched proteins in TurboID fusions compared to TurboID alone or empty vector controls, allowing us to compare the proxisome of IpgC and MxiE in both the off- and on-states of the T3SS.

Analyzing the IpgC and MxiE pathway using proximity labeling presents significant challenges, particularly in data interpretation. To ensure accuracy, results must be stringently filtered to account for various sources of false positives. A key consideration is that the expressed fusion proteins are localized in the *Shigella* cytoplasm, which naturally contains biotinylated proteins such as BccP, as well as biotin biosynthesis enzymes like BioB and BioD whose expression is regulated by BirA. These endogenous proteins, involved in biotin homeostasis, are common contaminants in bacterial systems studied with promiscuous biotin ligases and were therefore excluded during data analysis. Another major source of contamination includes translation-associated proteins, which are frequently detected in mass spectrometry analyses due to their high abundance. Since TurboID fusions do not require activation and can initiate biotinylation immediately after translation, proteins involved in the translation process, such as ribosomal proteins, are susceptible to non-specific labeling, further contributing to background noise in the dataset. Finally, a significant class of false positives that is unique to this biological system is the T3SS substrates. These substrates depend on MxiE and IpgC for expression, and they are highly abundant in strains with active Type III secretion. Additionally, since transcription and translation are often coupled in bacteria, proteins encoded by MxiE- and IpgC-regulated genes could transiently be in close proximity to them during synthesis, which could lead to their enrichment as false positives. These factors greatly refined our dataset, as rigorous

filtering was necessary to remove non-specific biotinylated proteins, substantially reducing the number of proteins considered for further analysis.

3.3.6 Expansion of the IpgC proxisome in the T3SS on-State and identification of vicinal proteins

SAINT analysis demonstrated that IpgC's proxisome expands upon T3SS activation, as evidenced by a greater number of enriched proteins in the on-state compared to the off-state (Figure 3.7A). Among the identified proteins, IpgC's well-established binding partners; IpaB, IpaC, and MxiE, were confidently detected in both TurboID fusions validating again the functionality of both IpgC-TurboID fusions. In contrast, MxiE's proxisome remained relatively unchanged, with a comparable number of enriched proteins in both states (Figure 3.7B). While MxiE's known interactor OspD1 was detected, IpgC itself was not identified as a proximal protein of MxiE. The alignment of the IpgC dataset with the established T3SS regulatory model (Figure 1.2) prompted us to focus our analysis on the state-dependent changes in its proxisome.

To refine the list of IpgC-associated proteins, we prioritized candidates that exhibited a significant fold-change (FC > 2) between T3SS states in at least one of the TurboID fusion orientations. Among these, several known or suspected T3SS-related proteins were identified. For example, IpaB and IpaC were slightly enriched in the off-state ($1 < FC_{off/on} < 2$), while MxiE showed a marked increase in the on-state (FC on/off 2.7 and 3.5 across the two fusion orientations), consistent with established T3SS signaling models. The detection of these known components supports the validity of our proximity labeling approach and confirms that the TurboID fusions captured biologically relevant changes in protein association across T3SS states. Interestingly, we also observed an unexpected enrichment of OspD1 in the off-state (FC off/on 1.2 and 1.6), despite its known interaction with MxiE rather than IpgC. Importantly, all these interactions were identified as high-confidence hits in both IpgC-TurboID fusions datasets (false discovery rate (FDR) < 10%), reinforcing the reliability of these findings.

Most remaining proteins associated with IpgC were enriched in the on-state. Notably, we identified the ATPase SpaL (FC on/off ≈ 1.3), a component of the T3SS sorting platform, and the phage shock protein A, PspA (FC on/off ≈ 2.7). These proteins were only significantly enriched in the IpgC-TurboID fusion dataset (FDR < 10%). IpgC's proxisome also included several

transcriptional regulators. Among them, VirF (FC on/off ≈ 1.2), a main regulator of the *Shigella* T3SS regulatory cascade, was detected as a vicinal protein. Additionally, McbR (FC on/off ≈ 2.7), SlyA (FC on/off ≈ 2.0), and CysB (FC on/off ≈ 4.0), were also enriched around IpgC in the on-state. The identification of these transcriptional regulators suggests that IpgC may have a wider impact on transcriptional regulation beyond its well-characterized interactions with the transcription activator MxiE.

In addition to its T3SS associated and transcription factors, IpgC's proxisome contained several housekeeping proteins, most of which were metabolic enzymes enriched in the on-state. STRING network analysis of these enzymes revealed a highly connected protein-protein interaction (PPI) network with a significant enrichment *p*-value of 8.32×10^{-8} † (Figure 3.8). Further gene ontology (GO) analysis identified six main enriched GO terms. While there is no direct evidence that these metabolic enzymes physically interact with IpgC, their presence in the dataset suggests that T3SS activation may be accompanied by certain metabolic shifts.

† *Note: The enrichment p-value reported here (8.32×10^{-8}) differs from that in the published version of this analysis (3.2×10^{-9}), likely due to differences in STRING database version or input parameters used at the time of analysis.*

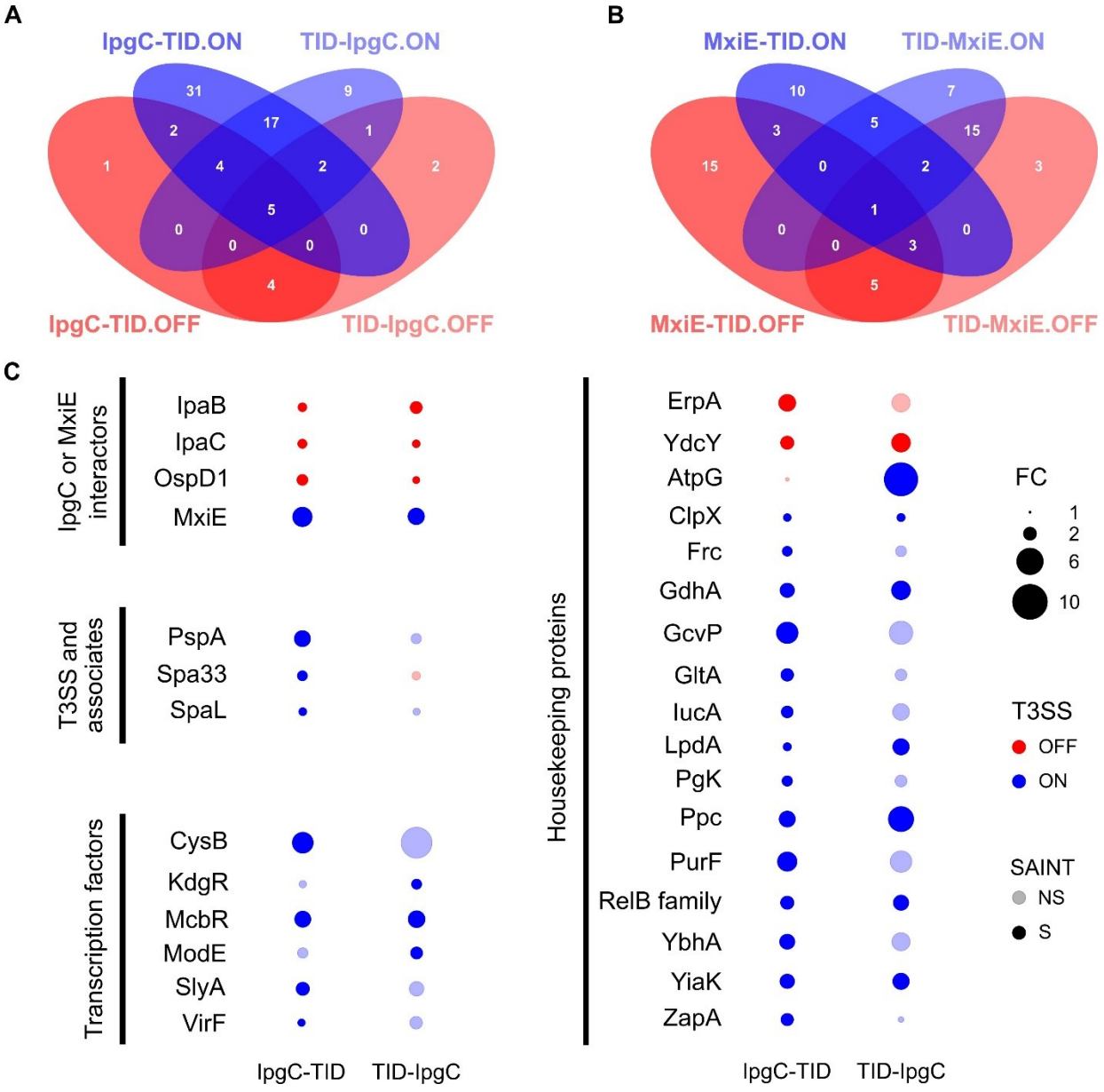
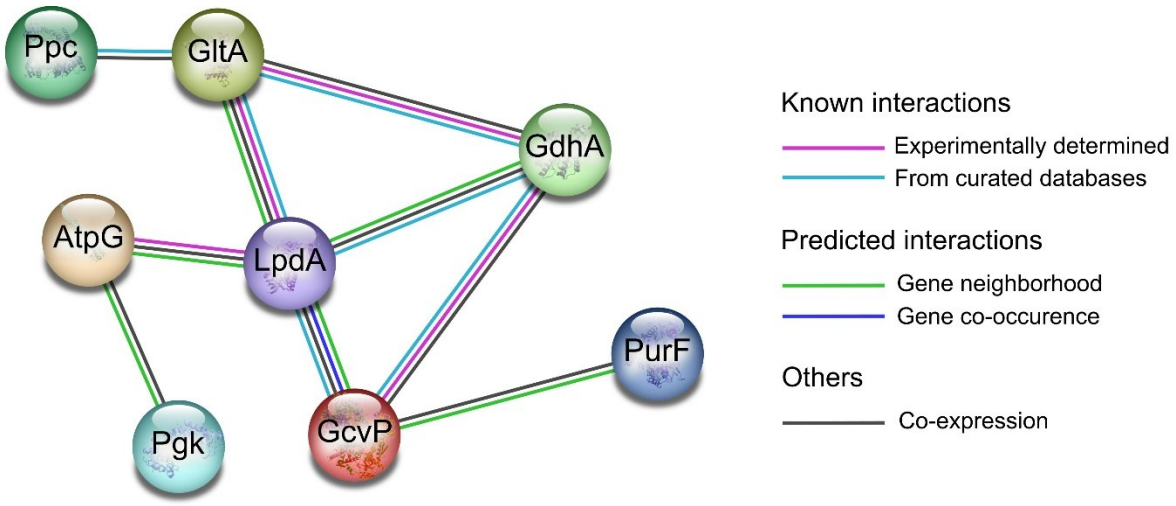


Figure 3.7. Mapping the IpgC and MxiE proxisomes using TurboID.
Venn diagrams illustrating the overlap of proteins detected in the off- and on-states of the T3SS for TurboID fusions with (A) IpgC (n = 5) and (B) MxiE (n = 6). (C) Dot plot depicting the fold change (FC) and SAINT significance [(FDR) < 10%] of high-confidence proteins identified with IpgC-TurboID-Myc (IpgC-TID) and TurboID-Myc-IpgC (TID-IpgC). The size of the dot reveals the fold change value, the color corresponds to the T3SS off-state (red) or on-state (blue) in which FC > 1, and the opacity indicates the significance (S) or non-significance (NS) of the enrichment of a given protein with both protein fusions using SAINT. Identified proteins are categorized into four functional groups: IpgC or MxiE interactors, T3SS and associates, transcription factors, and housekeeping proteins.



GO-term	Description	Proteins in Network	Strength	FDR
GO:0019464	Glycine decarboxylation	LpdA, GcvP	2.41	0.03
GO:0046034	ATP metabolic process	LpdA, AtpG, PgK	1.73	0.03
GO:0006099	Tricarboxylic acid cycle	LpdA, GltA, PpC	1.73	0.03
GO:0009150	Purine ribonucleotide metabolic process	LpdA, AtpG, PgK, PurF	1.42	0.03
GO:1901605	Alpha-amino acid metabolic process	LpdA, GdhA, GcvP, PurF	1.02	0.05
GO:0019752	Carboxylic acid metabolic process	LpdA, PgK, GdhA, GcvP, PurF, PpC	0.77	0.03

Figure 3.8. STRING network analysis of housekeeping proteins identified in the IpgC proxisome.
The network represents the analysis of a subset of housekeeping proteins identified in Figure 3.7, panel C using *E. coli* K-12 database and a medium confidence of 0.4. The colors of the lines connecting the nodes indicate the type of interaction evidence. The table summarizes the six most enriched Gene Ontology (GO) terms related to this network and their corresponding proteins.

3.3.7 AlphaFold predictions support known IpgC interactions but show low confidence for novel candidates

To evaluate the structural feasibility of IpgC interactions with various T3SS associated proteins and transcription factors that passed the SAINT screening, AlphaFold was used to generate structural models of their complexes with IpgC. All candidate proteins were modeled in their monomeric form, and for transcription factors with known functional oligomeric states, the corresponding dimeric or tetrameric forms were also modeled (Bartoli et al., 2020; Grunden et al., 1999; LeBoeuf & Grove, 2025; Lord et al., 2014; Nasser et al., 1992). The reliability of these models was assessed using the predicted template modeling (pTM) score, which evaluates the accuracy of the overall fold, and the interface predicted template modeling (iPTM) score, which assesses confidence in the relative positioning of subunits within the complex (Jumper et al., 2021; Varadi et al., 2022). The results indicated that the ipTM scores for all newly detected proteins were below 0.8, suggesting low confidence in strong interactions between IpgC and these candidates (Table 3.1). Surprisingly, the known interactors IpaB, IpaC, and MxiE yielded ipTM scores ranging from 0.76 to 0.78. Although these values fall below the high-confidence threshold of 0.8, they remain within the intermediate confidence range (0.6 - 0.8). Given that the ipTM scores for known interactors were not substantially high, the low scores observed for the new candidate proteins cannot be used to support or discard their potential interaction with IpgC.

Table 3.1. AlphaFold-predicted iPTM and pTM scores for IpgC interactions with various T3SS components and transcription factors identified as significant by SAINT.
The pTM score evaluates the reliability of the overall structural fold, with values above 0.5 indicating a potential resemblance to the true structure. The iPTM score assesses the predicted spatial arrangement of protein subunits within a complex, where values above 0.8 suggest high confidence, those between 0.6 and 0.8 fall into an uncertainty zone, while scores below 0.6 indicate unreliability of the predictions (EMBL-EBI, 2025).The proteins were assessed using their functional oligomeric states.

Protein	UniProt Accession	Oligomeric state	ipTM score	pTM score
IpaB	P18011	monomer	0.78	0.49
IpaC	P18012	monomer	0.76	0.42
OspD1	Q6XW07	monomer	0.23	0.42
MxiE	P0A2S7	monomer	0.76	0.76
PspA	P0AFM8	monomer	0.24	0.42
SpaO (Spa33)	P0A1K9	monomer	0.14	0.36
SpaL (Spa47)	P0A1C1	monomer	0.13	0.69
CysB	A0A4P7TNH5	monomer	0.1	0.47
		dimer	0.45	0.55
		tetramer	0.3	0.41
KdgR	A0A4P7TM43	monomer	0.1	0.47
		dimer	0.33	0.38
McbR	A0A2Y5AZW9	monomer	0.38	0.68
		dimer	0.56	0.69
ModE	A0A4P7TSI9	monomer	0.17	0.39
		dimer	0.21	0.36
SlyA	P0A8W4	monomer	0.13	0.46
		dimer	0.21	0.39
VirF	P0A2T1	monomer	0.17	0.5

3.3.8 IpgC absence affects acid resistance in *Shigella flexneri*

Analyzing IpgC-TurboID mass spectrometry data, we found that SlyA was significantly enriched in the T3SS on-state, with an FC on/off ratio of approximately 2. This discovery caught our interest because there is no known direct interaction between IpgC and SlyA. SlyA is primarily recognized for its role in activating acid resistance by promoting the glutamate-dependent acid resistance (GDAR) pathway (*gadA*, *gadB*), which helps neutralize acidic environments. Additionally, when overexpressed, SlyA has been shown to restore the expression of T3SS components in non-permissive conditions (Weatherspoon-Griffin & Wing, 2016). The identification of SlyA and several other transcription factors, as well as MxiE, within the IpgC proxisome in the T3SS on-state led us to consider whether IpgC might be involved in broader regulatory processes. Based on these observations, we decided to investigate the potential connection between IpgC and acid resistance. We hypothesized that IpgC may interact with SlyA in the on-state, which would have consequences on acid resistance (AR) in *Shigella*.

To determine the optimal conditions for assessing the effect of IpgC on acid resistance, we exposed the WT (T3SS inactive) and Δ*ipaD* (T3SS active) control strains to varying pH levels and acid exposure durations, followed by viability assessment using a *Shigella* spot assay with serial dilutions. Based on preliminary testing, pH 2.5 and a 2-hour (T3) exposure time were selected for further experiments (Figure 3.9), as these conditions produced a measurable survival difference between the two strains while maintaining bacterial viability. This experiment was conducted both in the absence and presence of glutamate, a key component of the GDAR pathway (Foster, 2004; Ojha et al., 2021). We then repeated the experiment adding Δ*ipgC* and Δ*ipaD* Δ*ipgC* to the tested strains. Results showed reduced AR in Δ*ipaD* Δ*ipgC* compared to both Δ*ipaD* and Δ*ipgC* in the absence of glutamate. In most replicates, Δ*ipaD* Δ*ipgC* exhibited almost no growth after 2 hours of acid exposure, suggesting that IpgC presence is important for AR specifically in the active state of T3SS. In contrast, the absence of IpgC in the Δ*ipgC* (inactive) strain did not result in a similar drastic effect on AR (Figure 3.10). In the presence of 50 mM glutamate, AR was restored in all strains as indicated by their growth under low pH conditions. Spot intensities were measured by densitometry using ImageJ by defining regions of interest around each spot, and densities were averaged across three biological replicates. Values were expressed relative to T_0 to indicate growth or decline in viability. Differences in colony density and growth dynamics, particularly the tendency of sparse colonies to grow larger than

crowded ones, combined with variability in imaging quality, contributed to relatively large standard deviations and limited the statistical significance in ANOVA analysis. For this reason, results are presented primarily qualitatively. In future experiments, increasing the number of biological replicates and using colony counting could reduce variability and allow more robust quantitative analysis. Nonetheless, this assay highlights the pronounced acid sensitivity of the Δ*ipaD* Δ*ipgC* mutant, as evidenced by its consistent lack of growth under acidic conditions.

These findings suggest that the absence of IpgC, as seen in the Δ*ipaD* Δ*ipgC* strain, might be hindering SlyA's ability to activate acid resistance pathways. However, the ability of glutamate to rescue AR could be explained in multiple ways. One possibility is that glutamate directly fuels the GDAR enzymatic pathway, compensating for reduced transcriptional activation of GDAR genes. Alternatively, the results could indicate that the GDAR pathway itself remains functional, as it is rescuing AR when glutamate is added, hence IpgC could be influencing AR through a separate, non-GDAR acid resistance mechanism. Supporting this idea, our proteomic analysis revealed enrichment in the active state of Formyl-coenzyme A transferase (Frc), a metabolic enzyme associated with oxalate-dependent acid resistance and previously shown to be downregulated in *slyA* mutants (Fontenot et al., 2013; Zhang et al., 2018). This suggests that IpgC may affect AR by modulating control over alternative stress response pathways, with glutamate supplementation compensating by activating GDAR.

The identification of both known interactors and novel candidates highlights the effectiveness of TurboID-mediated proximity labeling in mapping the proxisome of IpgC in *Shigella flexneri*. Notably, the increased association of IpgC with T3SS structural and regulatory components suggests a dual role in mediating interactions with both the secretion machinery and transcriptional regulators in a state-dependent manner. Furthermore, the identification of potential links to SlyA and acid resistance raises the possibility that IpgC contributes to this pathway, which is crucial for *Shigella* survival in the host environment. These findings lay the groundwork for further exploration of the identified candidates, which provides novel insights into IpgC's functions beyond its established roles in the T3SS.

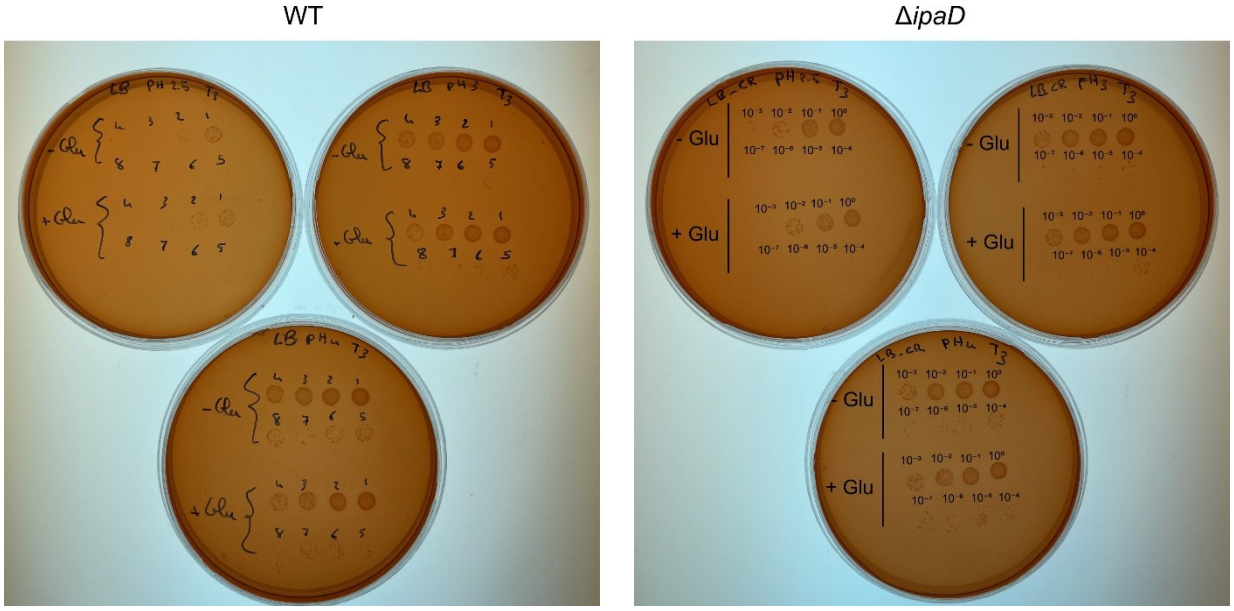


Figure 3.9. Survival of *Shigella* strains in their active and inactive Type III secretion states after acid exposure at different pH levels, with and without glutamate.
Serial dilutions of WT and Δ*ipaD* strains (1 to 10^{-7}) incubated in acidic LB pH 2.5-4 for 2 hours (T3) in the presence and absence of glutamate (50 mM).

