Resolving Cytosolic Interactions and Injectisome Binding Dynamics of the *Yersinia enterocolitica* Sorting Platform and ATPase via Single-Molecule Tracking

Joshua R Prindle

B.S. in Biochemistry, Clarke University, 2018

B.A. in English, Clarke University, 2018

A Dissertation presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Chemistry

University of Virginia

August 2023

Andreas Gahlmann

Linda Columbus

David Cafiso

Marcos Pires

Hervé Agaisse

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. Andreas Gahlmann. Coming into graduate school, my expertise in fluorescence microscopy and computer programming was non-existent. However, simply observing how Andreas interacted with the remainder of his research group showed me, time and time again, that he was truly passionate about teaching his students. He was patient as I navigated the training process, always making sure to take time out of his day to help in any way I needed. As the years progressed, so too did our professional relationship. I will never forget the countless hours spent talking shop in his office, arguing over what should and shouldn't go into our manuscripts. It was these moments that made me feel like I belonged – that I was a scientist. Ultimately, Andreas saw potential in me, as he does in all of his students. I will carry the lessons he taught with me in all of my future endeavors, and I wish him the best of luck in the future research of his lab.

I would also like to thank the students who helped me in the Gahlmann Lab, all of whom I now consider my friend. When I first joined, Dr. Julian Rocha was in his final year of graduate school. Julian trained me on the microscope and on the data processing, always making sure to answer even the simplest of questions. I will never forget our countless lunch dates, eating delicious food and talking, not as colleagues, but as friends. In my second and third year, Dr. Alecia Achimovich and I bonded over our affection for cats (we basically had a competition for who could collect the most). Alecia also provided me the necessary help for all things molecular biology, which was instrumental in my research. Finally, I must thank Dr. Yibo Wang, who joined the Gahlmann Lab alongside me, and Olivia de Cuba. No matter what, Yibo always showed patience and kindness. I was inspired

by his ability to keep a level head, even in the most stressful of situations, and his willingness to put his work to the side to lend a helping hand. Similarly, I was inspired by Olivia's ability to navigate her research like a seasoned professional. I will never forget our countless walks together, the numerous cups of coffee, and the friendship that we developed. The expertise she brought to the lab helped me tremendously, and I thank her for that. I wish her the best of luck as she comes to the end of her graduate school experience and carries on with the T3SS project.

Of course, I also need to thank my mother and father. My father, who is the hardest-working man I know, helped me throughout the entirety of my academic journey. He always knew what I was learning, why I was learning it, and what my next steps were. Even as he was going through cancer treatment in my fourth year of graduate school, he would inquire about what my next experiment would be. He kept me focused, and more importantly, allowed me to talk about and follow my passion for science. My mother, who raised me since I was a baby, helped me stay calm. Since grade school, she referred to me as "her little scholar." She encouraged me to find enjoyment in life, which helped me stay relaxed even in the most stressful of times.

Finally, I need to thank my significant other, Sadie. We began dating in our first year of graduate school and have been inseparable ever since. She supported me every step of the way, even while managing her own research. She has the biggest heart and, through thick and thin, loves me endlessly. Making it to this point would certainly not have been possible without her. She is my best friend, and I cannot wait to see what the future holds for us.

ABSTRACT

The membrane-embedded injectisome, the structural component of the virulence-associated type III secretion system (T3SS), is used by gram-negative bacterial pathogens to inject species-specific effector proteins into eukaryotic host cells. The cytosolic injectisome proteins, also known as the sorting platform and ATPase, are required for export of effectors and display both stationary, injectisome-bound populations and freely diffusing cytosolic populations. In *Yersinia enterocolitica,* evidence supports an effector protein shuttling model, in which effector proteins are shuttled from the cytosol to the injectisome prior to secretion through the hollow needle complex. However, how the cytosolic injectisome proteins interact with each other in the cytosol and associate with membrane-embedded injectisomes remains unclear. Additionally, it is difficult to reconcile the observed exchange rate of *Ye*SctQ, the primary component of the sorting platform, with reported rates of substrate secretion. To determine the subcomplexes formed by the sorting platform and ATPase in the cytosol of live *Y. enterocolitica*, I utilized 3D single-molecule localization and tracking microscopy. Specifically, I developed a novel data analysis pipeline that relies on the diffusion coefficient spectrum, which provided a straightforward approach towards interpreting complex intracellular diffusion data. I also developed a data processing pipeline for 2D single-molecule bound-time data, which allowed for quantitative analysis of an individual protein's binding kinetics with stationary injectisomes. Results support a model in which distinct complexes readily form among the sorting platform and ATPase in the cytosol of live *Y. enterocolitica,* where the propensity for complex formation changes in the presence or absence of injectisomes. Additionally, bound-time analysis of *Ye*SctQ supports the effector protein shuttling mechanism. Our data

indicates that *Ye*SctQ can deliver effector proteins to the injectisome at a maximum secretion rate of approximately one effector protein every 0.6 seconds. Further research is needed to support these findings, which has important implications for how the sorting platform and ATPase functionally regulate secretion.