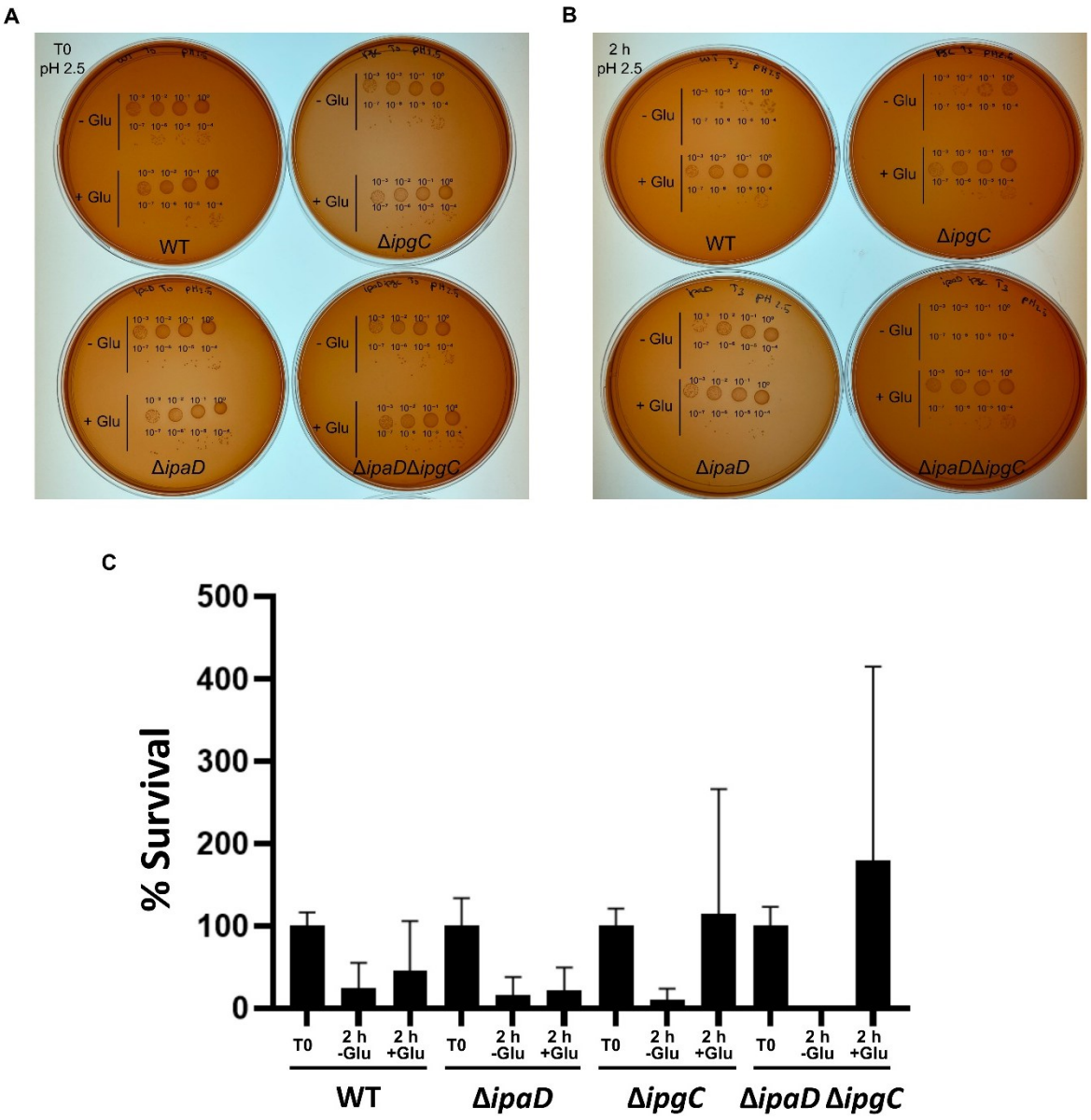


Figure 3.10. IpgC absence increases acid sensitivity in *Shigella flexneri* when the T3SS is active.

Serial dilutions of the different bacterial strains incubated in acidic LB pH 2.5 at (A) T0 and (B) after 2 hours in the presence and absence of glutamate (50 mM). (C) Survival and growth of the four strains after two hours incubation with acidic LB (pH 2.5) are expressed as a percentage relative to their initial population at T0. Values above 100% indicate bacterial growth, while values below 100% reflect a decline in survival. Each data point represents the mean percentage (n = 3) for growth in the 10^{-4} dilution. Spot intensities were measured using ImageJ and statistical significance was determined using one-way analysis of variance. Error bars represent the standard deviation of the mean.

3.4 Discussion

Understanding the molecular landscape and proximal protein environment associated with the T3SS in *Shigella flexneri* is essential for deciphering the mechanisms underlying its virulence. In this study, we utilized TurboID-mediated proximity labeling to characterize the proxisomes of the regulatory proteins IpgC and MxiE, which are central to T3SS activation and regulation. IpgC acts as a chaperone and transcriptional co-activator, while MxiE functions as a transcriptional activator; together, they coordinate the expression of late T3SS effectors in response to host cell contact. It is important to note that proximity labeling identifies the proxisome, a collection of proteins in the immediate vicinity of the bait protein, which may include direct and indirect interactors, as well as non-interacting proximal proteins. These proteins provide insight into the local cellular context and processes surrounding the bait under specific conditions, rather than defining a strict interactome. Building on the previous chapter, where we confirmed TurboID's effectiveness in rapidly biotinylating cytoplasmic proteins in *S. flexneri*, we fused TurboID to IpgC and MxiE to evaluate its capacity to label proximal cytosolic proteins under both active and inactive T3SS conditions. This strategy enabled us to capture dynamic changes in the local protein environment of IpgC, offering insights into how its proxisome composition shifts in response to T3SS activation.

When selecting a plasmid backbone for our constructs, we prioritized expression efficiency and modularity. We chose a pUC18 derivative, which carries a high-copy origin of replication known to enhance gene expression. This decision was supported by findings from the pUdOs study (Manigat et al., 2024), which showed that plasmids with a pUC origin (e.g., pUdO1c) produced approximately 4.8-fold higher GFP fluorescence, reflecting stronger reporter gene expression, compared to those with a p15A origin (e.g., pUdO2c). Since pUC18 and pUdO1c share the same high-copy origin, while pSU-based plasmids use the lower-copy p15A origin, this comparison supports our decision to prioritize expression efficiency by using a pUC-based backbone. We paired this backbone with the *rpsM* promoter, a well-established constitutive promoter that drives the expression of the ribosomal protein S13 (Valdivia & Falkow, 1996). This configuration enabled the generation of TurboID fusions with IpgC and MxiE. Given the uncertainty about how fusion orientation might affect the function of IpgC and MxiE, their localization, and interaction patterns, we constructed two variants of each fusion, placing TurboID at either the N- or C-terminus. By employing this strategy, we were able to

observe differences in expression and complementation depending on the positioning of the biotin ligase. Both IpgC-TurboID and TurboID-IpgC were expressed at levels comparable to endogenous IpgC. Although some degradation was observed in TurboID-IpgC, both fusion constructs successfully restored *ipaH* expression in Δ*ipaD* Δ*ipgC* mutants. Interestingly, a slight increase in IpaH levels was detected in the off-state, suggesting that the fusion protein may have altered IpgC's interactions with its partners IpaB and IpaC, leading to enhanced IpgC-MxiE complex formation and increased *ipaH* transcription, even in the absence of T3SS activation. Such subtle changes are expected with fusion proteins, as the addition of a tag or protein can alter bait proteins stability or their interactions. Nevertheless, IpgC-TurboID fusions retained functionality and effectively complemented *ipgC* deletion strains. In contrast to IpgC, evaluating MxiE-TurboID fusions expression relative to the native MxiE levels in the WT strain was not possible due to the lack of an MxiE specific antibody. Detection of MxiE-TurboID fusions relied on the Myc tag, which introduced uncertainty about whether the fusion proteins were expressed at levels comparable to native MxiE. Consequently, we could not assess potential discrepancies in expression that might affect MxiE's functionality or interaction profile. However, we were still able to evaluate *mxiE* complementation using *ipaH* expression as a functional readout. MxiE-TurboID fusions displayed differences in stability and functionality. TurboID-MxiE remained stable and successfully complemented *ipaH* expression in the active strain, whereas MxiE-TurboID exhibited significant degradation and failed to restore *ipaH* expression under activating conditions. Like the IpgC fusions, a slight increase in IpaH levels was observed in the Δ*mxiE* off-state, likely due to MxiE overexpression or reduced inhibition by OspD1. Although the IpgC-TurboID and TurboID-MxiE fusions displayed the best preservation of the native function of the bait proteins, we proceeded with all fusion constructs for mass spectrometry analysis to capture a broader range of proximal proteins and potential interactions. This approach ensured that any possible differences in labeling due to TurboID positioning would not limit the scope of our findings.

Using the raw mass spectrometry data, we generated heatmaps to compare the biotinylated proteomes across replicates of the IpgC and MxiE fusion constructs. For IpgC, samples consistently clustered by T3SS activation state, regardless of whether TurboID was fused at the N- or C-terminus. This suggests that IpgC's proximal protein environment is primarily shaped by the activation status of the T3SS, rather than the orientation of the fusion.

In contrast, MxiE replicates displayed inconsistent clustering, with some grouping by T3SS state and others by fusion orientation. This variability reflects the biotinylation assay results, where MxiE-TurboID fusions produced patterns nearly indistinguishable from free TurboID and showed minimal differences between T3SS on and off states. Furthermore, as mentioned before, MxiE complementation was only partial, and its expression levels could not be directly compared to native protein levels due to the absence of a specific antibody. This limitation, combined with the indistinct biotinylation patterns and inconsistent clustering behavior, suggests that the fusion of TurboID to MxiE may have altered its function or interaction profile. These observations reinforce the importance of using multiple fusion orientations and validating expression levels when applying proximity labeling.

To refine the mass spectrometry dataset, we applied the SAINT algorithm, a statistical tool valuable in proximity labeling studies due to the inherently high background and promiscuous labeling associated with these techniques (Gingras et al., 2019; Mannix et al., 2019; Olson et al., 2019; Saner et al., 2025). SAINT compares proteins detected in IpgC and MxiE TurboID fusion samples to those found in empty vector and promiscuous TurboID controls, assigning confidence scores that help filter out statistically insignificant hits (Choi et al., 2011). This approach allowed us to focus on proteins reliably enriched under our experimental conditions and compare the proxisomes of IpgC and MxiE in both active and inactive T3SS states. By reducing the initial dataset by over 95%, SAINT yielded 78 and 69 high confidence enriched proteins for IpgC and MxiE, respectively. While SAINT analysis helped refine our dataset by filtering out statistically insignificant hits, it also underscored the importance of accounting for biological variability that could still lead to false-positive identifications. One major challenge was the dynamic nature of the T3SS system, which undergoes significant shifts in protein localization and concentration between its active and inactive states. These changes, driven by differential expression and secretion of T3SS substrates, complicated the interpretation of proximity-based enrichment. A clear example of this challenge was OspD1, which was unexpectedly enriched in the on-state in MxiE proxisome, despite predictions of off-state enrichment. This discrepancy stems from its cytoplasmic accumulation in the off-state, which led to high background biotinylation in the free TurboID control and ultimately its exclusion from the MxiE proxisome dataset. In contrast, during the on-state, OspD1 is secreted, reducing its cytosolic concentration and detectability. This dynamic affected the fold-change calculations,

which compare both experimental and control samples across T3SS states. More broadly, T3SS substrates fall into two categories that contribute to false positives: those that accumulate in the cytoplasm prior to secretion (e.g., OspD1, IpgD, IpaA), and those whose expression is upregulated upon T3SS activation (e.g., IpaHs, VirA, OspE1). In both cases, apparent enrichment reflects changes in cytosolic concentration rather than true proximity to bait proteins. This also explains the modest enrichment of MxiE in the IpgC proxisome under T3SS-on conditions, despite its known interaction with IpgC. These observations underscore the importance of contextualizing proximity labeling data within the chosen biological system, especially when secretion or compartmentalization plays a major role.

In addition to the biological variability introduced by the T3SS system itself, several technical factors also influenced our analysis. To use IpgC and MxiE as bait proteins, we employed deletion strains lacking the respective genes and complemented them with TurboID fusion constructs. This approach introduced complexity, as the expression and function of the fused proteins may not fully replicate those of their endogenous counterparts, particularly given the central roles of IpgC and MxiE in T3SS regulation. To mitigate this, we generated both N- and C-terminal TurboID fusions to maximize the likelihood of retaining native functionality. To mimic infection and activate the T3SS, we used Δ*ipaD* and *ipaB4* knockout strains. However, this design required the exclusion of IpaD and IpaB from our analysis, as their absence in the active strains would artificially increase their fold change in favor of the inactive condition, potentially misclassifying them as off-state-specific enriched proteins. Another important consideration in our analysis was the background biotinylation occurring in the *Shigella* cytoplasm. TurboID is constitutively active almost immediately after its translation, resulting in non-specific labeling, particularly of highly abundant proteins involved in translation, such as ribosomal proteins. Also, TurboID exhibits biotinylation activity even before exogenous biotin is added to the media. Our biotinylation assays confirmed this early activity, showing that 32-39% of total biotinylation in our IpgC-TurboID samples occurred prior to the supplementation of the medium with an additional 50 μM biotin. Finally, we accounted for naturally biotinylated proteins in the cytoplasm, such as BccP, which is biotinylated by the endogenous BirA enzyme under normal cellular conditions, as well as enzymes involved in biotin biosynthesis, including BioB and BioD.

Our mass spectrometry data revealed that the IpgC proxisome was notably more diverse in the T3SS on-state. This increased diversity may reflect a shift in IpgC's functional availability: in the off-state, IpgC is typically sequestered by its cargo proteins, IpaB and IpaC, limiting its interaction with other cellular components. Upon T3SS activation, these cargos are secreted, potentially freeing IpgC to interact with MxiE or diffuse throughout the cytoplasm, where it may engage in additional, yet-to-be-characterized roles. In contrast, MxiE showed similar protein enrichment in both states. IpgC-TurboID fusions dataset successfully identified well-known binding partners of IpgC, such as IpaC and IpaB, enriched in the off-state, and MxiE in the on-state, all detected with high confidence, supporting the validity of the experimental approach and IpgC functionality. In contrast, the MxiE-TurboID dataset did not identify IpgC, MxiE's known binding partner, as a high-confidence enriched protein. Consequently, our analysis focused more on the IpgC dataset, which aligned more closely with established T3SS regulatory models. In addition to the well-characterized interactors that validate our biotinylation assay, IpgC-TurboID dataset revealed enrichment in other T3SS associated proteins. Among these, SpaO (Spa33) and SpaL (Spa47), which are both core components of the cytoplasmic sorting platform. This platform orchestrates the hierarchical recruitment and secretion of T3SS substrates. Spa33 has been proposed to participate in effector shuttling by interacting with effectors or chaperone-effector complexes, guiding them to the export apparatus. Its enrichment in the IpgC proxisome could also reflect an interaction with this chaperone or transient proximity during the shuttling process, particularly when IpgC is bound to IpaB and IpaC (Marcos-Vilchis et al., 2025; Wimmi et al., 2024). SpaL, an essential ATPase, energizes the translocation of effectors into host cells. Its homologs, such as EscN in EPEC, FliI in *S. enterica* flagellar assembly, HrcN in *Xanthomonas campestris*, and InvC/SsaN in *Salmonella enterica* (T3SS-1 and T3SS-2), have been shown to interact with chaperone-effector complexes (Akeda & Galán, 2005; Gauthier & Finlay, 2003; Lorenz & Büttner, 2009; Majewski et al., 2019; Y. Yoshida et al., 2014, p. 2). Therefore, SpaL's enrichment near IpgC may similarly reflect transient interactions with the chaperone-effector complex prior to secretion. Additionally, PspA, a component of the phage shock protein (Psp) stress response system, was enriched in the IpgC-TurboID dataset. Although no direct interaction between PspA and IpgC has been described, its presence may reflect a functional link between stress response and T3SS activity. In *Yersinia enterocolitica*, membrane-bound components of the Psp system (PspB and PspC), which are co-transcribed with PspA,

have been implicated in alleviating membrane stress induced by T3SS activation, though their mechanism remains unclear (Darwin, 2013; Horstman & Darwin, 2012).

Beyond T3SS-associated proteins, the IpgC proxisome in the on-state also included six transcription factors: VirF, KdgR, CysB, ModE, McbR and SlyA. Their presence may reflect direct or indirect interactions with IpgC, or co-localization due to spatial proximity during transcriptional events involving the IpgC-MxiE complex. VirF is particularly notable due to its role in the VirF/VirB/MxiE transcriptional cascade, in which MxiE and IpgC participate in a negative feedback loop that downregulates VirB, thereby modulating the expression of virulence genes during infection by interfering with VirF's activation of the *virB* promoter (McKenna et al., 2022). This regulatory mechanism may explain VirF's presence in the IpgC proxisome. Other transcription factors detected may also have functional relevance to T3SS regulation. KdgR, a member of the IclR family, has been implicated in T3SS regulation in *Xanthomonas oryzae* (Lu et al., 2011; Molina-Henares et al., 2006). While CysB, a LysR-type transcriptional regulator, has been associated with T3SS gene expression in *Pseudomonas aeruginosa* and other species (M. Chen et al., 2022; Song et al., 2019). ModE is a transcriptional regulator that controls expression of the molybdate-specific transport system (modABCD) in *E. coli*, including regulation of *modA*, which encodes the molybdate-binding protein. In *P. aeruginosa*, *modA* mutant has been associated with reduced virulence, likely due to impaired molybdenum metabolism, which is essential for host adaptation. Although no direct connection between this protein and the T3SS has been established, ModA has been reported to be secreted via the H2-type VI secretion system (H2-T6SS) in *P. aeruginosa* (Grunden et al., 1996; Wang et al., 2021).

A particularly intriguing subset of transcription factors enriched in the IpgC proxisome under T3SS on-conditions includes SlyA and McbR, both well-known regulators of stress response. Their proximity to IpgC suggests a potential link between virulence regulation and environmental adaptation, especially acid resistance (AR). Both McbR and SlyA play central roles in mediating stress adaptation and survival in acidic environments, regulating genes such as *rmf* (ribosome modulation factor), which promotes ribosome dimerization into 100S particles, preserving ribosomes during stationary phase and facilitating rapid recovery from nutrient stress (Sebastian & Finkel, 2022; H. Yoshida et al., 2018). They also regulate *gadX*, a key activator of the glutamate-dependent acid resistance (GDAR) system (Tucker et al., 2003; H.

Yoshida et al., 2018). SlyA directly activates GDAR by binding to the promoter region of the glutamate decarboxylase *gadA* (Zhang et al., 2018). Notably, RNA sequencing data from *S. flexneri* strains with an active T3SS revealed moderate upregulation (1.3 to 2-fold) of several acid resistance genes, including GDAR genes (*gadA*, *gadB*, *gadC*) and *hyaB*, which encodes a subunit of the hydrogenase 1 enzyme complex involved in hydrogen uptake and previously shown to be downregulated in *slyA* mutants (Silué et al., 2020; Zhang et al., 2018). While these findings suggest a potential interaction between T3SS activity and acid resistance, the modest fold changes warrant further investigation to determine their physiological relevance. This connection is further supported by our phenotypic analysis. The T3SS-active double knockout strain (Δ*ipaD* Δ*ipgC*) exhibited severe acid sensitivity, with minimal to no growth following two hours of low pH exposure. This phenotype was markedly more pronounced than in either single mutant, and particularly striking compared to the Δ*ipgC* strain alone, which showed only moderate impairment in acid resistance under glutamate-deprived conditions. These observations suggest that IpgC may contribute to acid resistance in a T3SS-dependent manner, potentially through regulatory functions. The rescue of growth upon glutamate supplementation across all strains indicates that the GDAR system remains at least partially functional and capable of alleviating acid stress when glutamate is available. However, the differential acid sensitivity observed in the absence of IpgC, especially under low-glutamate conditions, points to a more complex mechanism. Proteomic analysis identified Formyl-coenzyme A transferase (Frc) in proximity to IpgC under T3SS-on conditions. Frc homologs in *E.coli* have been implicated in oxalate-dependent acid resistance and maintaining pH homeostasis during acid stress. Frc was previously found to be downregulated in *slyA* mutants and proposed to be under SlyA control in regulatory models (Fontenot et al., 2013; Zhang et al., 2018), though direct regulation has not been conclusively demonstrated. Given that both the GDAR genes and potentially Frc may be influenced by SlyA, we hypothesize that IpgC could act as a supporting or regulatory partner that facilitates SlyA-mediated transcriptional activation during acid stress. Although few studies have explored the interplay between SlyA and the T3SS in *S. flexneri*, one report showed that exogenous expression of *slyA* could compensate for the loss of *virB*, restoring virulence-associated phenotypes (Weatherspoon-Griffin & Wing, 2016). However, these findings were based on overexpression from high-copy plasmids, which may not accurately reflect native regulatory interactions. This underscores the need for deeper investigation into the potential role

of IpgC in modulating SlyA activity, particularly within the context of T3SS-dependent virulence and stress adaptation.

To further assess the likelihood of IpgC interactions with transcription factors detected in its proxisome, we employed AlphaFold to predict protein-protein interactions (Jumper et al., 2021; Varadi et al., 2022). Despite testing various oligomeric states, confidence scores were consistently low. To better interpret these results, we used IpgC's known binding partners, IpaB and IpaC, as controls. Surprisingly, even these well-characterized interactions yielded scores within an uncertainty zone. In contrast, the prediction for the IpgC-MxiE complex was notably better, falling near the confidence threshold and aligning with their known functional relationship. This pattern suggests that AlphaFold may struggle to model interactions involving highly disordered proteins, such as T3SS effectors like IpaB and IpaC, which contain coiled-coil motifs and disordered N-terminal regions essential for secretion (Gazi et al., 2009). These regions are frequently truncated or unresolved in crystallographic studies due to their flexibility (Gazi et al., 2009; Lee et al., 2004; Lilic et al., 2006). Unfortunately, they seem also detrimental to the accuracy of structural prediction. Transcription factors like SlyA and McbR are structurally more stable, yet their low prediction scores raise questions about the nature of their proximity to IpgC. Taken together, these results suggest that while AlphaFold may be capable of detecting plausible interfaces under favorable structural conditions, as seen with MxiE, it may not reliably capture interactions involving flexible or transient complexes. Therefore, low confidence scores should be interpreted cautiously and supplemented with experimental validation. Given the diversity of transcription factors detected in the IpgC proxisome, it is unlikely that IpgC directly binds all of them. However, our observations, particularly the acid sensitivity phenotype of the Δ*ipaD* Δ*ipgC* strain and the presence of AR regulators in the proxisome, suggest a possible relationship between IpgC and SlyA. However, AlphaFold predictions yielded low confidence scores for this interaction, raising questions about the nature of their association. Several models could explain this proximity: IpgC may interact with SlyA indirectly, through shared complexes or signaling pathways; it may interact directly, either in its free form or in complex with its known partner MxiE; or it may not bind SlyA at all, but instead co-localize with it and the other transcription factors due to spatial proximity during DNA binding events involving the IpgC-MxiE complex. While our data does not establish a direct interaction between IpgC and SlyA,

they are consistent with a model in which IpgC influences acid resistance through regulatory networks that include SlyA.

In addition to transcriptional regulators, our analysis identified a number of housekeeping proteins, predominantly metabolic enzymes, enriched in proximity to IpgC, particularly under T3SS-active conditions. While these proteins are typically abundant in the cytoplasm, their variable detection patterns suggest potential metabolic changes associated with T3SS activation. The functional relevance of many of these enzymes in the context of T3SS remains uncertain. However, three proteins, ClpX, ErpA, and GltA, have previously been implicated, albeit indirectly, in T3SS-related processes. ClpX, a component of the ClpXP protease complex, has been linked to T3SS regulation in *Yersinia pestis* (Jackson et al., 2004). ErpA, an iron-sulfur (Fe-S) cluster carrier, is connected to T3SS regulation via the Fe-S sensor IscR, which modulates T3SS expression in several species such as *Salmonella enterica* and *Yersinia pseudotuberculosis* (Miller et al., 2014; Vergnes et al., 2017). In *Shigella*, deletion of the *isc* operon regulated by IscR, resulted in an avirulent phenotype, indicating a role for this pathway in virulence (Runyen-Janecky et al., 2008). GltA, a citrate synthase, has been associated with T3SS modulation in *Pseudomonas aeruginosa*, where its deletion led to increased T3SS gene expression and enhanced antibiotic tolerance, potentially via the stringent response (H. Chen et al., 2023). Additionally, GltA was part of a highly connected STRING network involving other metabolic enzymes detected in our dataset, including AtpG, GcvP, GdhA, Pgk, Ppc, and PurF. Gene ontology (GO) analysis of this group revealed six significantly enriched biological processes, pointing to a coordinated shift in core metabolic activities. Although direct physical interactions between these enzymes and IpgC have not been established, their co-enrichment under T3SS-on conditions raises the possibility that virulence activation may be accompanied by metabolic remodeling. This potential coupling underscores how the metabolic state of the cell could be shaped by pathogenic mechanisms, reflecting a broader integration of metabolism and virulence regulation.

In summary, this study advances our understanding of the molecular landscape surrounding IpgC in *Shigella flexneri* by applying TurboID-mediated proximity labeling to map its proxisome under varying T3SS activation states. The identification of well-established interactors, such as IpaB, IpaC, and MxiE, and novel proximal proteins, highlights the utility of

TurboID in capturing dynamic protein environments within the bacterial cytoplasm. Beyond reaffirming IpgC's established roles in effector chaperoning and transcriptional co-activation, our data reveal previously unrecognized associations with T3SS structural components, transcriptional regulators, and cellular processes. Notably, the enrichment of acid resistance regulators such as SlyA, alongside phenotypic evidence of acid sensitivity in Δ*ipaD* Δ*ipgC* strains, suggests a link between virulence regulation and environmental adaptation. This connection points to a broader role for IpgC in coordinating stress responses during host infection. Furthermore, the detection of metabolic enzymes in proximity to IpgC under T3SS-active conditions implies that virulence activation may be coupled with metabolic remodeling. Collectively, these findings validate TurboID as a proximity labeling tool in the *Shigella* cytoplasm and lay the groundwork for future investigations into the extended regulatory functions of IpgC. Exploring the roles of newly identified interactors may uncover novel functions of IpgC and provide deeper insights into its contributions to *Shigella* pathogenesis.

Chapter 4 : Conclusion

Shigella flexneri remains a serious public health concern, with rising antibiotic resistance further complicating treatment efforts. Central to its pathogenicity is the Type III Secretion System, a syringe-like apparatus that injects effector proteins directly into host cells. These proteins manipulate host cell processes to aid bacterial invasion, dampen immune responses, and ensure intracellular survival. Although the structure and function of the T3SS have been studied extensively, key aspects of its regulation and dynamic protein-protein interactions are not fully understood. Traditional techniques often fall short in detecting the transient and low-affinity interactions that govern this system especially in the complex and crowded bacterial cytoplasm. To address this gap, we applied proximity-dependent biotinylation using TurboID to study the proxisome of the T3SS chaperone IpgC, that is, the set of proteins near IpgC, which may include direct interactors, indirect partners, and colocalized proteins. TurboID demonstrated rapid labeling kinetics compared to other biotin ligases we tested. In addition, despite TurboID being derived from *E. coli BirA* (Branon et al., 2018), our data suggest it had undergone enough structural divergence during its engineering to exhibit reduced affinity for native biotin-related partners such as BccP and the biotin operon. This loss of native functions and increased promiscuity made it well-suited for cytoplasmic expression in *Shigella*. From this approach, we were able to reconstruct a snapshot of the IpgC proxisome. The discovery that IpgC's proxisome diversifies when the T3SS is active suggests that IpgC may have broader regulatory or functional roles beyond its recognized chaperone activity. A graphical summary of these findings is provided in Figure 4.1, showing the differential composition of the IpgC proxisome across T3SS activity states.

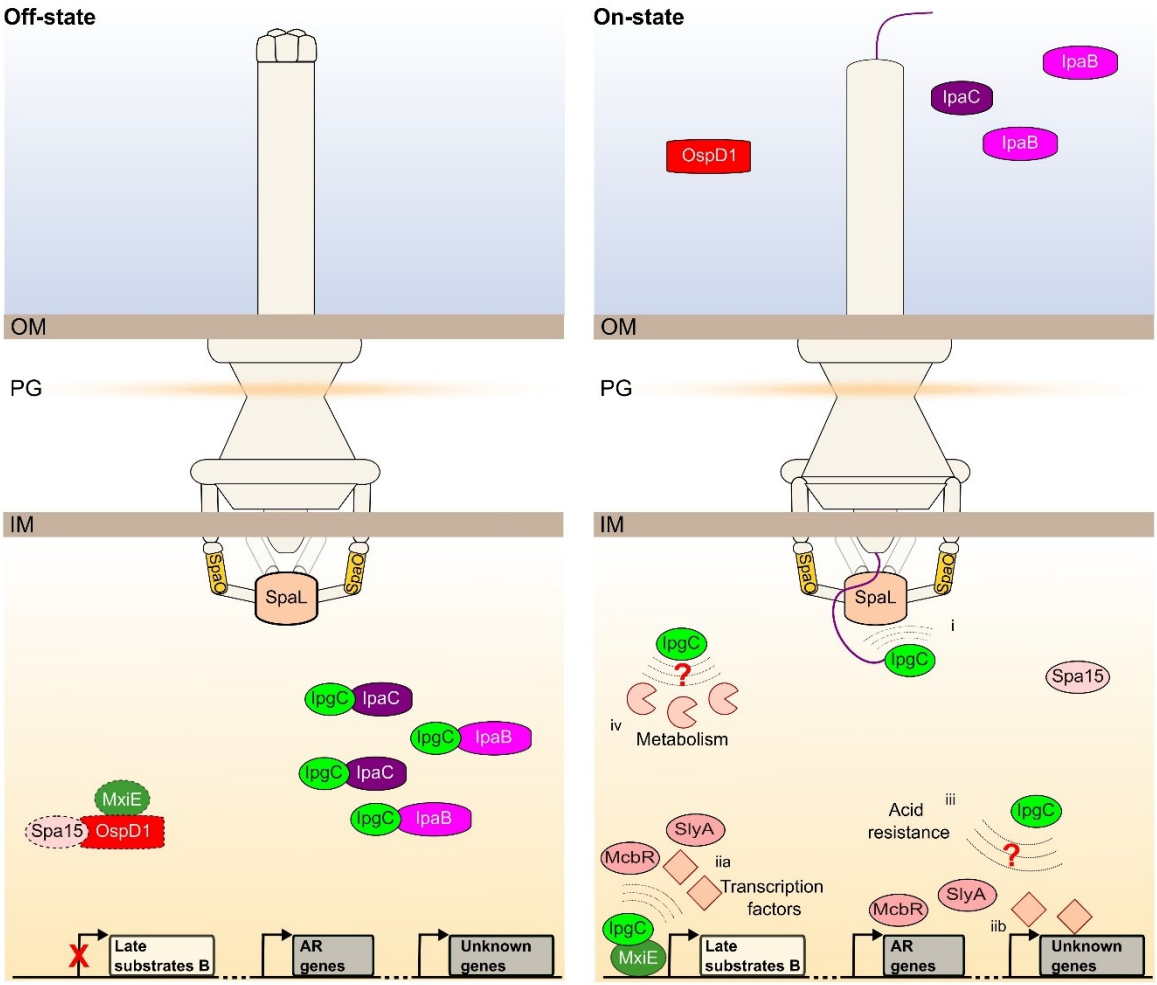


Figure 4.1. The dynamic composition of the IpgC proxisome in *Shigella flexneri* under T3SS on and off-states.
This model is proposed based on our TurboID proximity labeling findings, which extend the established T3SS signaling model. It is known that in the off-state, MxiE and IpgC form inhibitory complexes with Spa15-OspD1 and IpaB or IpaC, respectively. However, in the on-state, IpgC transports its cargo, IpaBC, to the T3SS sorting platform and becomes available for interaction with MxiE. This interaction leads to the formation of a transcriptional activation complex that drives the expression of MxiE box-containing genes, encoding T3SS late substrates B (Parsot et al., 2005). Our TurboID data support the model that IpgC forms complexes with IpaBC and MxiE. However, interactions between MxiE, OspD1, and Spa15 (shown as dashed lines) could not be confirmed due to technical limitations in the MxiE-TurboID experiments. Furthermore, the increase in free cytoplasmic IpgC concentration upon cargo release in the on-state leads to an enrichment of the IpgC proxisome, suggesting that MxiE does not fully titrate all the IpgC released by the secretion of IpaB and IpaC. The proximity of IpgC to the ATPase SpaL and the cytoplasmic sorting platform component SpaO during cargo delivery likely explains their presence in the IpgC proxisome in the on-state (i). IpgC may also interact with these components, either alone or in a complex with its cargo, as other chaperones have been

shown to interact with SpaL and SpaO. Additionally, we predict that IpgC diffuses to other areas of the cell, where it colocalizes with transcription factors through DNA binding in complex with MxiE (iia) or in its free form, with or without DNA binding (iib). Our acid resistance assay suggests that IpgC may also influence acid resistance, either through interactions with AR transcriptional regulators such as SlyA and McbR (iia-b), or through other unidentified mechanisms (iii). The mobility of IpgC may also allow it to interact with metabolic enzymes in other cellular regions (iv). Our findings suggest potential new connections between T3SS signaling and cellular processes mediated by IpgC, to be further explored. This schematic is adapted from reference (Haidar-Ahmad et al., 2024). OM: outer membrane; PG: peptidoglycan cell wall; IM: inner membrane

Proximity labeling using biotin ligases has proven effective in the cytosolic environment of *Shigella flexneri*, as demonstrated by our mapping of the IpgC proxisome using TurboID. This success highlights the potential of proximity labeling to capture dynamic protein interactions within the T3SS, where many regulatory and effector-related processes occur. Furthermore, extending this technique to other cytosolic chaperones may provide insight into how effectors are recruited and delivered to the injectisome, and whether these chaperones physically interact with apparatus components or the sorting platform. For example, SpaL and its homologs within the ATPase complex have been shown to associate with chaperone-effector pairs in other species (Akeda & Galán, 2005; Gauthier & Finlay, 2003; Lorenz & Büttner, 2009; Majewski et al., 2019; Y. Yoshida et al., 2014), suggesting that additional, yet unidentified, interactions may exist among T3SS components.

The T3SS spans several structural layers including the cytoplasm, inner membrane, and periplasm, requiring tightly coordinated assembly and interaction of its components across these domains. By fusing biotin ligases, such as TurboID or smaller variants like miniTurbo and BioID2, to components of the secretion apparatus, particularly sorting platform proteins, we can capture localized protein enrichments and interactions that occur at distinct stages of T3SS assembly and activation. This strategy is supported by previous structural studies of the *Salmonella Typhimurium* T3SS, where fusions with fluorescent proteins such as GFP were successfully applied to sorting platform components including SpaO, OrgA, OrgB, and InvC (Spa33, MxiK, MxiN, and SpaL respectively in *Shigella*), without disrupting their localization or function (B. Hu et al., 2017). These fusion constructs enabled *in situ* mapping of the injectisome architecture using cryo-electron tomography, demonstrating that T3SS proteins can tolerate

tagging at either terminus. While biotin ligases like TurboID are somewhat larger than typical fluorescent proteins (35 kDa vs. ~27 kDa), their successful use in our bacterial system, and the availability of smaller ligases, suggest that similar fusion strategies are feasible for T3SS components. This opens the door to using proximity-dependent labeling across different compartments of the T3SS, providing a better understanding of this system and its dynamic changes.

We envision several promising research directions that can build upon the foundation established in this study. One area of interest involves the metabolic enzymes identified in the IpgC proxisome in the on-state of the T3SS. Although direct physical interactions between IpgC and these enzymes remain unconfirmed, their co-enrichment suggests a potential link between virulence activation and metabolic remodeling. Future studies could investigate whether this association reflects metabolic shifts that accompany T3SS activation, or whether metabolic state and virulence regulation are coordinated through shared stress response pathways, energy demands, or metabolite signaling. Additionally, integrating proximity labeling with metabolomics and transcriptomics could reveal how metabolic pathways and T3SS activation are coordinated. Proximity labeling experiments using TurboID fusions to metabolic regulators or stress sensors may uncover transient or condition-specific interactions. Investigating whether metabolic enzymes cluster near T3SS components or sorting platforms during activation may also shed light on spatial coordination between metabolism and virulence.

Another key area involves investigating the observed proximity between IpgC and transcriptional regulators such as SlyA and McbR. Preliminary data from our acid resistance assays suggest that IpgC may play a role in modulating acid stress responses, potentially through interaction with SlyA in the T3SS on-state. This opens exciting avenues for further validation, which we have already begun to pursue using deletion mutants and qPCR-based expression profiling. Additionally, we are employing a transcription-based secretion activity reporter (Manigat et al., 2024) to monitor the expression of genes involved in the GDAR system in *ipgC* knockout strains exposed to acidic environments.

Beyond functional validation, biotin ligases such as TurboID offer exciting potential for dissecting stress-induced changes in the T3SS network. By fusing biotin ligases to IpgC or structural components of the secretion system and applying mass spectrometry to analyze

labeling patterns under acid stress, we could determine whether stress response regulators like SlyA are differentially recruited. This approach may uncover previously unrecognized regulators that are specifically involved in maintaining T3SS activity under hostile environmental conditions. Moreover, it could help reveal whether acid stress alters T3SS assembly or effector dynamics in a state-dependent manner. However, using proximity labeling under acidic conditions presents significant challenges. A study in *E. coli* has shown that while glutamate decarboxylases elevate intracellular pH during acid stress, the internal environment can still drop to between pH 4.2 and 4.7 (Richard & Foster, 2004). Such low pH conditions could influence both the structure and function of the biotin ligases as well as their target proteins. Even if the proteins themselves remain stable, the efficiency of lysine biotinylation may be compromised due to protonation of amine groups, potentially reducing labeling yield. Therefore, extensive testing and optimization will be required to ensure that proximity labeling under acidic conditions provides reliable and interpretable data.

As with any emerging methodology, optimizing the application of proximity labeling in bacteria, particularly within a complex system like the T3SS, is an evolving process. Given the scarcity of studies utilizing biotin ligases in the bacterial cytoplasm, our work serves as a foundation for further methodological refinement. One area for improvement is the protein extraction step. In our study, a mild detergent was used during cell lysis, which, while effective for solubilizing cytosolic proteins like the known IpgC binding partners, may have limited the recovery of hydrophobic membrane-associated T3SS components. Future protocols could explore alternative detergents or mechanical techniques better suited for membrane proteins, potentially expanding the coverage of T3SS components. Another critical consideration is the relative size of TurboID (~35 kDa), which is nearly twice the size of bait proteins like IpgC (~18 kDa) and MxiE (~25 kDa). This size imbalance may hinder the natural folding, localization, or interaction dynamics of the bait protein, thereby influencing the quality or specificity of biotinylation. In the case of MxiE, this phenomenon may be the source of the apparent dysfunction of its TurboID fusions. Despite its established role in T3SS regulation, MxiE-TurboID fusions failed to significantly enrich interacting proteins under either T3SS-on or off conditions. Moreover, biotinylation patterns from MxiE-TurboID were indistinguishable from those of free TurboID, suggesting limited functional engagement with nearby proteins. These findings, combined with reduced expression and incomplete functional complementation

observed in our assays, indicate that the fusion may have disrupted MxiE's stability or activity. To address this, future studies could try incorporating flexible linkers of varying lengths between the biotin ligase and the bait protein to reduce structural interference and maintain proper folding of the bait protein. Optimizing linker design may not only enhance the functional integrity of the fusion protein but also improve the spatial reach of biotinylation.

In addition to the constraints posed by the fusion with the biotin ligase, background biotinylation remains a challenge in bacterial systems, particularly when TurboID is expressed constitutively. In our current setup, biotinylation occurred even prior to the biotin pulse, limiting our control over labeling. To address this, we refined our experimental design by testing different time points and growth conditions to capture more meaningful hits, however, background labeling was still high, and it posed a particular challenge when attempting to visualize biotinylated proteins via fluorescence microscopy. Due to the small size of *Shigella* and the high labeling promiscuity of TurboID, nearly the entire bacterial cell appeared labeled (results not shown), making it difficult to resolve subcellular localization or draw conclusions about specific labeling events. As a result, the approach of visualizing the spatial distribution of the IpgC proxisome was deemed impractical. This observation emphasized the importance of including a free TurboID control in our mass spectrometry experiments, which enabled the subtraction of background signals. A promising strategy to address this issue involves the use of split-ligase systems, such as Split-TurboID that involves a split at residue L73/G74 dividing the enzyme into two fragments that can be reconstituted when brought into proximity by bait-prey interaction or colocalization. While Split-TurboID, similarly to Split-BioID, is less active than its full-length counterpart, it has demonstrated detectable biotinylation after just 30 minutes of biotin incubation, offering improved labeling control with reduced background (Cho, Branon, Rajeev, et al., 2020; De Munter et al., 2017; Kwak et al., 2020; Schopp et al., 2017). Combined with optimized linker designs and improved sample processing protocols, these strategies provide a path toward more refined and interpretable proximity labeling datasets in bacterial systems. Finally, tool refinement remains a critical avenue for expanding the precision of proximity labeling in bacteria. While TurboID proved to be a powerful tool, we also explored other biotin ligases such as BioID2, valued for its smaller size and reduced background activity. Our designing of a quadruple mutant of BioID2 highlighted how small changes can enhance the enzyme performance.

UltraID and MicroID2 are more recently engineered biotin ligases derived from *Aquifex aeolicus*, the same bacterium from which BioID2 was developed, but with molecular weights below 20 kDa, making them some of the smallest proximity labeling biotin ligases available. They are truncated variants of BioID2 with introduced mutations in the biotin binding domain, making them both smaller and more promiscuous, enhancing their utility in compact or sensitive bacterial systems (B. S. Johnson et al., 2022; Kubitz et al., 2022). Interestingly, two of the mutations we introduced (K115 and L67) were also part of the initial screening of UltraID variants developed by another group (Kubitz et al., 2022). In our study K115 was found to increase activity when introduced alone, however, its performance declined when combined with other, more potent mutations. In the UltraID study, the variant containing the K115 mutation exhibited increased activity but showed decreased thermal stability and was ultimately discarded. The L67 mutation, which emerged as the key driver of increased activity in our BioID2 mutant, was not pursued further in the UltraID study due to an unintended stop codon mutation that extended the variant by 22 amino acids (Kubitz et al., 2022). As a result, the true impact of L67 on ligase performance remains unexplored in their system. Given its strong effect in our construct, reintroducing L67 into their optimized variants, such as UltraID, or MicroID2, could be valuable to determine whether it similarly enhances activity without compromising specificity. This presents a compelling direction for future research, where insights from parallel engineering efforts can be cross applied to refine and expand the proximity labeling toolkit for bacterial systems.

In summary, this work demonstrates the feasibility and utility of using TurboID for proximity labeling in the cytoplasm of *Shigella flexneri*, offering new insights into the molecular organization and regulatory complexity of the Type III Secretion System. By applying this approach to the T3SS chaperone IpgC, we uncovered a dynamic and state-dependent proxisome, suggesting potential roles for IpgC beyond classical effector chaperoning. The findings lay the foundation for exploring how T3SS activity intersects with stress response pathways and other cellular networks. Moreover, this study highlights the importance of refining proximity labeling strategies to overcome technical challenges inherent to bacterial systems. Ultimately, by advancing our understanding of *Shigella*'s virulence machinery at the molecular level, this work contributes to the broader effort of identifying novel therapeutic targets. In light of rising antimicrobial resistance and limited treatment options for shigellosis, such insights are essential

for guiding the development of alternative therapeutic approaches, offering new avenues for intervention beyond conventional antibiotics.

Chapter 5 : Experimental Procedures

5.1 Preface

The experimental procedures described in Sections 5.2 to 5.13 and the corresponding supplemental tables are adapted from a previously published study in mSphere (Haidar-Ahmad et al., 2024), with modifications and additional data where necessary to support the broader scope of this thesis. As the lead author of the original publication, I retain the rights to reuse this content in accordance with the journal's policy and the terms of the Creative Commons Attribution License (CC BY 4.0), which permits reuse provided the original published version is properly cited. Raw data and Significance Analysis of INTeractome (SAINT) results for IpgC and MxiE, along with fold changes of proteins identified in their respective proxisomes, are provided in the online supplementary material of the publication (Haidar-Ahmad et al., 2024). The mass spectrometry data generated in the study are publicly available through the Mass Spectrometry Interactive Virtual Environment (MassIVE) at: ftp://massive.ucsd.edu/v08/MSV000095473/. I would also like to acknowledge the contributions of Kyle Tomaro, for his assistance in writing Sections 5.11 and 5.12, which detail the bioinformatics analyses performed during this work.

5.2 Bacterial strains

We used *Shigella flexneri* serotype 5a strain M90T with acquired streptomycin resistance as the WT strain (Allaoui et al., 1992). Its isogenic mutant strains Δ*ipaD,* Δ*mxiE*, *ipaB4*, and *ipaB4* Δ*mxiE* were a kind gift from Philippe Sansonetti and Claude Parsot (Mavris, Page, et al., 2002; Ménard et al., 1993) (see Table 5.S.1 for strains description). *S. flexneri* was grown on tryptic soy agar (TSA) and tryptic soy broth (TSB). *E. coli* BL21 Rosetta strain (Novagen), a kind gift of Roberto Chica, was grown on Luria-Bertani agar or broth. Media were supplemented with 100 µg/mL ampicillin, 30 µg/mL chloramphenicol, 12.5 µg/mL zeocin, 50 µg/mL gentamicin, or 50 µg/mL kanamycin as appropriate for plasmid selection and maintenance or to confirm the genotype of mutant strains. Δ*ipgC* and Δ*ipaD* Δ*ipgC* were generated by allelic exchange (Datsenko & Wanner, 2000), in the WT and isogenic Δ*ipaD* strains described above, and inspired by the genetic makeup of the original nonpolar Δ*ipgC* mutant (Ménard, Sansonetti,

Parsot, et al., 1994). In brief, the zeocin (*zeo*) resistance cassette was inserted between base pair 128 and 453 of *ipgC*, leaving 42 codons and 5 codons upstream and downstream of the *zeo* resistance cassette, respectively. The *zeo* amplicon was obtained with primers GMO28/29 using the template pZeo/SV2+ (see Table 5.S.2 for primers list). Mutant clones were isolated on TSA supplemented with 12.5 µg/mL zeocin and 0.01% Congo red (Fisher Scientific). The correct insertion of zeo was confirmed by colony screening PCR using primers GMO24/27 and GMO25/26 for the 5′ and 3′ junctions, respectively.