TABLE OF CONTENTS

Chapter 1
INTRODUCTION

1.1 Virulence-associated T3SS in gastrointestinal pathogens

1.1.1 Overview of the injectisome

Bacterial secretion systems are a class of membrane-embedded, macromolecular complexes that are responsible for secreting proteins from the bacterial cytosol into the extracellular space, into other bacteria, or into plant and eukaryotic cells. To date, nine different secretion systems have been identified. These secretion systems play crucial roles in bacterial survival, virulence, cell-to-cell communication, and interactions with host organisms, contributing to processes such as pathogenesis, cell motility, biofilm formation, and nutrient acquisition (1,2).

The injectisome, otherwise known as the virulence-associated type III secretion system (*v*T3SS), is expressed by many bacterial pathogens, including *Escherichia coli*, *Salmonella*, *Pseudomonas*, *Shigella*, and *Yersinia*, which are responsible for widespread human disease both historically and currently. The injectisome connects the bacterial cytosol with the eukaryotic cytosol, essentially acting as a molecular syringe, which allows pathogens to inject effector proteins into host cells (3,4), thereby manipulating host cellular processes and promoting infection. In particular, secretion of virulent effectors into host cells enables pathogens to evade host defenses, modulate immune responses, and establish a conducive environment for their survival and replication within the host organism (5). For example, in *Yersinia enterocolitica*, the injectisome secretes the effector protein YopE, which functions as a GTPase-activating protein, disrupting the actin cytoskeleton and impairing the host cell's ability to phagocytose the bacteria (6). While the proteins of the injectisome have been (structurally) characterized

and share homology among pathogens, the effector proteins exhibit species-specificity, varying based on the specific site and mode of pathogenesis within the host organism.

With the rising threat of antibiotic resistance (7), the injectisome, a target specific to pathogenic bacteria, may serve as an attractive drug target for combatting infection. Modern-day antibiotics are generally either bacteriostatic or bactericidal, creating selective pressure on bacteria to mutate as a survival mechanism against these drugs. Targeting the injectisome may be beneficial, because disrupting this system could not only combat antibiotic-resistant bacterial strains but also potentially reduce the selective pressure imposed by traditional antibiotics (8). Additionally, re-programming of the injectisome for the delivery of biomolecules is attractive for a variety of biomedical applications (9).

1.1.2 Structure of the injectisome

In its fully-assembled state, the injectisome is composed of multiple subcomplexes and spans the inner- and outer-bacterial cell membranes and the host cell membrane. These subcomplexes are known as the needle complex and translocon pore, the inner- and outer-membrane rings, the export apparatus, and the cytosolic complex (**Fig. 1.1a**). Specifically, the needle complex protrudes from the bacterial cell wall (approximately 60 nm) into the extracellular space and is composed of the structural protein, SctF (10). The needle complex is capped with the needle tip protein, SctA. SctB and SctE form a translocon pore in the host cell membrane (11). At the base of the needle complex are SctC, SctD, and SctJ, which form the inner- and outer-membrane rings that stably anchor the injectisome in the bacterial cell wall. The export apparatus is embedded in the inner-membrane and is composed of SctR, SctS, SctT, SctU, and SctV. Finally, at the base of the injectisome is

the cytosolic complex, which is composed of SctK, SctQ, and SctL, and the ATPase, SctN (12–14). The cytosolic complex can be further classified as a combination of the sorting platform, which includes SctK, SctQ, and SctL, and the ATPase, composed of hexameric SctN.

1.1.3 Assembly of the injectisome and effector protein secretion

Injectisome assembly follows a two-step process: first, the establishment of the primary secretion-competent machinery, and second, the assembly of the needle filament, tip, and translocon pore, which is dependent on type III secretion. The principle secretion-competent machinery consists of the export apparatus and inner- and outer-membranes rings. Two assembly models have been proposed (3,12,15). In the inside-out model of assembly, the export apparatus proteins SctR, SctS, and SctT form a hexameric structure in the inner-membrane, followed by the addition of SctU and SctV (16). The fully-assembled export apparatus then serves as a nucleation point for the assembly of the inner-membrane ring. Specifically, SctJ and SctD collectively organize into two 24-membered rings, forming a nested configuration, where SctJ resides within the interior of the ring surrounding the export apparatus. The outer-membrane ring, formed independently by SctC, then combines with the formed export apparatus/inner-membrane ring structure, permitting binding of the cytosolic complex (17,18). An inside-out assembly model is favored in *Salmonella* and *E. coli*, while *Y. enterocolitica* favors an outside-in assembly model. The key difference in the outside-in model is that the outer-membrane ring is formed prior to the incorporation of the inner-membrane ring and export apparatus (19).