5.3 Plasmids

The coding sequence of *bioID*, *bioID2*, *turboID*, and *miniTurbo* [Addgene plasmids #35700 and #74224, kind gift from Kyle Roux (Kim et al., 2016; Roux et al., 2012); #107169 and #107170, kind gift from Alice Ting (Branon et al., 2018)] were amplified by PCR using primers HMIO 607/609, 11/14, 249/251, and 247/251, respectively, which inserted EcoRI (5′-end), and XbaI restriction sites and a Myc tag (3′-end). The resulting amplicons were digested with EcoRI and XbaI and then subcloned by ligation into a similarly digested pSU2.1 (Silué & Campbell-Valois, 2022)], a derivative of pSU2718 (GenBank: M64731.1), yielding pSU2.1 *bioID* (pNHA1), *bioID2* (pNHA2), *turboID* (pNHA3), or *miniTurbo* (pNHA4). These pSU2.1 derivatives possess a *lac* promoter that is constitutively expressed due to the absence of the *lacI* on the plasmid. Mutagenesis PCR was used to introduce the following mutations into *bioID2* (pNHA2): K19A using primers HMIO390/391, S112A using HMIO392/393, D114A using HMIO394/396, E227A using HMIO397/398, and the double mutation L66A/L67A using HMIO423/424.

To construct the plasmids needed for protein purification, *turboID, bioID, and bioID2* (WT or *bioID2* quadruple mutant), were amplified using HMIO408/409, HMIO404/405 and HMIO406/407. Then they were ligated into BamHI and KpnI of pQE-80L (Qiagen) to yield pQE-80L *turboID* (pNHA5), pQE-80L *bioID* (pNHA6), and pQE-80L *bioID2* (WT pNHA21, or mutant). To obtain pQE-80L *birA* (pNHA7), a G118R mutation was introduced into pNHA6 by mutagenesis PCR with primers HMIO650/651 (Choi-Rhee et al., 2004). Finally, *bccP* was cloned from *Shigella flexneri* M90T by PCR using primers HMIO402/403 and inserted into pQE-80L *MBP-icaR* in place of *icaR* [Silué et Campbell-Valois, unpublished, and (Silué &

Campbell-Valois, 2022)], yielding pQE-80L *MBP-bccP* (pNHA8), using BamHI and PstI restriction sites.

To obtain pSU2.1 *ipgC-turboID-myc*, the *ipgC* amplicon was obtained with primers FXO179/180, digested with BglII and BamHI and ligated through BamHI into pSU2.1 (pNHA15); the *turboID-myc* amplicon was then obtained with primers HMIO250/251 and ligated 3′ of *ipgC* into BamHI-XbaI restriction sites (pNHA18). Similarly, to obtain pSU2.1 *ipgC-bioID-myc*, the *bioID-myc* amplicon was obtained with primers HMIO404/609 and ligated 3′ of *ipgC* into BamHI-XbaI restriction sites (pNHA16). To obtain pSU2.1 *ipgC-bioID2-myc*, the *bioID2-myc* amplicon was obtained with primers HMIO406/14 and ligated 3′ of *ipgC* into BamHI-XbaI restriction sites (pNHA17). Additionally, the following mutations were introduced into *bioID2* in the *ipgC-bioID2-myc* construct (pNHA17): L67A using HMIO69/70, K115A using HMIO75/76, and the double mutation L66A/L67A using HMIO423/424. To test the effect of extending the linkers in IpgC fusion with BioID or TurboID, we used primers HMIO637/638 for *ipgC-L-bioID* (pNHA19), and HMIO639/638 for *ipgC-L-turboID*, in *ipgC-bioID-myc* and *ipgC-turboID-myc* (pNHA20), respectively.

To yield a satisfactory level of expression, *turboID-myc* and *ipgC-turboID-myc* were then subcloned through XhoI and XbaI overhangs into pUC18.1rp (pNHA9), a derivative of pUC18 with the constitutive *rpsM* promoter, yielding pUC18.1rp *turboID-myc* (pNHA10), and pUC18.1rp *ipgC-turboID-myc* (pNHA11). For the pUC18.1 *ipgC-myc* (pNHA22) and pUC18.1rp *ipgC-myc* (pNHA23) controls, *ipgC-myc* was similarly subcloned into pUC18.1 and pUC18.1rp using the same XhoI/XbaI cloning strategy. The MxiE amplicon obtained with HMIO699/701 was swapped with *ipgC* from pUC18.1rp *ipgC-turboID-myc* through ligation of EcoRI and BamHI digested fragments to yield pUC18.1rp *mxiE-turboID-myc* (pNHA13). Finally, the *turboID-myc-ipgC* and *turboID-myc-mxiE* were derived in two steps from pUC18.1rp *turboID-myc*. First, the entry plasmid was obtained by introducing KpnI, SacI, and BamHI restriction sites downstream of *myc* in pUC18.1rp *turboID-myc* by mutagenesis PCR with primers HMIO694/695. Second, *mxiE* and *ipgC* amplicons generated with primers HMIO698/699 and HMIO734/735 were ligated into this entry plasmid using KpnI and BamHI and KpnI and SacI, respectively, yielding pUC18.1rp *turboID-myc-ipgC* (pNHA12) and pUC18.1rp *turboID-myc-mxiE* (pNHA14). Notably, IpgC/MxiE-TurboID-Myc and TurboID-

Myc-IpgC/MxiE protein fusions have a linker between the TurboID and the bait proteins that measures 5 amino acids (GSGGG) and 14 amino acids (SGEQKLISEEDLGT; the Myc tag is underlined), respectively. Therefore, both N-terminal and C-terminal TurboID fusions with IpgC or MxiE have the same molecular weights, i.e. 54 kDa and 66 kDa, respectively. Phusion High-Fidelity DNA polymerase was used to generate all amplicons used for cloning (Thermo Scientific, #F530) and restriction enzymes were purchased from Thermo Scientific. All constructs were verified by Sanger sequencing (Génome Québec). For practical reasons, the main plasmids described above were given short names that are used in the rest of the methods (see Table 5.S.3 for plasmids list).

5.4 Comparison of the different promiscuous biotin ligases by immunoblotting

Shigella flexneri strain M90T WT harboring either pSU2.1, pNHA1, pNHA2 (or its mutants), pNHA3, or pNHA4 (Table 5.S.3) were grown overnight in TSB supplemented with the corresponding antibiotics at 30°C with shaking at 250 rpm. The next day, the overnight cultures were subcultured 1/50 in fresh TSB supplemented with half the regular concentration of antibiotics and incubated at 37°C until they reached an optical density at 600 nm in a 1 cm light path (OD600) ≈ 1. Then, 50 µM biotin (Sigma-Aldrich, B4501) was added to each tube, and they were incubated at 37°C for the indicated time. The volume of cell culture from each strain was adjusted based on the OD600, pelleted, resuspended in Laemmli buffer (1×), and heated for 10 min at 95°C. Samples were run on 12% SDS-PAGE gels and then transferred to a nitrocellulose membrane (Bio-Rad, 1620115) to measure biotinylation of the cell lysate. The membrane was blocked in phosphate-buffered saline (PBS) with 1% bovine serum albumin (BSA) and 0.2% TritonX for 10 min then incubated with 1/25000 streptavidin-horseradish peroxidase conjugate (Cytiva, RPN1231) for 40 min at room temperature followed by three washes of 10 min in PBS. Notably, the endogenous biotinylated protein BccP (MW 17 kDa) masked the signal of the lower active BioID2. This phenomenon is not problematic with the highly active BioID and TurboID. Nevertheless, we generally avoided imaging biotinylated protein with MW < 20 kDa by adjusting the migration time to elute them from the gel or by discarding this section of the membrane. To evaluate protein expression, the same samples were run on separate gels and then transferred to a polyvinylidene difluoride (PVDF) membrane (Bio-Rad, 1620177).

Immunoblotting was performed using 1/10,000 mouse anti-Myc (Genescript, A00704) and 1/1,000 anti-RecA (MBL, ARM191) as the primary antibodies, and 1/20,000 anti-mouse IgG-HRP (Jackson Immunoresearch, 115-035-003) as the secondary antibody. Both nitrocellulose and PVDF membranes were soaked in Clarity Western ECL substrate (Bio-Rad, 170-5061), and images were acquired using a ChemiDoc XRS+ system (Bio-Rad). The activity of the four biotin ligases (Figure 2.9) was assessed by quantification of biotinylated proteins at 10 min, measuring pixel intensities in the region of interest (ROI) from 28 to 150 kDa for each lane ($n = 3$) with ImageJ (version 1.54f). To better quantify the expression and activity of TurboID and miniTurbo, three replicates were repeated on the same membrane and activity was evaluated while considering the expression level of the two biotin ligases and error propagation. Graphs were plotted using GraphPad Prism showing data as mean ± standard deviation. Comparisons of multiple samples were performed using one-way analysis of variance (ANOVA) followed by Tukey's test. Comparisons between TurboID and miniTurbo were performed using unpaired Student's *t*-test.

5.5 Purification of biotin ligases and BccP

pQE-80L-derived plasmids pNHA5-7, pNHA5 (and its quadruple mutant), harboring 6×His-biotin ligase and the plasmid pNHA8 harboring *MBP-bccP* were introduced into BL21 Rosetta *E. coli* strain. Prior to purification, Coomassie-stained SDS-PAGE gels were used to assess the expression and solubility of the four proteins expressed in BL21 Rosetta cells. For each protein, samples were collected from: (i) total cell lysate before IPTG induction, (ii) total cell lysate after IPTG induction, (iii) insoluble fraction after cell lysis and centrifugation, and (iv) soluble fraction after cell lysis and centrifugation. Cultures were grown in terrific broth (TB) and 2xYT for the biotin ligases and *bccP*, respectively, both supplemented with ampicillin. Protein expression was induced with 1 mM isopropyl-β-D-thiogalactoside at OD600 ≈ 0.6 and incubated for 24 hours at 16°C. The cells were then collected by centrifugation and resuspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, 1 mM phenylmethylsulfonyl fluoride [PMSF], 150 µg/mL lysozyme [BioShop, LYS702], 10 µg/mL DNase I [Sigma-Aldrich DN25], 10 mM $MgCl_2$, pH 7.4) for biotin ligases, or (50 mM tris-HCl, 150 mM NaCl, 1 mM PMSF, 150 µg/mL lysozyme, 10 µg/mL DNase I, 10 mM $MgCl_2$, pH 8.0)

for BccP. Cells were sonicated and clarified by centrifugation. The soluble cell lysate was incubated with HisPur resin (Thermo Scientific, 88221) for the Histag, or amylose resin (New England Biolabs, E8021S) for the MBP, at room temperature for 1 hour with end-over-end rotation and then transferred into Pierce Centrifuge Columns (Thermo Scientific, 89868). HisPur resin was then washed with 50 mM sodium phosphate, 300 mM NaCl, 25 mM imidazole, pH 7.4, and proteins were eluted with 50 mM sodium phosphate, 300 mM NaCl, 250 mM imidazole, pH 7.4, and desalted using Zeba Spin Desalting Columns (Thermo Scientific, 89890) in a storage buffer (40 mM Tris-HCl [pH 8.0], 100 mM KCl, 10% glycerol). The same procedure was used with the amylose resin, except the wash (20 mM Tris-HCl [pH 7.4], 200 mM NaCl, 1 mM EDTA) and elution buffer (20 mM Tris-HCl [pH 7.4], 200 mM NaCl, 1 mM EDTA, 10 mM maltose) were modified as indicated. Protein samples were frozen in an ethanol dry ice bath and then stored at -80°C until further use. The use of 2xYT medium instead of TB was crucial to obtain a significant fraction of purified BccP in its unbiotinylated form. However, an additional purification step on streptavidin-agarose beads (Thermo Scientific, 20353) was required to remove biotinylated BccP. Briefly, the eluate from the amylose resin was incubated at room temperature with streptavidin-agarose using an end-over-end rotator to capture holo BccP (biotinylated). The beads were pelleted by centrifugation and the resulting apo BccP-enriched supernatant was used to perform the biochemical assays described below.

5.6 *In vitro* biotinylation assay to compare BccP binding activity of BioID, BioID2 and TurboID

This assay was performed as previously described (Choi-Rhee et al., 2004) by incubating 20 nM of the purified biotin ligases with 2 μM BccP or RNase A (Sigma-Aldrich, R6513) for different time points ranging from 3 min to 24 hours at 37°C in a reaction buffer (40 mM Tris-HCl [pH 8.0], 3 mM ATP, 5.5 mM $MgCl_2$, 100 mM KCl, 1.4 mM β-mercaptoethanol, and 5 μM biotin). BSA (HyClone, SH30574.01) was also tested under the same conditions to compare with RNase A. The reaction was stopped by adding Laemmli buffer and heating at 95°C for 10 min. Samples were run on a 12% SDS-PAGE gel then analyzed by Western blotting with 1/25000 streptavidin-horseradish peroxidase conjugate. Band intensities (n = 3) were measured using ImageJ by defining an equal ROI around the BccP biotinylated band at each time point. Time

course graphs were plotted and fitted using KaleidaGraph version 5.0 (Synergy Software), using the equation for the level of biotinylation B in function of time, t:

$$B(t) = \frac{B^I * t}{t^{50} + t}$$

where B^I is the level of biotinylation when $t \rightarrow \infty$, and t^{50} is the midpoint of biotinylation, i.e., the time at which 50% of the protein is biotinylated.

5.7 Electrophoretic mobility shift assay

The 112 bp biotin operator was obtained from *Shigella flexneri* strain M90T by PCR with primers HMIO702-703 described in reference (Chakravartty & Cronan, 2012). The PCR product was run on a 2% polyacrylamide gel and purified with a gel and PCR purification kit to obtain the biotin operator DNA probe. The gel shift assay was performed by incubating for 30 min at room temperature, 40 nM of this DNA probe with different concentrations (50 nM, 150 nM, 250 nM) of each biotin ligase tested, in a reaction buffer (50 mM tris-HCl pH 8.0, 1 mM EDTA, 50 mM NaCl, 1 mM ATP, 1 mM $MgCl_2$, 1 μM biotin, 10% glycerol). A 5% polyacrylamide retardation gel was prerun in 0.5× tris-acetate-EDTA buffer (TAE) at room temperature for 30 min (100V) before the reaction mixtures were loaded onto it and run at 4°C for 45 min (100V). The gels were washed with ddH_2O and stained with 1 μg/mL ethidium bromide for 20 min then visualized using a ChemiDoc XRS+ system. A Coomassie-stained SDS gel was prepared in parallel as a loading control for the samples. Band intensities (n = 3) were quantified in ImageJ by defining a ROI around the free DNA band at 112 bp in the positive control lane. The same ROI position was used to measure intensity in the corresponding region of the other lanes, representing where unshifted DNA would be expected. The intensity of the free DNA band in the positive control was considered 100%. For graphing purposes, the values were inverted so that the positive control was assigned 0%, and higher percentages correspond to an increasing fraction of DNA shifted from the ROI, representing DNA bound by the biotin ligase. Graphs were plotted using GraphPad Prism showing data as mean ± standard deviation. DNA shift in each sample was compared to the free DNA using one-way ANOVA followed by Dunnett's multiple comparisons test.

5.8 Complementation and biotinylation assay for the TurboID fusions with IpgC and MxiE

The strains Δ*ipgC* and Δ*ipaD* Δ*ipgC* harboring pUC18.1rp-derived plasmids pNHA10, 11, and 12 and strains Δ*mxiE* and *ipaB4* Δ*mxiE* harboring pUC18.1rp-derived plasmids pNHA10, 13, and 14 were used to test for the complementation. *Shigella flexneri* Str M90T WT, Δ*ipaD*, and *ipaB4* harboring pNHA9 were used as the reference for the off-state and on-state of the T3SS. These strains were incubated overnight at 30 degrees with shaking at 250 rpm in TSB supplemented with ampicillin. The next day, a 1/50 subculture was prepared in TSB and incubated at 37°C until an OD600 ≈ 1 was reached. Cells were quantified based on their OD, pelleted, resuspended in 1× Laemmli buffer, and heated for 10 min at 95°C. Samples were run on a 12% SDS-PAGE gel then analyzed by Western blot, as described in section 5.4, using the mouse anti-Myc primary antibodies, 1/10,000 rabbit polyclonal anti-IpgC (Ménard, Sansonetti, Parsot, et al., 1994), and 1/3,000 rabbit anti-IpaH (Mavris, Page, et al., 2002). The biotinylation assay was performed, as described in section 5.4, with minor modifications; bacterial cultures were supplemented with 50 μM biotin, except for the no supplemental biotin control, and incubated at 37°C for 10 and 20 min. Protein samples were prepared and analyzed by Western blot using streptavidin-horseradish peroxidase conjugate, mouse anti-Myc and anti-RecA primary antibodies, and 1/20,000 anti-mouse IgG-HRP and anti-rabbit IgG-HRP secondary antibodies (Jackson Immunoresearch, 111-035-003). Quantification of biotinylated proteins by each fusion at different time points was assessed in each strain for the IpgC-TurboID fusions, by measuring pixel intensities in the ROI from 28 to 150 kDa for each lane (n = 3) with ImageJ. To calculate the percentage of biotinylation events that occurred in the 10 min of incubation with the supplemental biotin, we subtracted relative pixel intensities in the no supplemental biotin control (%) from pixel intensities in the 10 min with supplemental biotin (%) for each strain discussed.

5.9 Preparation of proteins for mass spectrometry analyses

The strains described in the complementation experiment were grown overnight at 30°C in TSB supplemented with ampicillin. The next day, a 1/50 dilution was prepared in 30 mL TSB supplemented with 50 mg/mL of ampicillin and incubated at 37°C until an OD600 ≈ 1. Next, 50 μM biotin was added to each flask, which were then incubated at 37°C for 10 min. Samples were cooled on ice for 15 min and washed with buffer W (50 mM HEPES [Fisher Scientific, BP310; pH 7.3], 200 mM NaCl, 10% glycerol, 1 mM PMSF [Sigma-Aldrich P7626], and 2 mM EDTA),

then resuspended in 4 mL of lysis buffer (buffer W supplemented with 1% [wt/vol] n-dodecyl-β-D-maltoside [Cayman Chemical Company, 16494]). Cells were lysed by sonication with 30 s pulses (0.85 cycle, 80% amplitude) repeated four times with a 30 s rest on ice between pulses. Purification of biotinylated proteins and collection of tryptic peptides were adapted from reference (Filip et al., 2021). Samples were centrifuged at 17,000 × *g* at 4°C for 20 min, and the resulting supernatant was mixed with 50 µL of PBS-washed streptavidin-agarose beads (Thermo Scientific, 20353) by end-over-end rotation at room temperature for 1.5 hours. The beads were then collected by centrifugation at 2,000 × *g* for 2 min and washed three times with 1% SDS using Bio-Spin Columns (Bio-Rad, 7326204), followed by three washes with 6 M urea (Bio-Rad, 161-0731). Protein-loaded streptavidin beads were washed one time with PBS and five times with 50 mM ammonium bicarbonate (ABC) (Thermo Scientific, 393212500) in Bio-Spin Columns then transferred to microcentrifuge tubes and pelleted at 2,000 × *g* for 2 min. The beads were resuspended in 10 mM DTT (Thermo Scientific, R0861) in ABC and heated for 30 min at 56°C followed by the addition of iodoacetamide (Thermo Scientific, 122270050) at a final concentration of 25 mM and an incubation in the dark with rotation at room temperature for 30 min. The beads were washed in ABC supplemented with 10 mM DTT for 5 min followed by three washes with 50 mM triethylammonium bicarbonate or TEAB (Thermo Scientific, 90114), pH 8.5. The proteins were then digested off the bead with 10 µg/mL trypsin (Promega, V5113) in TEAB while shaking overnight at 37°C. Digested peptides were recovered by centrifugation and filtration using Bio-Spin Columns. The samples were dried using vacuum centrifugation at 1,500 rpm and room temperature.

5.10 Mass spectrometry-based proteomics analyses

The samples were analyzed at the John L. Holmes Mass Spectrometry Facility (Faculty of Science, University of Ottawa), as previously described (Risha et al., 2020). Briefly, the tryptic peptides obtained in the previous section were analyzed by an Orbitrap Fusion mass spectrometer (Thermo Scientific) coupled to a Dionex UltiMate 3000 RSLCnano system (Thermo Scientific) with an in-house-packed 70 µm ID × 150 mm length separation column (Polymicro Technology), Luna C18(2), 3 µm, 100 Å (Phenomenex). Samples were injected and separated through a gradient solution of 0.1% formic acid and acetonitrile in deionized water.

They were applied to the column for a total of 105 min at a flowrate of 0.30 µL/min, which divided as follows. The peptides were separated with 2% acetonitrile for the first 7 min, then the gradient was linearly increased to 38% for the next 70 min, then from 38 to 98% for 5 min, remained at 98% for 10 min, then gradually decreased from 98 to 2% acetonitrile for 3 min, and finally washed with 2% acetonitrile for 10 min. Peptides were ionized using nano-electrospray ionization (ESI) at an ion source temperature of 250°C with an ion spray voltage of 2.1 kV. Survey scans between 300 and 2,000 *m*/*z* were acquired at 60K resolution, and precursors with charge state +2 to +7 were filtered according to the monoisotopic precursor selection. The dynamic exclusion time was set to 30 s with a tolerance of 10 ppm. Automatic gain control settings were set to 5×10^5 for full Fourier transform mass spectrometry (FTMS) scans and 1×10^4 for tandem mass spectrometry (MS/MS) scans. Precursors were isolated using a 2 *m/z* isolation frame, and collision-induced dissociation fragmentation was performed with a collision energy of 35%.

5.11 Mass spectrometry computational data analysis

The raw mass spectrometry data were processed using the Trans-Proteomic Pipeline (v6.3.0 Arcus, Build 202305021110-8944 for WindowsNT-x8664) (Deutsch et al., 2023), as follows. The raw files were converted to mzML using ProteoWizard's msconvert (version 3.0.22340) (Chambers et al., 2012). Comet (release 2023.01 rev.2) (Eng et al., 2013) was used for peptide-spectrum matching against a protein sequence database made of the proteome of *S. flexneri* 5a strain M90T (NZ_CP037923.1) and its virulence plasmid pWR100 (NC_024996.1) from the NCBI RefSeq database along with the "common contaminants" from the Max Planck Institute and their reversed counterpart for a target-decoy database search (10,269 sequences total). Semi-tryptic peptides with at most three miscleavages were searched by Comet with a precursor mass tolerance of 20 ppm. Carbamidomethylation (C) was considered as a static modification and oxidation (M) and deamidation (NQ) were considered as variable modifications. Confidence of Comet peptide-spectrum match results was assessed using PeptideProphet (Keller et al., 2002), ProteinProphet, and iProphet (Nesvizhskii et al., 2003; Shteynberg et al., 2011). Label-free quantification was achieved using StPeter (version 1.4.0) (Hoopmann et al., 2018) with 1% FDR cutoffs and a probability threshold associated with a

calculated 1% FDR for the input data, protein samples were set at 20 µg, and degenerate peptide quantification was allowed. The protein lists were exported to include only the sequences under the calculated 1% FDR threshold in each replicate with the DECOY and contaminant sequences filtered out.

5.12 Computational assessment of proxisome

The SAINT software package (version 2.5.0) (Choi et al., 2011) was used to calculate the probability of the detected proteins being enriched in the test versus control samples using spectral counts. The test samples consisted of five (IpgC) and six (MxiE) biological replicates for each TurboID fusion (N- or C-terminal). The control samples consisted of two or three biological replicates of the respective knockout strains used to mimic the on-state and off-state of the T3SS (Δ*ipgC*, Δ*ipaD* Δ*ipgC*, Δ*mxiE*, and *ipaB4* Δ*mxiE*) expressing the free TurboID, and of non-TurboID expressing strains (M90T WT and Δ*ipaD*). For each bait, two SAINT analyses were run corresponding to the off-and on-state, each including both N- and C-terminal TurboID fusions as independent test groups. SAINT was run with recommended settings. Ten percent FDR cutoffs including, at a minimum, all interactions with a probability > 90% was used to identify confident vicinal proteins. The FC of a protein p was calculated using LFQ intensities of p with this formula:

FC = [LFQ(p) test on/LFQ(p) control on]/[LFQ(p) test off/LFQ(p) control off]

in which “on” and “off” indicate the LFQ in the on-state and off-state, respectively, tests were either the N-terminal or C-terminal fusions of TurboID with IpgC or MxiE and the controls were the corresponding free TurboID, as described above. Proteins that were undetected in one or more of the samples (LFQ = 0) were imputed with LFQ = 1 to perform the FC calculation. To plot these data, FC values for proteins FC < 1 (i.e., those enriched in the off-state) were inverted to obtain FC > 1, and proteins enriched in the off-state and on-state were identified with red and blue dots, respectively.

5.13 STRING network

A STRING network of selected proteins found in the proxisome of IpgC according to TurboID was generated using STRING v.12.0 taking into consideration known interactions that were experimentally determined or detected in curated databases, as well as predicted interactions based on gene neighboring or co-occurrence, and co-expression. An FDR stringency of 1% and a medium confidence of 0.400 were applied using the *E. coli* K-12 database. All proteins showed in Figure 3.7.C and conserved in *E. coli* K-12 were used to generate the initial network. The most highly interconnected region of this network is shown in Figure 3.8 The GO term analyses were performed on this subnetwork with the whole *E. coli* K-12 genome with threshold FDR ≤ 5% and strength > 0.01.

5.14 Acid resistance

Acid resistance assays were performed based on the protocol reported by (Ojha et al., 2021) with adjustments of experimental conditions. WT, Δ*ipaD*, Δ*ipgC*, and Δ*ipaD* Δ*ipgC S. flexneri* strains were subcultured from overnight LB cultures at a 1:100 dilution and grown with shaking at 37°C in 3 mL LB broth until reaching an OD600 ≈ 1. Cultures containing 5×10^8 cells were pelleted by centrifugation and resuspended in LB adjusted to different pH, with or without 50 mM glutamic acid (Fisher Chemical, A125-100) where indicated. After 2 hours of acid exposure at 37°C, serial dilutions (10^0 to 10^{-7}) were prepared in PBS (pH 7.3), and 5 µL of each dilution was spotted onto LB agar plates supplemented with 0.01% Congo red. Survival and growth were assessed after overnight incubation and expressed as a percentage relative to the initial population at T_0, with values above 100% indicating growth and values below 100% indicating a decline in viability. Spot intensities (n = 3) were measured using ImageJ by defining an equal ROI around each spot and measuring its density, providing a relative measure of growth and survival under each condition. Graphs were plotted using GraphPad Prism showing data as mean ± standard deviation. growth and survival were compared to T_0 using one-way ANOVA.

References

Abram, K., Udaondo, Z., Bleker, C., Wanchai, V., Wassenaar, T. M., Robeson, M. S., & Ussery, D. W. (2021). Mash-based analyses of *Escherichia coli* genomes reveal 14 distinct phylogroups. *Communications Biology*, *4*(1), 117. https://doi.org/10.1038/s42003-020-01626-5

Adikaram, P. R., & Beckett, D. (2012). Functional versatility of a single protein surface in two protein:protein interactions. *Journal of Molecular Biology*, *419*(3–4), 223–233. https://doi.org/10.1016/j.jmb.2012.03.010

Agaisse, H. (2016). Molecular and Cellular Mechanisms of *Shigella flexneri* Dissemination. *Frontiers in Cellular and Infection Microbiology*, *6*. https://doi.org/10.3389/fcimb.2016.00029

Akeda, Y., & Galán, J. E. (2005). Chaperone release and unfolding of substrates in type III secretion. *Nature*, *437*(7060), 911–915. https://doi.org/10.1038/nature03992

Akhondi, H., Goldin, J., & Simonsen, K. A. (2025). Bacterial Diarrhea. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK551643/

Alberta Health Services. (2025, March 28). *Shigella Information—Edmonton Zone Medical Officers of Health*. Alberta Health Services. https://www.albertahealthservices.ca/moh/Page19158.aspx

Allaoui, A., Mounier, J., Prévost, M. C., Sansonetti, P. J., & Parsot, C. (1992). *icsB*: A *Shigella flexneri* virulence gene necessary for the lysis of protrusions during intercellular spread. *Molecular Microbiology*, *6*(12), 1605–1616. https://doi.org/10.1111/j.1365-2958.1992.tb00885.x

Anderson, M., Sansonetti, P. J., & Marteyn, B. S. (2016). *Shigella* Diversity and Changing Landscape: Insights for the Twenty-First Century. *Frontiers in Cellular and Infection Microbiology*, *6*, 45. https://doi.org/10.3389/fcimb.2016.00045

Aslam, A., Hashmi, M. F., & Okafor, C. N. (2025). Shigellosis. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK482337/

Azevedo, A. M., Martins, V. C., Prazeres, D. M. F., Vojinović, V., Cabral, J. M. S., & Fonseca, L. P. (2003). Horseradish peroxidase: A valuable tool in biotechnology. In *Biotechnology Annual Review* (pp. 199–247). Elsevier. https://doi.org/10.1016/s1387-2656(03)09003-3

Bagautdinov, B., Matsuura, Y., Bagautdinova, S., & Kunishima, N. (2008). Protein Biotinylation Visualized by a Complex Structure of Biotin Protein Ligase with a Substrate. *Journal of Biological Chemistry*, *283*(21), 14739–14750. https://doi.org/10.1074/jbc.M709116200

Bajunaid, W., Haidar-Ahmad, N., Kottarampatel, A. H., Ourida Manigat, F., Silué, N., F. Tchagang, C., Tomaro, K., & Campbell-Valois, F.-X. (2020). The T3SS of *Shigella*: Expression, Structure, Function, and Role in Vacuole Escape. *Microorganisms*, *8*(12), 1933. https://doi.org/10.3390/microorganisms8121933

Baker, S., & Scott, T. A. (2023). Antimicrobial-resistant *Shigella*: Where do we go next? *Nature Reviews Microbiology*, *21*(7), 409–410. https://doi.org/10.1038/s41579-023-00906-1

Barbagallo, M., Di Martino, M. L., Marcocci, L., Pietrangeli, P., De Carolis, E., Casalino, M., Colonna, B., & Prosseda, G. (2011). A New Piece of the *Shigella* Pathogenicity Puzzle: Spermidine Accumulationby Silencing of the *speG* Gene. *PLoS ONE*, *6*(11), e27226. https://doi.org/10.1371/journal.pone.0027226

Barceloux, D. G. (2008). *Shigella* Species (Shiga Enterotoxins). *Medical toxicology of natural substances: Foods, fungi, medicinal herbs, plants, and venomous animals*. (pp. 150-155). John Wiley & Sons. https://doi.org/10.1002/9780470330319.ch20

Barta, M. L., Tachiyama, S., Muthuramalingam, M., Arizmendi, O., Villanueva, C. E., Ramyar, K. X., Geisbrecht, B. V., Lovell, S., Battaile, K. P., Picking, W. L., & Picking, W. D. (2018). Using disruptive insertional mutagenesis to identify the *in situ* structure-function landscape of the *Shigella* translocator protein IpaB. *Protein Science : A Publication of the Protein Society*, *27*(8), 1392–1406. https://doi.org/10.1002/pro.3428

Bartoli, J., Viala, J. P., & Bouveret, E. (2020). SlyA Transcriptional Regulator Is Not Directly Affected by ppGpp Levels. *Frontiers in Microbiology*, *11*, 1856. https://doi.org/10.3389/fmicb.2020.01856

Beloin, C., & Dorman, C. J. (2003). An extended role for the nucleoid structuring protein H-NS in the virulence gene regulatory cascade of *Shigella flexneri*. *Molecular Microbiology*, *47*(3), 825–838. https://doi.org/10.1046/j.1365-2958.2003.03347.x

Bennish, M. L., & Ahmed, S. (2020). Shigellosis. In *Hunter's Tropical Medicine and Emerging Infectious Diseases* (pp. 492–499). Elsevier. https://doi.org/10.1016/b978-0-323-55512-8.00048-x

Bernardini, M. L., Mounier, J., d'Hauteville, H., Coquis-Rondon, M., & Sansonetti, P. J. (1989). Identification of *icsA*, a plasmid locus of *Shigella flexneri* that governs bacterial intra- and intercellular spread through interaction with F-actin. *Proceedings of the National Academy of Sciences*, *86*(10), 3867–3871. https://doi.org/10.1073/pnas.86.10.3867

Blocker, A., Jouihri, N., Larquet, E., Gounon, P., Ebel, F., Parsot, C., Sansonetti, P., & Allaoui, A. (2001). Structure and composition of the *Shigella flexneri* 'needle complex', a part of its type III secreton. *Molecular Microbiology*, *39*(3), 652–663. https://doi.org/10.1046/j.1365-2958.2001.02200.x

Bongrand, C., Sansonetti, P. J., & Parsot, C. (2012). Characterization of the Promoter, MxiE Box and 5′ UTR of Genes Controlled by the Activity of the Type III Secretion Apparatus in *Shigella flexneri*. *PLoS ONE*, *7*(3), e32862. https://doi.org/10.1371/journal.pone.0032862

Bosch, J. A., Chen, C., & Perrimon, N. (2021). Proximity-dependent labeling methods for proteomic profiling in living cells: An update. *WIREs Developmental Biology*, *10*(1). https://doi.org/10.1002/wdev.392

Botteaux, A., Sory, M. P., Biskri, L., Parsot, C., & Allaoui, A. (2009). MxiC is secreted by and controls the substrate specificity of the *Shigella flexneri* type III secretion apparatus. *Molecular Microbiology*, *71*(2), 449–460. https://doi.org/10.1111/j.1365-2958.2008.06537.x

Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., Feldman, J. L., Perrimon, N., & Ting, A. Y. (2018). Efficient proximity labeling in living cells and organisms with TurboID. *Nature Biotechnology*, *36*(9), 880–887. https://doi.org/10.1038/nbt.4201

Braun, M. E., Follador, R., & Kemmler, S. J. (2021). *Shigella-tetravalent (shigella4v) bioconjugate*. https://patentscope.wipo.int/search/en/WO2021255684

Bravo, V., Puhar, A., Sansonetti, P., Parsot, C., & Toro, C. S. (2015). Distinct Mutations Led to Inactivation of Type 1 Fimbriae Expression in *Shigella* spp. *PLOS ONE*, *10*(3), e0121785. https://doi.org/10.1371/journal.pone.0121785

Broach, W. H., Egan, N., Wing, H. J., Payne, S. M., & Murphy, E. R. (2012). VirF-Independent Regulation of *Shigella virB* Transcription is Mediated by the Small RNA RyhB. *PLoS ONE*, *7*(6), e38592. https://doi.org/10.1371/journal.pone.0038592

Buchrieser, C., Glaser, P., Rusniok, C., Nedjari, H., D'Hauteville, H., Kunst, F., Sansonetti, P., & Parsot, C. (2000). The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of *Shigella flexneri*. *Molecular Microbiology*, *38*(4), 760–771. https://doi.org/10.1046/j.1365-2958.2000.02179.x

Campbell-Valois, F.-X., Schnupf, P., Nigro, G., Sachse, M., Sansonetti, P. J., & Parsot, C. (2014). A Fluorescent Reporter Reveals On/Off Regulation of the *Shigella* Type III Secretion Apparatus during Entry and Cell-to-Cell Spread. *Cell Host & Microbe*, *15*(2), 177–189. https://doi.org/10.1016/j.chom.2014.01.005

Cashel, M., & Gallant, J. (1969). Two compounds implicated in the function of the RC gene of *Escherichia coli*. *Nature*, *221*(5183), 838–841. https://doi.org/10.1038/221838a0

Chakravartty, V., & Cronan, J. E. (2012). Altered Regulation of *Escherichia coli* Biotin Biosynthesis in BirA Superrepressor Mutant Strains. *Journal of Bacteriology*, *194*(5), 1113–1126. https://doi.org/10.1128/JB.06549-11

Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M.-Y., Paulse, C., Creasy, D., … Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, *30*(10), 918–920. https://doi.org/10.1038/nbt.2377

Chapman-Smith, A., & Cronan, J. E. (1999). The enzymatic biotinylation of proteins: A post-translational modification of exceptional specificity. *Trends in Biochemical Sciences*, *24*(9), 359–363. https://doi.org/10.1016/s0968-0004(99)01438-3

Chen, H., Gong, X., Fan, Z., Xia, Y., Jin, Y., Bai, F., Cheng, Z., Pan, X., & Wu, W. (2023). *Pseudomonas aeruginosa* Citrate Synthase GltA Influences Antibiotic Tolerance and the Type III Secretion System through the Stringent Response. *Microbiology Spectrum*, *11*(1), e03239-22. https://doi.org/10.1128/spectrum.03239-22

Chen, M., Zhang, W., Han, L., Ru, X., Cao, Y., Hikichi, Y., Ohnishi, K., Pan, G., & Zhang, Y. (2022). A CysB regulator positively regulates cysteine synthesis, expression of type III secretion system genes, and pathogenicity in *Ralstonia solanacearum*. *Molecular Plant Pathology*, *23*(5), 679–692. https://doi.org/10.1111/mpp.13189

Cherradi, Y., Hachani, A., & Allaoui, A. (2014). Spa13 of *Shigella flexneri* has a dual role: Chaperone escort and export gate-activator switch of the type III secretion system. *Microbiology*, *160*(1), 130–141. https://doi.org/10.1099/mic.0.071712-0

Cheung, M., Shen, D., Makino, F., Kato, T., Roehrich, A. D., Martinez-Argudo, I., Walker, M. L., Murillo, I., Liu, X., Pain, M., Brown, J., Frazer, G., Mantell, J., Mina, P., Todd, T., Sessions, R. B., Namba, K., & Blocker, A. J. (2015). Three-dimensional electron microscopy reconstruction and cysteine-mediated crosslinking provide a model of the type III secretion system needle tip complex. *Molecular Microbiology*, *95*(1), 31–50. https://doi.org/10.1111/mmi.12843

Chiang, I.-L., Wang, Y., Fujii, S., Muegge, B. D., Lu, Q., Tarr, P. I., & Stappenbeck, T. S. (2021). Biofilm Formation and Virulence of *Shigella flexneri* Are Modulated by pH of Gastrointestinal Tract. *Infection and Immunity*, *89*(11), e0038721. https://doi.org/10.1128/IAI.00387-21

Cho, K. F., Branon, T. C., Rajeev, S., Svinkina, T., Udeshi, N. D., Thoudam, T., Kwak, C., Rhee, H.-W., Lee, I.-K., Carr, S. A., & Ting, A. Y. (2020). Split-TurboID enables contact-dependent proximity labeling in cells. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(22), 12143–12154. https://doi.org/10.1073/pnas.1919528117

Cho, K. F., Branon, T. C., Udeshi, N. D., Myers, S. A., Carr, S. A., & Ting, A. Y. (2020). Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nature Protocols*, *15*(12), 3971–3999. https://doi.org/10.1038/s41596-020-0399-0

Choi, H., Larsen, B., Lin, Z.-Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z. S., Tyers, M., Gingras, A.-C., & Nesvizhskii, A. I. (2011). SAINT: Probabilistic scoring of affinity purification–mass spectrometry data. *Nature Methods*, *8*(1), 70–73. https://doi.org/10.1038/nmeth.1541

Choi-Rhee, E., Schulman, H., & Cronan, J. E. (2004). Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein Science: A Publication of the Protein Society*, *13*(11), 3043–3050. https://doi.org/10.1110/ps.04911804

Coluccia, M., Béranger, A., Trirocco, R., Fanelli, G., Zanzi, F., Colonna, B., Grossi, M., Prosseda, G., & Pasqua, M. (2023). Role of the MDR Efflux Pump AcrAB in Epithelial Cell Invasion by *Shigella flexneri*. *Biomolecules*, *13*(5), 823. https://doi.org/10.3390/biom13050823

Cossart, P., & Sansonetti, P. J. (2004). Bacterial Invasion: The Paradigms of Enteroinvasive Pathogens. *Science*, *304*(5668), 242–248. https://doi.org/10.1126/science.1090124

Cronan, J. E. (2024). Biotin protein ligase as you like it: Either extraordinarily specific or promiscuous protein biotinylation. *Proteins: Structure, Function, and Bioinformatics*, *92*(4), 435–448. https://doi.org/10.1002/prot.26642

Darwin, A. J. (2013). Stress Relief during Host Infection: The Phage Shock Protein Response Supports Bacterial Virulence in Various Ways. *PLoS Pathogens*, *9*(7), e1003388. https://doi.org/10.1371/journal.ppat.1003388

Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, *97*(12), 6640–6645. https://doi.org/10.1073/pnas.120163297

Daugherty, A., Suvarnapunya, A. E., & Runyen-Janecky, L. (2012). The role of OxyR and SoxRS in oxidative stress survival in *Shigella flexneri*. *Microbiological Research*, *167*(4), 238–245. https://doi.org/10.1016/j.micres.2011.09.004

De Biase, D., & Pennacchietti, E. (2012). Glutamate decarboxylase-dependent acid resistance in orally acquired bacteria: Function, distribution and biomedical implications of the gadBC operon. *Molecular Microbiology*, *86*(4), 770–786. https://doi.org/10.1111/mmi.12020

De Munter, S., Görnemann, J., Derua, R., Lesage, B., Qian, J., Heroes, E., Waelkens, E., Van Eynde, A., Beullens, M., & Bollen, M. (2017). Split-BioID: A proximity biotinylation assay for dimerization-dependent protein interactions. *FEBS Letters*, *591*(2), 415–424. https://doi.org/10.1002/1873-3468.12548

Deng, W., Marshall, N. C., Rowland, J. L., McCoy, J. M., Worrall, L. J., Santos, A. S., Strynadka, N. C. J., & Finlay, B. B. (2017). Assembly, structure, function and regulation of type III secretion systems. *Nature Reviews. Microbiology*, *15*(6), 323–337. https://doi.org/10.1038/nrmicro.2017.20

Deutsch, E. W., Mendoza, L., Shteynberg, D. D., Hoopmann, M. R., Sun, Z., Eng, J. K., & Moritz, R. L. (2023). Trans-Proteomic Pipeline: Robust Mass Spectrometry-Based Proteomics Data Analysis Suite. *Journal of Proteome Research*, *22*(2), 615–624. https://doi.org/10.1021/acs.jproteome.2c00624

Dickenson, N. E., Arizmendi, O., Patil, M. K., Toth, R. T., Middaugh, C. R., Picking, W. D., & Picking, W. L. (2013). N-Terminus of IpaB Provides a Potential Anchor to the *Shigella* Type III Secretion System Tip Complex Protein IpaD. *Biochemistry*, *52*(49), 8790–8799. https://doi.org/10.1021/bi400755f

Dif, G., Djemouai, N., Bouras, N., & Zitouni, A. (2025). In-Depth Genome-Based Analysis of *Shigella* spp. and *Escherichia* spp.: Resolving Ambiguities and Unveiling Phylogenetic Relationships. *Current Microbiology*, *82*(4). https://doi.org/10.1007/s00284-025-04158-5

Duncan-Lowey, J. K., Wiscovitch, A. L., Wood, T. E., Goldberg, M. B., & Russo, B. C. (2020). *Shigella flexneri* Disruption of Cellular Tension Promotes Intercellular Spread. *Cell Reports*, *33*(8), 108409. https://doi.org/10.1016/j.celrep.2020.108409

DuPont, H. L., Levine, M. M., Hornick, R. B., & Formal, S. B. (1989). Inoculum size in shigellosis and implications for expected mode of transmission. *The Journal of Infectious Diseases*, *159*(6), 1126–1128. https://doi.org/10.1093/infdis/159.6.1126

Egile, C., Loisel, T. P., Laurent, V., Li, R., Pantaloni, D., Sansonetti, P. J., & Carlier, M.-F. (1999). Activation of the Cdc42 Effector N-Wasp by the *Shigella flexneri* Icsa Protein Promotes Actin Nucleation by Arp2/3 Complex and Bacterial Actin-Based Motility. *The Journal of Cell Biology*, *146*(6), 1319–1332. https://doi.org/10.1083/jcb.146.6.1319

Elowitz, M. B., Surette, M. G., Wolf, P.-E., Stock, J. B., & Leibler, S. (1999). Protein Mobility in the Cytoplasm of *Escherichia coli*. *Journal of Bacteriology*, *181*(1), 197–203. https://doi.org/10.1128/JB.181.1.197-203.1999

EMBL-EBI. (2025). *Confidence scores in AlphaFold-Multimer | AlphaFold*. https://www.ebi.ac.uk/training/online/courses/alphafold/inputs-and-outputs/evaluating-alphafolds-predicted-structures-using-confidence-scores/confidence-scores-in-alphafold-multimer/

Eng, J. K., Jahan, T. A., & Hoopmann, M. R. (2013). Comet: An open-source MS/MS sequence database search tool. *Proteomics*, *13*(1), 22–24. https://doi.org/10.1002/pmic.201200439

Ernst, N. H., Reeves, A. Z., Ramseyer, J. E., & Lesser, C. F. (2018). High-Throughput Screening of Type III Secretion Determinants Reveals a Major Chaperone-Independent Pathway. *mBio*, *9*(3). https://doi.org/10.1128/mbio.01050-18

Escobar-Paramo, P., Giudicelli, C., Parsot, C., & Denamur, E. (2003). The Evolutionary History of *Shigella* and Enteroinvasive *Escherichia coli* Revised. *Journal of Molecular Evolution*, *57*(2), 140–148. https://doi.org/10.1007/s00239-003-2460-3

Falconi, M., Colonna, B., Prosseda, G., Micheli, G., & Gualerzi, C. O. (1998). Thermoregulation of *Shigella* and *Escherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of *virF* promoter to transcriptional repressor H-NS. *The EMBO Journal*, *17*(23), 7033–7043. https://doi.org/10.1093/emboj/17.23.7033

Feng, Y., Chen, Z., & Liu, S.-L. (2011). Gene Decay in *Shigella* as an Incipient Stage of Host-Adaptation. *PLoS ONE*, *6*(11), e27754. https://doi.org/10.1371/journal.pone.0027754

Ferrari, M. L., Charova, S. N., Sansonetti, P. J., Mylonas, E., & Gazi, A. D. (2021). Structural Insights of *Shigella* Translocator IpaB and Its Chaperone IpgC in Solution. *Frontiers in Cellular and Infection Microbiology*, *11*. https://doi.org/10.3389/fcimb.2021.673122

Filip, R., Desrochers, G. F., Lefebvre, D. M., Reed, A., Singaravelu, R., Cravatt, B. F., & Pezacki, J. P. (2021). Profiling of MicroRNA Targets Using Activity-Based Protein Profiling: Linking Enzyme Activity to MicroRNA-185 Function. *Cell Chemical Biology*, *28*(2), 202-212.e6. https://doi.org/10.1016/j.chembiol.2020.12.009

Finka, A., & Goloubinoff, P. (2013). Proteomic data from human cell cultures refine mechanisms of chaperone-mediated protein homeostasis. *Cell Stress and Chaperones*, *18*(5), 591–605. https://doi.org/10.1007/s12192-013-0413-3

Flexner, S. (1990). On the etiology of tropical dysentery. In *Bull Johns Hopkins Hosp* (Vol. 11, pp. 231–242).

Fontenot, E. M., Ezelle, K. E., Gabreski, L. N., Giglio, E. R., McAfee, J. M., Mills, A. C., Qureshi, M. N., Salmon, K. M., & Toyota, C. G. (2013). YfdW and YfdU Are Required for Oxalate-Induced Acid Tolerance in *Escherichia coli* K-12. *Journal of Bacteriology*, *195*(7), 1446–1455. https://doi.org/10.1128/JB.01936-12

Foster, J. W. (2004). *Escherichia coli* acid resistance: Tales of an amateur acidophile. *Nature Reviews. Microbiology*, *2*(11), 898–907. https://doi.org/10.1038/nrmicro1021

Gajiwala, K. S., & Burley, S. K. (2000). HDEA, a periplasmic protein that supports acid resistance in pathogenic enteric bacteria. *Journal of Molecular Biology*, *295*(3), 605–612. https://doi.org/10.1006/jmbi.1999.3347

Gauthier, A., & Finlay, B. B. (2003). Translocated Intimin Receptor and Its Chaperone Interact with ATPase of the Type III Secretion Apparatus of Enteropathogenic *Escherichia coli*. *Journal of Bacteriology*, *185*(23), 6747–6755. https://doi.org/10.1128/JB.185.23.6747-6755.2003

Gazi, A. D., Charova, S. N., Panopoulos, N. J., & Kokkinidis, M. (2009). Coiled-coils in type III secretion systems: Structural flexibility, disorder and biological implications. *Cellular Microbiology*, *11*(5), 719–729. https://doi.org/10.1111/j.1462-5822.2009.01297.x

Gingras, A.-C., Abe, K. T., & Raught, B. (2019). Getting to know the neighborhood: Using proximity-dependent biotinylation to characterize protein complexes and map organelles. *Current Opinion in Chemical Biology*, *48*, 44–54. https://doi.org/10.1016/j.cbpa.2018.10.017

Government of Canada. (2024, January 29). *For health professionals: Shigellosis (Shigella)*. https://www.canada.ca/en/public-health/services/diseases/shigella/health-professionals.html

Grunden, A. M., Ray, R. M., Rosentel, J. K., Healy, F. G., & Shanmugam, K. T. (1996). Repression of the *Escherichia coli* modABCD (molybdate transport) operon by ModE. *Journal of Bacteriology*, *178*(3), 735–744. https://doi.org/10.1128/jb.178.3.735-744.1996

Grunden, A. M., Self, W. T., Villain, M., Blalock, J. E., & Shanmugam, K. T. (1999). An analysis of the binding of repressor protein ModE to modABCD (molybdate transport) operator/promoter DNA of *Escherichia coli*. *The Journal of Biological Chemistry*, *274*(34), 24308–24315. https://doi.org/10.1074/jbc.274.34.24308

Guo, J., Guo, S., Lu, S., Gong, J., Wang, L., Ding, L., Chen, Q., & Liu, W. (2023). The development of proximity labeling technology and its applications in mammals, plants, and microorganisms. *Cell Communication and Signaling*, *21*(1). https://doi.org/10.1186/s12964-023-01310-1

Habenstein, B., El Mammeri, N., Tolchard, J., Lamon, G., Tawani, A., Berbon, M., & Loquet, A. (2019). Structures of Type III Secretion System Needle Filaments. In *Current Topics in Microbiology and Immunology* (pp. 109–131). Springer International Publishing. https://doi.org/10.1007/82_2019_192

Haidar-Ahmad, N., Manigat, F. O., Silué, N., Pontier, S. M., & Campbell-Valois, F.-X. (2023). A Tale about *Shigella*: Evolution, Plasmid, and Virulence. *Microorganisms*, *11*(7), 1709. https://doi.org/10.3390/microorganisms11071709

Haidar-Ahmad, N., Tomaro, K., Lavallée-Adam, M., & Campbell-Valois, F.-X. (2024). The promiscuous biotin ligase TurboID reveals the proxisome of the T3SS chaperone IpgC in *Shigella flexneri*. *mSphere*, *9*(11), e00553-24. https://doi.org/10.1128/msphere.00553-24

Hall, C. P., Jadeja, N. B., Sebeck, N., & Agaisse, H. (2022). Characterization of MxiE- and H-NS-Dependent Expression of *ipaH7.8*, *ospC1*, *yccE*, and *yfdF* in *Shigella flexneri*. *mSphere*, *7*(6), e0048522. https://doi.org/10.1128/msphere.00485-22

Hawkey, J., Monk, J. M., Billman-Jacobe, H., Palsson, B., & Holt, K. E. (2020). Impact of insertion sequences on convergent evolution of *Shigella* species. *PLoS Genetics*, *16*(7), e1008931. https://doi.org/10.1371/journal.pgen.1008931

Hensley, C. T., Kamneva, O. K., Levy, K. M., Labahn, S. K., Africa, L. A., & Wing, H. J. (2011). Two promoters and two translation start sites control the expression of the *Shigella*

flexneri outer membrane protease IcsP. *Archives of Microbiology*, *193*(4), 263–274. https://doi.org/10.1007/s00203-010-0669-2

Herfurth, M., Müller, F., Søgaard-Andersen, L., & Glatter, T. (2023). A miniTurbo-based proximity labeling protocol to identify conditional protein interactomes *in vivo* in *Myxococcus xanthus*. *STAR Protocols*, *4*(4), 102657. https://doi.org/10.1016/j.xpro.2023.102657

High, N., Mounier, J., Prévost, M. C., & Sansonetti, P. J. (1992). IpaB of *Shigella flexneri* causes entry into epithelial cells and escape from the phagocytic vacuole. *The EMBO Journal*, *11*(5), 1991–1999. https://doi.org/10.1002/j.1460-2075.1992.tb05253.x

Hoopmann, M. R., Winget, J. M., Mendoza, L., & Moritz, R. L. (2018). StPeter: Seamless Label-Free Quantification with the Trans-Proteomic Pipeline. *Journal of Proteome Research*, *17*(3), 1314–1320. https://doi.org/10.1021/acs.jproteome.7b00786

Horstman, N. K., & Darwin, A. J. (2012). Phage shock proteins B and C prevent lethal cytoplasmic membrane permeability in *Yersinia enterocolitica*. *Molecular Microbiology*, *85*(3), 445–460. https://doi.org/10.1111/j.1365-2958.2012.08120.x

Hu, B., Lara-Tejero, M., Kong, Q., Galán, J. E., & Liu, J. (2017). In Situ Molecular Architecture of the *Salmonella* Type III Secretion Machine. *Cell*, *168*(6), 1065-1074.e10. https://doi.org/10.1016/j.cell.2017.02.022

Hu, B., Morado, D. R., Margolin, W., Rohde, J. R., Arizmendi, O., Picking, W. L., Picking, W. D., & Liu, J. (2015). Visualization of the type III secretion sorting platform of *Shigella flexneri*. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(4), 1047–1052. https://doi.org/10.1073/pnas.1411610112

Hu, J., Worrall, L. J., Hong, C., Vuckovic, M., Atkinson, C. E., Caveney, N., Yu, Z., & Strynadka, N. C. J. (2018). Cryo-EM analysis of the T3S injectisome reveals the structure of the needle and open secretin. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-06298-8

Hu, J., Worrall, L. J., & Strynadka, N. C. (2020). Towards capture of dynamic assembly and action of the T3SS at near atomic resolution. *Current Opinion in Structural Biology*, *61*, 71–78. https://doi.org/10.1016/j.sbi.2019.10.005

Izoré, T., Job, V., & Dessen, A. (2011). Biogenesis, Regulation, and Targeting of the Type III Secretion System. *Structure*, *19*(5), 603–612. https://doi.org/10.1016/j.str.2011.03.015

Jackson, M. W., Silva-Herzog, E., & Plano, G. V. (2004). The ATP-dependent ClpXP and Lon proteases regulate expression of the *Yersinia pestis* type III secretion system via regulated proteolysis of YmoA, a small histone-like protein. *Molecular Microbiology*, *54*(5), 1364–1378. https://doi.org/10.1111/j.1365-2958.2004.04353.x

Jensen, V. B., Harty, J. T., & Jones, B. D. (1998). Interactions of the invasive pathogens *Salmonella typhimurium*, *Listeria monocytogenes*, and *Shigella flexneri* with M cells and murine Peyer's patches. *Infection and Immunity*, *66*(8), 3758–3766. https://doi.org/10.1128/IAI.66.8.3758-3766.1998

Johannes, L., & Römer, W. (2010). Shiga toxins from cell biology to biomedical applications. *Nature Reviews Microbiology*, *8*(2), 105–116. https://doi.org/10.1038/nrmicro2279

Johnson, B. S., Chafin, L., Farkas, D., Adair, J., Elhance, A., Farkas, L., Bednash, J. S., & Londino, J. D. (2022). MicroID2: A Novel Biotin Ligase Enables Rapid Proximity-Dependent Proteomics. *Molecular & Cellular Proteomics*, *21*(7), 100256. https://doi.org/10.1016/j.mcpro.2022.100256

Johnson, S., Kuhlen, L., Deme, J. C., Abrusci, P., & Lea, S. M. (2019). The Structure of an Injectisome Export Gate Demonstrates Conservation of Architecture in the Core Export Gate between Flagellar and Virulence Type III Secretion Systems. *mBio*, *10*(3). https://doi.org/10.1128/mbio.00818-19

Johnsson, N., & Varshavsky, A. (1994). Split ubiquitin as a sensor of protein interactions *in vivo*. *Proceedings of the National Academy of Sciences*, *91*(22), 10340–10344. https://doi.org/10.1073/pnas.91.22.10340

Journet, L., Agrain, C., Broz, P., & Cornelis, G. R. (2003). The Needle Length of Bacterial Injectisomes Is Determined by a Molecular Ruler. *Science*, *302*(5651), 1757–1760. https://doi.org/10.1126/science.1091422

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kago, G., Turnbough, C. L., Salazar, J. C., & Payne, S. M. (2024). (P)ppGpp is required for virulence of *Shigella flexneri*. *Infection and Immunity*, *92*(1), e0033423. https://doi.org/10.1128/iai.00334-23

Kaminski, R. W., Pavlinac, P. B., Platts-Mills, J. A., Rogawski McQuade, E. T., Hausdorff, W. P., Isbrucker, R. A., Vannice, K. S., Cavaleri, M., Kochhar, S., Mehring-LeDoare, K., Enwere, G., Wilder-Smith, A., Kotloff, K. L., Sow, S., & Giersing, B. K. (2025). WHO Workshop Report: Regulatory Science to Inform Clinical Pathways for *Shigella* Vaccines Intended for Use in Children in Low- and Middle-Income Countries. *Vaccines*, *13*(5), 439. https://doi.org/10.3390/vaccines13050439

Kane, C. D., Schuch, R., Day, W. A., & Maurelli, A. T. (2002). MxiE Regulates Intracellular Expression of Factors Secreted by the *Shigella flexneri* 2a Type III Secretion System.

Journal of Bacteriology, *184*(16), 4409–4419. https://doi.org/10.1128/JB.184.16.4409-4419.2002

Kane, K. A., & Dorman, C. J. (2012). VirB-Mediated Positive Feedback Control of the Virulence Gene Regulatory Cascade of *Shigella flexneri*. *Journal of Bacteriology*, *194*(19), 5264–5273. https://doi.org/10.1128/JB.00800-12

Keller, A., Nesvizhskii, A. I., Kolker, E., & Aebersold, R. (2002). Empirical Statistical Model to Estimate the Accuracy of Peptide Identifications Made by MS/MS and Database Search. *Analytical Chemistry*, *74*(20), 5383–5392. https://doi.org/10.1021/ac025747h

Kido, K., Yamanaka, S., Nakano, S., Motani, K., Shinohara, S., Nozawa, A., Kosako, H., Ito, S., & Sawasaki, T. (2020). AirID, a novel proximity biotinylation enzyme, for analysis of protein-protein interactions. *eLife*, *9*, e54983. https://doi.org/10.7554/eLife.54983

Kim, D. I., Jensen, S. C., Noble, K. A., Kc, B., Roux, K. H., Motamedchaboki, K., & Roux, K. J. (2016). An improved smaller biotin ligase for BioID proximity labeling. *Molecular Biology of the Cell*, *27*(8), 1188–1196. https://doi.org/10.1091/mbc.E15-12-0844

Kim, D. I., & Roux, K. J. (2016). Filling the Void: Proximity-Based Labeling of Proteins in Living Cells. *Trends in Cell Biology*, *26*(11), 804–817. https://doi.org/10.1016/j.tcb.2016.09.004

Kothary, M. H., & Babu, U. S. (2001). Infective Dose of Foodborne Pathogens in Volunteers: A Review. *Journal of Food Safety*, *21*(1), 49–68. https://doi.org/10.1111/j.1745-4565.2001.tb00307.x

Kotloff, K. L., Riddle, M. S., Platts-Mills, J. A., Pavlinac, P., & Zaidi, A. K. M. (2018). Shigellosis. *The Lancet*, *391*(10122), 801–812. https://doi.org/10.1016/s0140-6736(17)33296-8

Kubitz, L., Bitsch, S., Zhao, X., Schmitt, K., Deweid, L., Roehrig, A., Barazzone, E. C., Valerius, O., Kolmar, H., & Béthune, J. (2022). Engineering of ultraID, a compact and hyperactive enzyme for proximity-dependent biotinylation in living cells. *Communications Biology*, *5*(1), 657. https://doi.org/10.1038/s42003-022-03604-5

Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galán, J. E., & Aizawa, S. I. (1998). Supramolecular structure of the *Salmonella typhimurium* type III protein secretion system. *Science (New York, N.Y.)*, *280*(5363), 602–605. https://doi.org/10.1126/science.280.5363.602

Kuhlen, L., Johnson, S., Zeitler, A., Bäurle, S., Deme, J. C., Caesar, J. J. E., Debo, R., Fisher, J., Wagner, S., & Lea, S. M. (2020). The substrate specificity switch FlhB assembles onto the export gate to regulate type three secretion. *Nature Communications*, *11*(1), 1296. https://doi.org/10.1038/s41467-020-15071-9

Kwak, C., Shin, S., Park, J.-S., Jung, M., Nhung, T. T. M., Kang, M.-G., Lee, C., Kwon, T.-H., Park, S. K., Mun, J. Y., Kim, J.-S., & Rhee, H.-W. (2020). Contact-ID, a tool for profiling organelle contact sites, reveals regulatory proteins of mitochondrial-associated membrane formation. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(22), 12109–12120. https://doi.org/10.1073/pnas.1916584117

Lam, S. S., Martell, J. D., Kamer, K. J., Deerinck, T. J., Ellisman, M. H., Mootha, V. K., & Ting, A. Y. (2015). Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nature Methods*, *12*(1), 51–54. https://doi.org/10.1038/nmeth.3179

Lampel, K. A., Formal, S. B., & Maurelli, A. T. (2018). A Brief History of *Shigella*. *EcoSal Plus*, *8*(1). https://doi.org/10.1128/ecosalplus.esp-0006-2017

Lan, R., & Reeves, P. R. (2002). *Escherichia coli* in disguise: Molecular origins of *Shigella*. *Microbes and Infection*, *4*(11), 1125–1132. https://doi.org/10.1016/s1286-4579(02)01637-4

Lara-Tejero, M., Kato, J., Wagner, S., Liu, X., & Galán, J. E. (2011). A Sorting Platform Determines the Order of Protein Secretion in Bacterial Type III Systems. *Science*, *331*(6021), 1188–1191. https://doi.org/10.1126/science.1201476

LeBoeuf, E. N., & Grove, A. (2025). CysB in the Multiverse of Functions: Regulatory Roles in Cysteine Biosynthesis and Beyond. *Frontiers in Bioscience (Landmark Edition)*, *30*(8), 36563. https://doi.org/10.31083/FBL36563

Lee, C. C., Wood, M. D., Ng, K., Andersen, C. B., Liu, Y., Luginbühl, P., Spraggon, G., & Katagiri, F. (2004). Crystal Structure of the Type III Effector AvrB from *Pseudomonas syringae*. *Structure*, *12*(3), 487–494. https://doi.org/10.1016/j.str.2004.02.013

Leow, C. H., Kazi, A., Chuah, C., Xu, Z., Singh, K. K. B., & Leow, C. Y. (2024). Shigellosis: The Road from Multiple Drug-Resistance (MDR) to Vaccine Development. In *Integrated Science* (pp. 209–237). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-64288-3_10

Leung, V., Ritchie, G., Stefanovic, A., Lee, C., Chorlton, S., Matic, N., Romney, M. G., Hayden, A., & Lowe, C. F. (2025). An Outbreak of Multidrug-Resistant *Shigella flexneri* Serotype 2a Among People Experiencing Homelessness in Vancouver. *Tropical Medicine and Infectious Disease*, *10*(5), 120. https://doi.org/10.3390/tropicalmed10050120

Lilic, M., Vujanac, M., & Stebbins, C. E. (2006). A Common Structural Motif in the Binding of Virulence Factors to Bacterial Secretion Chaperones. *Molecular Cell*, *21*(5), 653–664. https://doi.org/10.1016/j.molcel.2006.01.026

Lin, Z., Cai, X., Chen, M., Ye, L., Wu, Y., Wang, X., Lv, Z., Shang, Y., & Qu, D. (2018). Virulence and Stress Responses of *Shigella flexneri* Regulated by PhoP/PhoQ. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.02689

Lithgow, J. K., Haider, F., Roberts, I. S., & Green, J. (2007). Alternate SlyA and H-NS nucleoprotein complexes control *hlyE* expression in *Escherichia coli* K-12. *Molecular Microbiology*, *66*(3), 685–698. https://doi.org/10.1111/j.1365-2958.2007.05950.x

Liu, Q., Zheng, J., Sun, W., Huo, Y., Zhang, L., Hao, P., Wang, H., & Zhuang, M. (2018). A proximity-tagging system to identify membrane protein–protein interactions. *Nature Methods*, *15*(9), 715–722. https://doi.org/10.1038/s41592-018-0100-5

Lord, D. M., Uzgoren Baran, A., Soo, V. W. C., Wood, T. K., Peti, W., & Page, R. (2014). McbR/YncC: Implications for the mechanism of ligand and DNA binding by a bacterial GntR transcriptional regulator involved in biofilm formation. *Biochemistry*, *53*(46), 7223–7231. https://doi.org/10.1021/bi500871a

Lorenz, C., & Büttner, D. (2009). Functional Characterization of the Type III Secretion ATPase HrcN from the Plant Pathogen *Xanthomonas campestris* pv. Vesicatoria. *Journal of Bacteriology*, *191*(5), 1414–1428. https://doi.org/10.1128/JB.01446-08

Lu, Y., Rashidul, I. M., Hirata, H., & Tsuyumu, S. (2011). KdgR, an IClR Family Transcriptional Regulator, Inhibits Virulence Mainly by Repression of hrp Genes in *Xanthomonas oryzae* pv. Oryzae. *Journal of Bacteriology*, *193*(23), 6674–6682. https://doi.org/10.1128/JB.05714-11

Lunelli, M., Kamprad, A., Bürger, J., Mielke, T., Spahn, C. M. T., & Kolbe, M. (2020). Cryo-EM structure of the *Shigella* type III needle complex. *PLoS Pathogens*, *16*(2), e1008263. https://doi.org/10.1371/journal.ppat.1008263

Majewski, D. D., Worrall, L. J., Hong, C., Atkinson, C. E., Vuckovic, M., Watanabe, N., Yu, Z., & Strynadka, N. C. J. (2019). Cryo-EM structure of the homohexameric T3SS ATPase-central stalk complex reveals rotary ATPase-like asymmetry. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-08477-7

Manigat, F. O., Connell, L. B., Stewart, B. N., LePabic, A.-R., Tessier, C. J. G., Emlaw, J. R., Calvert, N. D., Rössl, A., Shuhendler, A. J., daCosta, C. J. B., & Campbell-Valois, F.-X. (2024). pUdOs: Concise Plasmids for Bacterial and Mammalian Cells. *ACS Synthetic Biology*, *13*(2), 485–497. https://doi.org/10.1021/acssynbio.3c00408

Mannix, K. M., Starble, R. M., Kaufman, R. S., & Cooley, L. (2019). Proximity labeling reveals novel interactomes in live Drosophila tissue. *Development*, *146*(14), dev176644. https://doi.org/10.1242/dev.176644

Marcos-Vilchis, A., Espinosa, N., Alvarez, A. F., Puente, J. L., Soto, J. E., & González-Pedrajo, B. (2025). On the role of the sorting platform in hierarchical type III secretion regulation in enteropathogenic *Escherichia coli*. *Journal of Bacteriology*, *207*(3), e00446-24. https://doi.org/10.1128/jb.00446-24

Martell, J. D., Deerinck, T. J., Sancak, Y., Poulos, T. L., Mootha, V. K., Sosinsky, G. E., Ellisman, M. H., & Ting, A. Y. (2012). Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nature Biotechnology*, *30*(11), 1143–1148. https://doi.org/10.1038/nbt.2375

Marteyn, B., West, N. P., Browning, D. F., Cole, J. A., Shaw, J. G., Palm, F., Mounier, J., Prévost, M.-C., Sansonetti, P., & Tang, C. M. (2010). Modulation of *Shigella* virulence in response to available oxygen *in vivo*. *Nature*, *465*(7296), 355–358. https://doi.org/10.1038/nature08970

Martyn, J. E., Pilla, G., Hollingshead, S., Winther, K. S., Lea, S., McVicker, G., & Tang, C. M. (2022). Maintenance of the *Shigella sonnei* Virulence Plasmid Is Dependent on Its Repertoire and Amino Acid Sequence of Toxin-Antitoxin Systems. *Journal of Bacteriology*, *204*(3). https://doi.org/10.1128/jb.00519-21

Mattock, E., & Blocker, A. J. (2017). How Do the Virulence Factors of *Shigella* Work Together to Cause Disease? *Frontiers in Cellular and Infection Microbiology*, *7*. https://doi.org/10.3389/fcimb.2017.00064

Maurelli, A. T. (2007). Black holes, antivirulence genes, and gene inactivation in the evolution of bacterial pathogens. *FEMS Microbiology Letters*, *267*(1), 1–8. https://doi.org/10.1111/j.1574-6968.2006.00526.x

Maurelli, A. T., Fernández, R. E., Bloch, C. A., Rode, C. K., & Fasano, A. (1998). "Black holes" and bacterial pathogenicity: A large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proceedings of the National Academy of Sciences*, *95*(7), 3943–3948. https://doi.org/10.1073/pnas.95.7.3943

Mavris, M., Page, A., Tournebize, R., Demers, B., Sansonetti, P., & Parsot, C. (2002). Regulation of transcription by the activity of the *Shigella flexneri* type III secretion apparatus. *Molecular Microbiology*, *43*(6), 1543–1553. https://doi.org/10.1046/j.1365-2958.2002.02836.x

Mavris, M., Sansonetti, P. J., & Parsot, C. (2002). Identification of the *cis* -Acting Site Involved in Activation of Promoters Regulated by Activity of the Type III Secretion Apparatus in *Shigella flexneri*. *Journal of Bacteriology*, *184*(24), 6751–6759. https://doi.org/10.1128/JB.184.24.6751-6759.2002

McKenna, J. A., Karney, M. M. A., Chan, D. K., Weatherspoon-Griffin, N., Becerra Larios, B., Pilonieta, M. C., Munson, G. P., & Wing, H. J. (2022). The AraC/XylS Protein MxiE and Its Coregulator IpgC Control a Negative Feedback Loop in the Transcriptional Cascade That Regulates Type III Secretion in *Shigella flexneri*. *Journal of Bacteriology*, *204*(7). https://doi.org/10.1128/jb.00137-22

McNorton, M. M., & Maier, R. J. (2012). Roles of H2 uptake hydrogenases in *Shigella flexneri* acid tolerance. *Microbiology (Reading, England)*, *158*(Pt 8), 2204–2212. https://doi.org/10.1099/mic.0.058248-0

McVicker, G., Hollingshead, S., Pilla, G., & Tang, C. M. (2019). Maintenance of the virulence plasmid in *Shigella flexneri* is influenced by Lon and two functional partitioning systems. *Molecular Microbiology*, *111*(5), 1355–1366. https://doi.org/10.1111/mmi.14225

Medicine NLo. (2020). *NCT04602975—A study to investigate the safety and immunogenicity of the SF2a-TT15 synthetic carbohydrate-based conjugate vaccine against Shigella Flexneri 2a (GlycoShig3)*. National Library of Medicine. https://clinicaltrials.gov/ct2/show/NCT04602975

Medicine NLo. (2021). *NCT05073003—A study on the safety and immune responses to the GVGH altSonflex1–2–3 vaccine against Shigellosis in adults, children, and infants.* National Library of Medicine. https://clinicaltrials.gov/ct2/show/NCT05073003

Mellouk, N., Weiner, A., Aulner, N., Schmitt, C., Elbaum, M., Shorte, S. L., Danckaert, A., & Enninga, J. (2014). *Shigella* Subverts the Host Recycling Compartment to Rupture Its Vacuole. *Cell Host & Microbe*, *16*(4), 517–530. https://doi.org/10.1016/j.chom.2014.09.005

Ménard, R., Sansonetti, P. J., & Parsot, C. (1993). Nonpolar mutagenesis of the ipa genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. *Journal of Bacteriology*, *175*(18), 5899–5906. https://doi.org/10.1128/jb.175.18.5899-5906.1993

Ménard, R., Sansonetti, P., & Parsot, C. (1994). The secretion of the *Shigella flexneri* Ipa invasins is activated by epithelial cells and controlled by IpaB and IpaD. *The EMBO Journal*, *13*(22), 5293–5302. https://doi.org/10.1002/j.1460-2075.1994.tb06863.x

Ménard, R., Sansonetti, P., Parsot, C., & Vasselon, T. (1994). Extracellular association and cytoplasmic partitioning of the IpaB and IpaC invasins of *S. flexneri*. *Cell*, *79*(3), 515–525. https://doi.org/10.1016/0092-8674(94)90260-7

Miller, H. K., Kwuan, L., Schwiesow, L., Bernick, D. L., Mettert, E., Ramirez, H. A., Ragle, J. M., Chan, P. P., Kiley, P. J., Lowe, T. M., & Auerbuch, V. (2014). IscR Is Essential for *Yersinia pseudotuberculosis* Type III Secretion and Virulence. *PLoS Pathogens*, *10*(6), e1004194. https://doi.org/10.1371/journal.ppat.1004194

Mo, Y., Fang, W., Li, H., Chen, J., Hu, X., Wang, B., Feng, Z., Shi, H., He, Y., Huang, D., Mo, Z., Ye, Q., & Du, L. (2021). Safety and Immunogenicity of a *Shigella* Bivalent Conjugate Vaccine (ZF0901) in 3-Month- to 5-Year-Old Children in China. *Vaccines*, *10*(1), 33. https://doi.org/10.3390/vaccines10010033

Molina-Henares, A. J., Krell, T., Eugenia Guazzaroni, M., Segura, A., & Ramos, J. L. (2006). Members of the IclR family of bacterial transcriptional regulators function as activators

and/or repressors. *FEMS Microbiology Reviews*, *30*(2), 157–186. https://doi.org/10.1111/j.1574-6976.2005.00008.x

Moss, J. E., Cardozo, T. J., Zychlinsky, A., & Groisman, E. A. (1999). The *selC*-associated SHI-2 pathogenicity island of *Shigella flexneri*. *Molecular Microbiology*, *33*(1), 74–83. https://doi.org/10.1046/j.1365-2958.1999.01449.x

Murphy, E. R., & Payne, S. M. (2007). RyhB, an Iron-Responsive Small RNA Molecule, Regulates *Shigella dysenteriae* Virulence. *Infection and Immunity*, *75*(7), 3470–3477. https://doi.org/10.1128/IAI.00112-07

Nakayama, S., & Watanabe, H. (1995). Involvement of *cpxA*, a sensor of a two-component regulatory system, in the pH-dependent regulation of expression of *Shigella sonnei virF* gene. *Journal of Bacteriology*, *177*(17), 5062–5069. https://doi.org/10.1128/jb.177.17.5062-5069.1995

Nasser, W., Reverchon, S., & Robert-Baudouy, J. (1992). Purification and functional characterization of the KdgR protein, a major repressor of pectinolysis genes of *Erwinia chrysanthemi*. *Molecular Microbiology*, *6*(2), 257–265. https://doi.org/10.1111/j.1365-2958.1992.tb02007.x

NCCID. (2024, October 28). *Shigellosis*. National Collaborating Centre for Infectious Diseases. https://nccid.ca/debrief/shigellosis/

Nemhauser, J. B. & Centers for Disease Control (U.S.) (Eds.). (2023). *CDC yellow book 2024: Health information for international travel*. Oxford University Press. https://doi.org/10.1093/oso/9780197570944.001.0001

Nesvizhskii, A. I., Keller, A., Kolker, E., & Aebersold, R. (2003). A Statistical Model for Identifying Proteins by Tandem Mass Spectrometry. *Analytical Chemistry*, *75*(17), 4646–4658. https://doi.org/10.1021/ac0341261

Nyarkoh, R., Odoom, A., & Donkor, E. S. (2024). Prevalence of *Shigella* species and antimicrobial resistance patterns in Africa: Systematic review and meta-analysis. *BMC Infectious Diseases*, *24*(1). https://doi.org/10.1186/s12879-024-09945-2

Ogawa, M., Yoshimori, T., Suzuki, T., Sagara, H., Mizushima, N., & Sasakawa, C. (2005). Escape of Intracellular *Shigella* from Autophagy. *Science*, *307*(5710), 727–731. https://doi.org/10.1126/science.1106036

Ojha, R., Dittmar, A. A., Severin, G. B., & Koestler, B. J. (2021). *Shigella flexneri* Diguanylate Cyclases Regulate Virulence. *Journal of Bacteriology*, *203*(23), e0024221. https://doi.org/10.1128/JB.00242-21

Okui, N. (2024). Dr. Kiyoshi Shiga (1871-1957): Outstanding Bacteriologist Who Discovered Dysentery Bacillus and Contributed Immensely to Public Health in Japan. *Cureus*, *16*(10), e71881. https://doi.org/10.7759/cureus.71881

Olive, A. J., Kenjale, R., Espina, M., Moore, D. S., Picking, W. L., & Picking, W. D. (2007). Bile Salts Stimulate Recruitment of IpaB to the *Shigella flexneri* Surface, Where It Colocalizes with IpaD at the Tip of the Type III Secretion Needle. *Infection and Immunity*, *75*(5), 2626–2629. https://doi.org/10.1128/IAI.01599-06

Olson, M. G., Widner, R. E., Jorgenson, L. M., Lawrence, A., Lagundzin, D., Woods, N. T., Ouellette, S. P., & Rucks, E. A. (2019). 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*(11), e00537-19. https://doi.org/10.1128/IAI.00537-19

Osler, W. (1892). *The principles and practice of medicine*. D Appleton and Company, New York,.

Page, A.-L., Ohayon, H., Sansonetti, P. J., & Parsot, C. (1999). The secreted IpaB and IpaC invasins and their cytoplasmic chaperone IpgC are required for intercellular dissemination of *Shigella flexneri*. *Cellular Microbiology*, *1*(2), 183–193. https://doi.org/10.1046/j.1462-5822.1999.00019.x

PAHO WHO. (2022, June 6). *Epidemiological Alert: Emergence and spread of Shigella sonnei with extreme resistance to antibiotics. Potential risk for Latin America and the Caribbean. 6 June 2022 - PAHO/WHO | Pan American Health Organization*. https://www.paho.org/en/documents/epidemiological-alert-emergence-and-spread-shigella-sonnei-extreme-resistance-antibiotics

Parsot, C., Ageron, E., Penno, C., Mavris, M., Jamoussi, K., d'Hauteville, H., Sansonetti, P., & Demers, B. (2005). A secreted anti-activator, OspD1, and its chaperone, Spa15, are involved in the control of transcription by the type III secretion apparatus activity in *Shigella flexneri*. *Molecular Microbiology*, *56*(6), 1627–1635. https://doi.org/10.1111/j.1365-2958.2005.04645.x

Parsot, C., Ménard, R., Gounon, P., & Sansonetti, P. J. (1995). Enhanced secretion through the *Shigella flexneri* Mxi-Spa translocon leads to assembly of extracellular proteins into macromolecular structures. *Molecular Microbiology*, *16*(2), 291–300. https://doi.org/10.1111/j.1365-2958.1995.tb02301.x

Pasqua, M., Grossi, M., Scinicariello, S., Aussel, L., Barras, F., Colonna, B., & Prosseda, G. (2019). The MFS efflux pump EmrKY contributes to the survival of *Shigella* within macrophages. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-39749-3

Pelletier, J. N., Campbell-Valois, F.-X., & Michnick, S. W. (1998). Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments.

Proceedings of the National Academy of Sciences, *95*(21), 12141–12146. https://doi.org/10.1073/pnas.95.21.12141

Picker, M., & Wing, H. (2016). H-NS, Its Family Members and Their Regulation of Virulence Genes in *Shigella* Species. *Genes*, *7*(12), 112. https://doi.org/10.3390/genes7120112

Picking, W. L., & Picking, W. D. (2016). The Many Faces of IpaB. *Frontiers in Cellular and Infection Microbiology*, *6*. https://doi.org/10.3389/fcimb.2016.00012

Pilla, G., McVicker, G., & Tang, C. M. (2017). Genetic plasticity of the *Shigella* virulence plasmid is mediated by intra- and inter-molecular events between insertion sequences. *PLOS Genetics*, *13*(9), e1007014. https://doi.org/10.1371/journal.pgen.1007014

Pinaud, L., Ferrari, M. L., Friedman, R., Jehmlich, N., Von Bergen, M., Phalipon, A., Sansonetti, P. J., & Campbell-Valois, F.-X. (2017). Identification of novel substrates of *Shigella* T3SA through analysis of its virulence plasmid-encoded secretome. *PLOS ONE*, *12*(10), e0186920. https://doi.org/10.1371/journal.pone.0186920

Portaliou, A. G., Tsolis, K. C., Loos, M. S., Zorzini, V., & Economou, A. (2016). Type III Secretion: Building and Operating a Remarkable Nanomachine. *Trends in Biochemical Sciences*, *41*(2), 175–189. https://doi.org/10.1016/j.tibs.2015.09.005

Prosseda, G., Di Martino, M. L., Campilongo, R., Fioravanti, R., Micheli, G., Casalino, M., & Colonna, B. (2012). Shedding of genes that interfere with the pathogenic lifestyle: The *Shigella* model. *Research in Microbiology*, *163*(6–7), 399–406. https://doi.org/10.1016/j.resmic.2012.07.004

Prunier, A.-L., Schuch, R., Fernández, R. E., Mumy, K. L., Kohler, H., McCormick, B. A., & Maurelli, A. T. (2007). *nadA* and *nadB* of *Shigella flexneri* 5a are antivirulence loci responsible for the synthesis of quinolinate, a small molecule inhibitor of *Shigella* pathogenicity. *Microbiology (Reading, England)*, *153*(Pt 7), 2363–2372. https://doi.org/10.1099/mic.0.2007/006916-0

Public Health Agency of Canada. (2024). *Shigella spp.: Infectious substances pathogen safety data sheet*. https://www.canada.ca/en/public-health/services/laboratory-biosafety-biosecurity/pathogen-safety-data-sheets-risk-assessment/shigella.html

Pupo, G. M., Lan, R., & Reeves, P. R. (2000). Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. *Proceedings of the National Academy of Sciences*, *97*(19), 10567–10572. https://doi.org/10.1073/pnas.180094797

Rajakumar, K., Sasakawa, C., & Adler, B. (1997). Use of a novel approach, termed island probing, identifies the *Shigella flexneri* she pathogenicity island which encodes a homolog of the immunoglobulin A protease-like family of proteins. *Infection and Immunity*, *65*(11), 4606–4614. https://doi.org/10.1128/iai.65.11.4606-4614.1997

Ramanathan, M., Majzoub, K., Rao, D. S., Neela, P. H., Zarnegar, B. J., Mondal, S., Roth, J. G., Gai, H., Kovalski, J. R., Siprashvili, Z., Palmer, T. D., Carette, J. E., & Khavari, P. A. (2018). RNA–protein interaction detection in living cells. *Nature Methods*, *15*(3), 207–212. https://doi.org/10.1038/nmeth.4601

Ramirez, C. A., Egetemaier, S., & Béthune, J. (2021). Context-Specific and Proximity-Dependent Labeling for the Proteomic Analysis of Spatiotemporally Defined Protein Complexes with Split-BioID. In *Methods in Molecular Biology* (pp. 303–318). Springer US. https://doi.org/10.1007/978-1-0716-1126-5_17

Ramos Moreno, A. C., Goncalves Ferreira, L., & Baquerizo Martinez, M. (2009). Enteroinvasive *Escherichia coli* vs. *Shigella flexneri*: How different patterns of gene expression affect virulence. *FEMS Microbiology Letters*, *301*(2), 156–163. https://doi.org/10.1111/j.1574-6968.2009.01815.x

Raso, M. M., Arato, V., Gasperini, G., & Micoli, F. (2023). Toward a *Shigella* Vaccine: Opportunities and Challenges to Fight an Antimicrobial-Resistant Pathogen. *International Journal of Molecular Sciences*, *24*(5), 4649. https://doi.org/10.3390/ijms24054649

Remy, O., Santin, Y. G., Jonckheere, V., Tesseur, C., Kaljević, J., Van Damme, P., & Laloux, G. (2024). Distinct dynamics and proximity networks of hub proteins at the prey-invading cell pole in a predatory bacterium. *Journal of Bacteriology*, *206*(4). https://doi.org/10.1128/jb.00014-24

Rey, C., Chang, Y.-Y., Latour-Lambert, P., Varet, H., Proux, C., Legendre, R., Coppée, J.-Y., & Enninga, J. (2020). Transcytosis subversion by M cell-to-enterocyte spread promotes *Shigella flexneri* and *Listeria monocytogenes* intracellular bacterial dissemination. *PLOS Pathogens*, *16*(4), e1008446. https://doi.org/10.1371/journal.ppat.1008446

Riccardi, E., Van Mastbergen, E. C., Navarre, W. W., & Vreede, J. (2019). Predicting the mechanism and rate of H-NS binding to AT-rich DNA. *PLOS Computational Biology*, *15*(3), e1006845. https://doi.org/10.1371/journal.pcbi.1006845

Richard, H., & Foster, J. W. (2004). *Escherichia coli* glutamate- and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential. *Journal of Bacteriology*, *186*(18), 6032–6041. https://doi.org/10.1128/JB.186.18.6032-6041.2004

Risha, Y., Minic, Z., Ghobadloo, S. M., & Berezovski, M. V. (2020). The proteomic analysis of breast cell line exosomes reveals disease patterns and potential biomarkers. *Scientific Reports*, *10*(1), 13572. https://doi.org/10.1038/s41598-020-70393-4

Roehrich, A. D., Bordignon, E., Mode, S., Shen, D.-K., Liu, X., Pain, M., Murillo, I., Martinez-Argudo, I., Sessions, R. B., & Blocker, A. J. (2017). Steps for *Shigella* Gatekeeper Protein MxiC Function in Hierarchical Type III Secretion Regulation. *The Journal of Biological Chemistry*, *292*(5), 1705–1723. https://doi.org/10.1074/jbc.M116.746826

Roux, K. J., Kim, D. I., Raida, M., & Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *Journal of Cell Biology*, *196*(6), 801–810. https://doi.org/10.1083/jcb.201112098

Runyen-Janecky, L., Daugherty, A., Lloyd, B., Wellington, C., Eskandarian, H., & Sagransky, M. (2008). Role and Regulation of Iron-Sulfur Cluster Biosynthesis Genes in *Shigella flexneri* Virulence. *Infection and Immunity*, *76*(3), 1083–1092. https://doi.org/10.1128/IAI.01211-07

Samavarchi-Tehrani, P., Samson, R., & Gingras, A.-C. (2020). Proximity Dependent Biotinylation: Key Enzymes and Adaptation to Proteomics Approaches. *Molecular & Cellular Proteomics: MCP*, *19*(5), 757–773. https://doi.org/10.1074/mcp.R120.001941

Saner, N., Uzun, C., Kırım, B. A., Özkan, S. N., Geiszler, D. J., Öztürk, E., Tunçbağ, N., & Özlü, N. (2025). Proximity Labeling and SILAC-Based Proteomic Approach Identifies Proteins at the Interface of Homotypic and Heterotypic Cancer Cell Interactions. *Molecular & Cellular Proteomics*, *24*(6), 100986. https://doi.org/10.1016/j.mcpro.2025.100986

Sansonetti, P. J., Kopecko, D. J., & Formal, S. B. (1982). Involvement of a plasmid in the invasive ability of *Shigella flexneri*. *Infection and Immunity*, *35*(3), 852–860. https://doi.org/10.1128/iai.35.3.852-860.1982

Schopp, I. M., Amaya Ramirez, C. C., Debeljak, J., Kreibich, E., Skribbe, M., Wild, K., & Béthune, J. (2017). Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. *Nature Communications*, *8*, 15690. https://doi.org/10.1038/ncomms15690

Scott, T. A., Baker, K. S., Trotter, C., Jenkins, C., Mostowy, S., Hawkey, J., Schmidt, H., Holt, K. E., Thomson, N. R., & Baker, S. (2025). *Shigella sonnei*: Epidemiology, evolution, pathogenesis, resistance and host interactions. *Nature Reviews Microbiology*, *23*(5), 303–317. https://doi.org/10.1038/s41579-024-01126-x

Sebastian, H., & Finkel, S. E. (2022). Absence of Ribosome Modulation Factor Alters Growth and Competitive Fitness of *Escherichia coli*. *Microbiology Spectrum*, *10*(2), e02239-21. https://doi.org/10.1128/spectrum.02239-21

Shen, B. A., Hustmyer, C. M., Roston, D., Wolfe, M. B., & Landick, R. (2022). Bacterial H-NS contacts DNA at the same irregularly spaced sites in both bridged and hemi-sequestered linear filaments. *iScience*, *25*(6), 104429. https://doi.org/10.1016/j.isci.2022.104429

Shen, D.-K., & Blocker, A. J. (2016). MxiA, MxiC and IpaD Regulate Substrate Selection and Secretion Mode in the T3SS of *Shigella flexneri*. *PloS One*, *11*(5), e0155141. https://doi.org/10.1371/journal.pone.0155141

Shere, K. D., Sallustio, S., Manessis, A., D'Aversa, T. G., & Goldberg, M. B. (1997). Disruption of IcsP, the major *Shigella* protease that cleaves IcsA, accelerates actin-based motility.

Molecular Microbiology, *25*(3), 451–462. https://doi.org/10.1046/j.1365-2958.1997.4681827.x

Shiga, K. (1936). The Trend of Prevention, Therapy and Epidemiology of Dysentery since the Discovery of Its Causative Organism. *New England Journal of Medicine*, *215*(26), 1205–1211. https://doi.org/10.1056/nejm193612242152602

Shkel, O., Kharkivska, Y., Kim, Y. K., & Lee, J. (2022). Proximity Labeling Techniques: A Multi-Omics Toolbox. *Chemistry – An Asian Journal*, *17*(2), e202101240. https://doi.org/10.1002/asia.202101240

Shteynberg, D., Deutsch, E. W., Lam, H., Eng, J. K., Sun, Z., Tasman, N., Mendoza, L., Moritz, R. L., Aebersold, R., & Nesvizhskii, A. I. (2011). iProphet: Multi-level Integrative Analysis of Shotgun Proteomic Data Improves Peptide and Protein Identification Rates and Error Estimates. *Molecular & Cellular Proteomics*, *10*(12), M111.007690. https://doi.org/10.1074/mcp.M111.007690

Silué, N., & Campbell-Valois, F.-X. (2022). *icaR* and *icaT* are Ancient Chromosome Genes Encoding Substrates of the Type III Secretion Apparatus in *Shigella flexneri*. *mSphere*, *7*(3), e0011522. https://doi.org/10.1128/msphere.00115-22

Silué, N., Marcantonio, E., & Campbell-Valois, F.-X. (2020). RNA-Seq analysis of the T3SA regulon in *Shigella flexneri* reveals two new chromosomal genes upregulated in the on-state. *Methods*, *176*, 71–81. https://doi.org/10.1016/j.ymeth.2019.03.017

Simoens, L., Jonckheere, V., & Van Damme, P. (2025). BioID in Bacteria: Selection of a Suitable Biotin Ligase for Proxeome Mapping. In P. Van Damme (Ed.), *Proximity-Dependent Protein Biotinylation* (Vol. 2953, pp. 17–39). Springer US. https://doi.org/10.1007/978-1-0716-4694-6_2

Small, P., Blankenhorn, D., Welty, D., Zinser, E., & Slonczewski, J. L. (1994). Acid and base resistance in *Escherichia coli* and *Shigella flexneri*: Role of *rpoS* and growth pH. *Journal of Bacteriology*, *176*(6), 1729–1737. https://doi.org/10.1128/jb.176.6.1729-1737.1994

Snauwaert, V., Jonckheere, V., & Van Damme, P. (2025). Unraveling N-Terminal Proteoform Interactomes via Multiplexed Recombineering in *Salmonella*. In P. Van Damme (Ed.), *Proximity-Dependent Protein Biotinylation* (Vol. 2953, pp. 81–102). Springer US. https://doi.org/10.1007/978-1-0716-4694-6_6

Song, Y., Yang, C., Chen, G., Zhang, Y., Seng, Z., Cai, Z., Zhang, C., Yang, L., Gan, J., & Liang, H. (2019). Molecular insights into the master regulator CysB-mediated bacterial virulence in *Pseudomonas aeruginosa*. *Molecular Microbiology*, *111*(5), 1195–1210. https://doi.org/10.1111/mmi.14200

Stenhouse, G. E., Keddy, K. H., Bengtsson, R. J., Hall, N., Smith, A. M., Thomas, J., Iturriza-Gómara, M., & Baker, K. S. (2023). The genomic epidemiology of shigellosis in South Africa. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-43345-5

Stévenin, V., Chang, Y.-Y., Le Toquin, Y., Duchateau, M., Gianetto, Q. G., Luk, C. H., Salles, A., Sohst, V., Matondo, M., Reiling, N., & Enninga, J. (2019). Dynamic Growth and Shrinkage of the *Salmonella*-Containing Vacuole Determines the Intracellular Pathogen Niche. *Cell Reports*, *29*(12), 3958-3973.e7. https://doi.org/10.1016/j.celrep.2019.11.049

Streaker, E. D., & Beckett, D. (2003). Coupling of Protein Assembly and DNA Binding: Biotin Repressor Dimerization Precedes Biotin Operator Binding. *Journal of Molecular Biology*, *325*(5), 937–948. https://doi.org/10.1016/S0022-2836(02)01308-6

Tesseur, C., Remy, O., Laloux, G., & Santin, Y. (2025). Identification of In Vivo Protein Networks Using miniTurbo Proximity Labeling in Bacteria. In *Methods in Molecular Biology* (pp. 71–80). Springer US. https://doi.org/10.1007/978-1-0716-4694-6_5

The, H. C., Thanh, D. P., Holt, K. E., Thomson, N. R., & Baker, S. (2016). The genomic signatures of *Shigella* evolution, adaptation and geographical spread. *Nature Reviews. Microbiology*, *14*(4), 235–250. https://doi.org/10.1038/nrmicro.2016.10

Torres-Vargas, C. E., Kronenberger, T., Roos, N., Dietsche, T., Poso, A., & Wagner, S. (2019). The inner rod of virulence-associated type III secretion systems constitutes a needle adapter of one helical turn that is deeply integrated into the system's export apparatus. *Molecular Microbiology*, *112*(3), 918–931. https://doi.org/10.1111/mmi.14327

Traxler, M. F., Zacharia, V. M., Marquardt, S., Summers, S. M., Nguyen, H., Stark, S. E., & Conway, T. (2011). Discretely calibrated regulatory loops controlled by ppGpp partition gene induction across the 'feast to famine' gradient in *Escherichia coli*. *Molecular Microbiology*, *79*(4), 830–845. https://doi.org/10.1111/j.1365-2958.2010.07498.x

Troeger, C., Blacker, B. F., Khalil, I. A., Rao, P. C., Cao, S., Zimsen, S. R., Albertson, S. B., Stanaway, J. D., Deshpande, A., Abebe, Z., Alvis-Guzman, N., Amare, A. T., Asgedom, S. W., Anteneh, Z. A., Antonio, C. A. T., Aremu, O., Asfaw, E. T., Atey, T. M., Atique, S., … Reiner, R. C. (2018). Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Infectious Diseases*, *18*(11), 1211–1228. https://doi.org/10.1016/S1473-3099(18)30362-1

Trofa, A. F., Ueno-Olsen, H., Oiwa, R., & Yoshikawa, M. (1999). Dr. Kiyoshi Shiga: Discoverer of the Dysentery Bacillus. *Clinical Infectious Diseases*, *29*(5), 1303–1306. https://doi.org/10.1086/313437

Tucker, D. L., Tucker, N., Ma, Z., Foster, J. W., Miranda, R. L., Cohen, P. S., & Conway, T. (2003). Genes of the GadX-GadW Regulon in *Escherichia coli*. *Journal of Bacteriology*, *185*(10), 3190–3201. https://doi.org/10.1128/JB.185.10.3190-3201.2003

Turbyfill, K. R., Clarkson, K. A., Oaks, E. V., Zurawski, D. V., Vortherms, A. R., & Kaminski, R. W. (2023). Development of the *Shigella flexneri* 2a, 3a, 6, and *S. sonnei* artificial Invaplex vaccines. *mSphere*, *8*(4). https://doi.org/10.1128/msphere.00073-23

Valdivia, R. H., & Falkow, S. (1996). Bacterial genetics by flow cytometry: Rapid isolation of *Salmonella typhimurium* acid-inducible promoters by differential fluorescence induction. *Molecular Microbiology*, *22*(2), 367–378. https://doi.org/10.1046/j.1365-2958.1996.00120.x

Valenzuela-Valderas, K. N., Farrashzadeh, E., Chang, Y.-Y., Shi, Y., Raudonis, R., Leung, B. M., Rohde, J. R., Enninga, J., & Cheng, Z. (2023). RACK1 promotes *Shigella flexneri* actin-mediated invasion, motility, and cell-to-cell spreading. *iScience*, *26*(11), 108216. https://doi.org/10.1016/j.isci.2023.108216

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., … Velankar, S. (2022). AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, *50*(D1), D439–D444. https://doi.org/10.1093/nar/gkab1061

Veenendaal, A. K. J., Hodgkinson, J. L., Schwarzer, L., Stabat, D., Zenk, S. F., & Blocker, A. J. (2007). The type III secretion system needle tip complex mediates host cell sensing and translocon insertion. *Molecular Microbiology*, *63*(6), 1719–1730. https://doi.org/10.1111/j.1365-2958.2007.05620.x

Venkatesan, M. M., Goldberg, M. B., Rose, D. J., Grotbeck, E. J., Burland, V., & Blattner, F. R. (2001). Complete DNA Sequence and Analysis of the Large Virulence Plasmid of *Shigella flexneri*. *Infection and Immunity*, *69*(5), 3271–3285. https://doi.org/10.1128/IAI.69.5.3271-3285.2001

Vergara-Irigaray, M., Fookes, M. C., Thomson, N. R., & Tang, C. M. (2014). RNA-seq analysis of the influence of anaerobiosis and FNR on *Shigella flexneri*. *BMC Genomics*, *15*(1), 438. https://doi.org/10.1186/1471-2164-15-438

Vergnes, A., Viala, J. P. M., Ouadah-Tsabet, R., Pocachard, B., Loiseau, L., Méresse, S., Barras, F., & Aussel, L. (2017). The iron-sulfur cluster sensor IscR is a negative regulator of Spi1 type III secretion system in *Salmonella enterica*. *Cellular Microbiology*, *19*(4), e12680. https://doi.org/10.1111/cmi.12680

Wagner, S., Grin, I., Malmsheimer, S., Singh, N., Torres-Vargas, C. E., & Westerhausen, S. (2018). Bacterial type III secretion systems: A complex device for the delivery of bacterial effector proteins into eukaryotic host cells. *FEMS Microbiology Letters*, *365*(19). https://doi.org/10.1093/femsle/fny201

Walker, R. I. (2025). Conserved antigens for enteric vaccines. *Vaccine*, *50*, 126828. https://doi.org/10.1016/j.vaccine.2025.126828

Wang, T., Du, X., Ji, L., Han, Y., Dang, J., Wen, J., Wang, Y., Pu, Q., Wu, M., & Liang, H. (2021). *Pseudomonas aeruginosa* T6SS-mediated molybdate transport contributes to bacterial competition during anaerobiosis. *Cell Reports*, *35*(2), 108957. https://doi.org/10.1016/j.celrep.2021.108957

Weatherspoon-Griffin, N., Picker, M. A., Pew, K. L., Park, H. S., Ginete, D. R., Karney, M. M., Usufzy, P., Castellanos, M. I., Duhart, J. C., Harrison, D. J., Socea, J. N., Karabachev, A. D., Hensley, C. T., Howerton, A. J., Ojeda-Daulo, R., Immak, J. A., & Wing, H. J. (2018). Insights into transcriptional silencing and anti-silencing in *Shigella flexneri*: A detailed molecular analysis of the *icsP* virulence locus. *Molecular Microbiology*, *108*(5), 505–518. https://doi.org/10.1111/mmi.13932

Weatherspoon-Griffin, N., & Wing, H. J. (2016). Characterization of SlyA in *Shigella flexneri* Identifies a Novel Role in Virulence. *Infection and Immunity*, *84*(4), 1073–1082. https://doi.org/10.1128/IAI.00806-15

WHO. (2022, March 24). *Extensively drug-resistant Shigella sonnei infections—Europe*. https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON364

Williams, P. C. M., & Berkley, J. A. (2018). Guidelines for the treatment of dysentery (shigellosis): A systematic review of the evidence. *Paediatrics and International Child Health*, *38*(sup1), S50–S65. https://doi.org/10.1080/20469047.2017.1409454

Wimmi, S., Balinovic, A., Brianceau, C., Pintor, K., Vielhauer, J., Turkowyd, B., Helbig, C., Fleck, M., Langenfeld, K., Kahnt, J., Glatter, T., Endesfelder, U., & Diepold, A. (2024). Cytosolic sorting platform complexes shuttle type III secretion system effectors to the injectisome in *Yersinia enterocolitica*. *Nature Microbiology*, *9*(1), 185–199. https://doi.org/10.1038/s41564-023-01545-1

Wu, S., Zhang, S., Liu, C.-M., Fernie, A. R., & Yan, S. (2025). Recent Advances in Mass Spectrometry-Based Protein Interactome Studies. *Molecular & Cellular Proteomics*, *24*(1), 100887. https://doi.org/10.1016/j.mcpro.2024.100887

Yang, F. (2005). Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery. *Nucleic Acids Research*, *33*(19), 6445–6458. https://doi.org/10.1093/nar/gki954

Yang, J., Nie, H., Chen, L., Zhang, X., Yang, F., Xu, X., Zhu, Y., Yu, J., & Jin, Q. (2007). Revisiting the Molecular Evolutionary History of *Shigella* spp. *Journal of Molecular Evolution*, *64*(1), 71–79. https://doi.org/10.1007/s00239-006-0052-8

Yoshida, H., Shimada, T., & Ishihama, A. (2018). Coordinated Hibernation of Transcriptional and Translational Apparatus during Growth Transition of *Escherichia coli* to Stationary

Phase. *mSystems*, *3*(5), 10.1128/msystems.00057-18. https://doi.org/10.1128/msystems.00057-18

Yoshida, Y., Miki, T., Ono, S., Haneda, T., Ito, M., & Okada, N. (2014). Functional Characterization of the Type III Secretion ATPase SsaN Encoded by *Salmonella* Pathogenicity Island 2. *PLoS ONE*, *9*(4), e94347. https://doi.org/10.1371/journal.pone.0094347

Yue, S., Xu, P., Cao, Z., & Zhuang, M. (2022). PUP-IT2 as an alternative strategy for PUP-IT proximity labeling. *Frontiers in Molecular Biosciences*, *9*, 1007720. https://doi.org/10.3389/fmolb.2022.1007720

Zhang, B., Ran, L., Wu, M., Li, Z., Jiang, J., Wang, Z., Cheng, S., Fu, J., & Liu, X. (2018). *Shigella flexneri* Regulator SlyA Controls Bacterial Acid Resistance by Directly Activating the Glutamate Decarboxylation System. *Frontiers in Microbiology*, *9*, 2071. https://doi.org/10.3389/fmicb.2018.02071

Appendix

Supplementary Tables for chapter 5

Table 5.S.1. *Shigella* strains used in this study

Bacterial Strain	Description	Antimicrobial resistance	Reference
S. flexneri M90T (WT)	*Shigella flexneri* serotype 5a strain M90T with a spontaneous streptomycin resistant mutation	streptomycin	(Allaoui et al., 1992)
S. flexneri Δ*ipaD*	Derivative of WT M90T with an *aphA3* cassette inserted after codon 130 of *ipaD*	kanamycin	(Ménard et al., 1993)
S. flexneri Δ*ipgC*	Derivative of WT M90T with a zeocin (*zeo*) resistance cassette inserted after codon 42 of *ipgC.*	zeocin	This study
S. flexneri Δ*ipaD* Δ*ipgC*	Derivative of Δ*ipaD* with a Zeocin (*zeo*) resistance cassette inserted after codon 42 of *ipgC.*	kanamycin & zeocin	This study
S. flexneri Δ*mxiE*	Derivative of WT M90T with an *aphA3* cassette inserted after codon 134 of *mxiE*	kanamycin	(Mavris, Page, et al., 2002)
S. flexneri ipaB4	Derivative of WT M90T with a gentamicin resistance cassette inserted after codon 337 of *ipaB*	gentamicin	(Mavris, Page, et al., 2002)
S. flexneri ipaB4 Δ*mxiE*	Derivative of Δ*mxiE* with a gentamicin resistance cassette inserted after codon 337 of *ipaB*	kanamycin & gentamicin	(Mavris, Page, et al., 2002)

Table 5.S.2. Primers used in this study

Primer	Sequence	Description
GMO24	CAGGCCAGGGTGTTGTCC	Check 5' end of *ipgC::zeo*; Reverse primer binding around the 60th codon of *zeo*; used with GMO27
GMO25	GGACAACACCCTGGCCTG	Check 3' end of *ipgC::zeo*; forward primer binding around the 66th codon of *zeo*; used with GMO26
GMO26	ATTTGCAGCTTGGATAGTATTGT	Check 3' end of *ipgC::zeo*; forward primer binding around the 23rd codon of *ipaB*; used with GMO25
GMO27	ACCGAAAATGAAAGCATCTCTA	Check 5' end of *ipgC::zeo*; Forward primer binding around the 12th codon of *ipgC*; used with GMO24
GMO28	GCAATTCCTGATGATATGATGGATGACA TTTATTCATATGTGACTAACTAGGAGGA ATAAATGAAAGCCAAGTTGACCAGT	Allelic exchange; binds 5' end of *zeo* and add *ipgC* homology arm
GMO29	GTGGTGCTTACATTATGCATAATAATTAC TCCTTGATATCTCAGTCCTGCTCCTC	Allelic exchange; binds 3' end of *zeo* and add *ipgC* homology arm
HMIO11	AGAGGAATTCAGAGGGTATTAATAATG AAAGAACAGAAGCTTATTAGCGAAGAG GATCTGTCCGGATTCAAGAACCTGATCT GGCTGA	Cloning; binds 5' end of *bioID2* and adds EcoRI and ribosomal binding site
HMIO14	AGAGTCTAGATTAGTTACAGATCCTCTT CGCTAATAAGCTTCTGTTCTCCGGAGCT TCTTCTCAGGCTGAACTC	Cloning; binds 3' end of *bioID2* and adds XbaI and Myc tag
HMIO247	AGAGGAATTCTAGGAGGTGCTAGCGCT ATGATCCCGCTGCTGAACGC	Cloning; binds 5' end of *miniTurbo* and adds EcoRI and ribosomal binding site
HMIO248	AGAGGGATCCGGCGGTGGCATCCCGCT	Cloning; binds 5' end of *miniTurbo* and

	GCTGAACGC	adds BamHI and 5 AA linker
HMIO249	AGAGGAATTCTAGGAGGTGCTAGCGCTATGAAAGACAATACTGTGCCTCTGAAGC	Cloning; binds 5' end of *turboID* and adds EcoRI and ribosomal binding site
FXO179	AGAGAGAGATCTTTAAATATCACCGAAAATGAAA	Cloning; binds 5' end of *ipgC* and adds BglII
FXO180	AGAGAGGGATCCCTCCTTGATATCCTGAATTGC	Cloning; binds 3' end of *ipgC* and adds BamHI
HMIO250	AGAGGGATCCGGCGGTGGCAAAGACAATACTGTGCCTCTGAAGC	Cloning; binds 5' end of *turboID* and adds BamHI and ribosomal binding site
HMIO251	AGAGTCTAGATTAGTTACAGATCCTCTTCGCTAATAAGCTTCTGTTCTCCGGACTTTTCGGCAGACCGCAGA	Cloning; binds 3'end of *turboID* and miniTurbo and adds XbaI and Myc tag (14 AA addition)
HMIO402	AGAGGGATCCGAAATCAGTGGTCACATCGTACG	Cloning; binds 5'end of *bccP* and adds BamHI
HMIO403	AGAGCTGCAGTTACTCGATGACGACCAGCGG	Cloning; binds 3'end of *bccP* and adds PstI
HMIO404	AGAGGGATCCAAGGACAACACCGTGCCCCT	Cloning; binds 5'end of *bioID* and adds BamHI
HMIO405	AGAGGGTACCTTACTTCTCTGCGCTTCTCAGGGAGAT	Cloning; binds 3'end of *bioID* and adds KpnI
HMIO406	AGAGGGATCCTTCAAGAACCTGATCTGGCTGAAG	Cloning; binds 5'end of bioID2 and adds BamHI
HMIO408	AGAGGGATCCAAAGACAATACTGTGCCTCTGAAG	Cloning; binds 5'end of *turboID* and adds BamHI
HMIO409	AGAGGGTACCTTACTTTTCGGCAGACCGCAG	Cloning; binds 3'end of *turboID* and adds KpnI
HMIO607	AGAGGAATTCTAGGAGGTGCTAGCGCT	Cloning; binds 5' end of *bioID* and adds

	ATGAAGGACAACACCGTGCCCCT	EcoRI and ribosomal binding site
HMIO609	AGAGTCTAGATTAGTTACAGATCCTCTT CGCTAATAAGCTTCTGTTCTCCGGACTT CTCTGCGCTTCTCAGGGAGAT	Cloning; binds 3'end of *bioID* and adds XbaI and Myc tag (14 AA addition)
HMIO650	CGCAGAGGACGGAAGTGGT	Mutagenesis; forward primer to introduce G118R to revert *bioID* to *birA*
HMIO651	TCCTCTGCCAGCCTGCTG	Mutagenesis; reverse primer to introduce G118R to revert *bioID* to *birA*
HMIO694	CTCGGTACCCAGATCCTCTTCGCTAATA AGC	Mutagenesis; forward primer to introduce KpnI-SacI-BamHI at the 3' of the Myc tag in pUC18.1rp *TurboID-myc*
HMIO695	CTCGGATCCTAACTAATCTAGAGTCGAC CTGC	Mutagenesis; reverser primer to introduce KpnI-SacI-BamHI at the 3' of the Myc tag in pUC18.1rp *TurboID-myc*
HMIO698	AGAGAGGGTACCATGGGATCAAGTAAA TATAAAGGTCTAAATACAAGTA	Cloning; binds 5'end of *mxiE* and adds KpnI
HMIO699	AGAGAGGGATCCCTCGAGCCCAATTTT TTCATTTATTTTTTTCA	Cloning; binds 3'end of *mxiE* and adds BamHI
HMIO701	AGAGAGGAATTCATGGGATCAAGTAAA TATAAAGGTCTAAATACAAGTA	Cloning; binds 5'end of *mxiE* and adds EcoRI
HMIO702	GGCAAGATCGTCCGTTGTCATAATCGAC	PCR; forward primer for the amplification of the biotin operator from str. M90T
HMIO703	GCCATGGGGCTTCTCCAAAACGTG	PCR; reverse primer for the amplification of the biotin operator from str. M90T
HMIO734	AGAGAGGGTACCTCTTTAAATATCACCG AAAATGAAAGCATCTCTACTG	Cloning; binds 5'end of *ipgC* and adds KpnI
HMIO735	AGAGAGGAGCTCTTACTCCTTGATATCC TGAATTGCGTCCAAG	Cloning; binds 3'end of *ipgC* and adds SacI

HMIO390	GCGGAGTGGAACGTGAGCTACGG	Mutagenesis; forward primer to introduce K19A to *bioID2*
HMIO391	CAGTCTCTCCTGGGTGCTGTCCAC	Mutagenesis; reverse primer to introduce K19A to *bioID2*
HMIO89	GTTCTCGAACTCCTTGGGG	Mutagenesis; forward primer to introduce L66A to *bioID2*
HMIO90	GGACCTATATTGATAGCACAGAAAT	Mutagenesis; reverse primer to introduce L66A to *bioID2*
HMIO69	GCGCAGCTGCCCCTGGTG	Mutagenesis; forward primer to introduce L67A to *bioID2*
HMIO70	CAGGTTCTCGAACTC	Mutagenesis; reverse primer to introduce L67A to *bioID2*
HMIO392	GCGAAGGACAAGCTGATCGTGGG	Mutagenesis; forward primer to introduce S112A to *bioID2*
HMIO393	CAGCTCGCACAGCACGCCG	Mutagenesis; reverse primer to introduce S112A to *bioID2*
HMIO394	GCGAAGCTGATCGTGGGCATC	Mutagenesis; forward primer to introduce D114A to *bioID2*
HMIO396	CTTGCTCAGCTCGCACAGCAC	Mutagenesis; reverse primer to introduce D114A to *bioID2*
HMIO75	GCGCTGATCGTGGGCATCG	Mutagenesis; forward primer to introduce K115A to *bioID2*
HMIO76	GTCCTTGCTCAGCTCGCA	Mutagenesis; reverse primer to introduce K115A to *bioID2*
HMIO397	GCGTTCAGCCTGAGAAGAAGC	Mutagenesis; forward primer to introduce E227A to *bioID2*
HMIO398	GCCGCTCAGGATCTCCTTGAT	Mutagenesis; reverse primer to introduce E227A to *bioID2*

HMIO423	GCCGCGCAGCTGCCCCTGGTGCTG	Mutagenesis; forward primer to introduce L66A L67A to *bioID2*
HMIO424	GTTCTCGAACTCCTTGGGGTTCAG	Mutagenesis; reverse primer to introduce L66A L67A to *bioID2*
HMIO637	GGAGGCAGTAAGGACAACACCGTGCCC	Forward primer to extend the linker between *ipgC* and *bioID* from 5 to 10 AA
HMIO639	GGAGGCAGTAAAGACAATACTGTGCCTCTGAAGC	Forward primer to extend the linker between *ipgC* and *turboID* from 5 to 10 AA
HMIO638	ACCTCCGCCACCGCCGGATCC	Reverse primer to extend the linker between *ipgC* and *bioID* or *turboID* from 5 to 10 AA
HMIO407	AGAGGGTACCTTAGCTTCTTCTCAGGCTGAACTCG	Cloning; binds 3'end of *bioID2* and adds KpnI

Table 5.S.3. Main plasmids used in this study

Plasmid ID	Plasmid name	Promoter	Reference	Addgene ID
pSU2.1	pSU2.1	Constitutive *lac* promoter	(Silué & Campbell-Valois, 2022)	N/A
pNHA1	pSU2.1 *bioID-myc*	Constitutive *lac* promoter	This study	N/A
pNHA2	pSU2.1 *bioID2-myc*	Constitutive *lac* promoter	This study	N/A
pNHA3	pSU2.1 t*urboID-myc*	Constitutive *lac* promoter	This study	N/A
pNHA4	pSU2.1 *miniTurbo-myc*	Constitutive *lac* promoter	This study	N/A
pNHA5	pQE-80L *turboID*	Inducible T5/*lac* promoter	This study	223647
pNHA6	pQE-80L *bioID*	Inducible T5/*lac* promoter	This study	223648
pNHA7	pQE-80L *birA*	Inducible T5/*lac* promoter	This study	223649
pNHA8	pQE-80L *MBP-bccP*	Inducible T5/*lac* promoter	This study	223650
pNHA9	pUC18.1rp	Constitutive *rpsM* promoter	This study	223642
pNHA10	pUC18.1rp t*urboID-myc*	Constitutive *rpsM* promoter	This study	223554
pNHA11	pUC18.1rp *ipgC-turboID-myc*	Constitutive *rpsM* promoter	This study	223643

pNHA12	pUC18.1rp t*urboID-myc-ipgC*	Constitutive *rpsM* promoter	This study	223644
pNHA13	pUC18.1rp *mxiE-turboID-myc*	Constitutive *rpsM* promoter	This study	223645
pNHA14	pUC18.1rp *turboID-myc-mxiE*	Constitutive *rpsM* promoter	This study	223646
pNHA15	pSU2.1 *ipgC-myc*	Constitutive *lac* promoter	This study	N/A
pNHA16	pSU2.1 *ipgC-bioID-myc*	Constitutive *lac* promoter	This study	N/A
pNHA17	pSU2.1 *ipgC-bioID2-myc*	Constitutive *lac* promoter	This study	N/A
pNHA18	pSU2.1 *ipgC-turboID-myc*	Constitutive *lac* promoter	This study	N/A
pNHA19	pSU2.1 *ipgC*-L-*bioID-myc*	Constitutive *lac* promoter	This study	N/A
pNHA20	pSU2.1 *ipgC*-L-*turboID-myc*	Constitutive *lac* promoter	This study	N/A
pNHA21	pQE-80L *bioID2*	Inducible T5/*lac* promoter	This study	N/A
pUC18.1	pUC18.1	Constitutive *lac* promoter	(Silué & Campbell-Valois, 2022)	N/A
pNHA22	pUC18.1 *ipgC-myc*	Constitutive *lac* promoter	This study	N/A
pNHA23	pUC18.1rp *ipgC-myc*	Constitutive *rpsM* promoter	This study	N/A