Formation of the secretion-competent machinery permits type III secretion-dependent assembly of the needle filament, tip, and translocon pore, which occurs in a hierarchal manner. After formation of the inner rod complex, composed of SctI, SctF is secreted through the injectisome and forms the needle by polymerizing into a helical structure, creating a channel that protrudes from the bacterial surface (20–23). Intermittent secretion of the needle length regulator protein, SctP, modulates the length of the needle by controlling the polymerization of needle subunits (24,25). Upon reaching the desired length, the needle is capped with the needle tip protein, SctA (26), and the injectisome switches from early- to middle-secretion substrates, a process that is in part due to the export apparatus protein, SctU (27–32). The middle-secretion substrates or translocases, SctB and SctE, are secreted through the needle, where they interact with and form the translocon pore in the host cell membrane (14,33,34). With a conduit between the bacterial and eukaryotic cell cytosol established, the injectisome finally switches its substrate selectivity to the late-secretion substrates, otherwise known as virulent effectors.

1.1.4 Binding properties and molecular organization of the sorting platform

The sorting platform is not embedded within the inner-membrane but instead transiently associates with SctD through its interaction with SctK. The transient nature of this complex has made its in vitro purification and characterization challenging. By capturing multiple two-dimensional images of a rapidly frozen, thin specimen from different angles (35), visualization of the sorting platform through cryogenic electron tomography (cryo-ET) with fully-assembled injectisomes has been possible (36).

Recent studies suggest distinct binding properties of the sorting platform in situ, which are contingent upon the bacterial species of origin. Subtomogram averages of the *Shigella flexneri* and *Salmonella* Typhimurium sorting platform complexes suggest that they are structurally similar. *S.* Typhimurium injectisomes show binding of the sorting platform leads to structural changes in the inner membrane ring, composed of 24 *Se*SctD proteins, to accommodate 6-fold symmetry. Specifically, helices from 4 *Se*SctD proteins arrange to associate with an individual *Se*SctK, with 6 *Se*SctK proteins binding to each injectisome (37). Axial scanning of the *S.* Typhimurium and *S. flexneri* sorting platform subtomogram averages suggest 6 distinct puncta consistent with the presence of separate complexes, or "pods," where each sorting platform pod follows a SctK-SctQ-SctL sequence of interactions, with the hexameric ATPase at the base of the complex (37,38) (**Fig. 1.1b**). Unlike the observation of distinct puncta observed for *Se*SctD in *S.* Typhimurium, however, *Sf*SctD in *S. flexneri* and *Ye*SctD in *Y. enterocolitica* form a continuous ring-like structure in the inner-membrane (39,40). Additionally, subtomogram averages of the *Y. enterocolitica* injectisome show continuous, ring-like densities in the region of *Ye*SctQ (40) (**Fig. 1.1c**). However, it is worth noting that this observation was limited by the resolution (~10 Å) of cryo-ET, and thus, the existence of distinct pods cannot be ruled out. Additionally, the choice of an initial reference structure for alignment during the iterative reconstruction process can introduce bias. A poor choice may lead to convergence towards local minima instead of the global minimum, and the final reconstructed structure may reflect the features present in the poorly chosen initial reference rather than the true structure of the specimen (41,42).

Despite these potentials for bias from cryo-ET studies, in vitro and in vivo evidence support the above findings and suggest that the stoichiometric combination of proteins that compose each pod complex differs among gastrointestinal pathogens. Photocrosslinking followed by purification and characterization of sorting platform subcomplexes in *S.* Typhimurium revealed that only an individual *Se*SctQ is required for binding with *Se*SctK and *Se*SctL. The presence of *Se*SctL, however, is required for the interaction between *Se*SctQ and *Se*SctK, suggesting that a conformational shift of *Se*SctQ induced by *Se*SctL permits binding with *Se*SctK (43). Two-hybrid analyses coupled with biophysical characterization of sorting platform subcomplexes in *S. flexneri* suggests the formation of distinct *Sf*SctK:*Sf*SctQ and *Sf*SctQ:*Sf*SctL complexes, with *Sf*SctQ adopting different conformational states for each (44). In *Y. enterocolitica*, live-cell fluorescence microscopy data of eGFP-tagged *Ye*SctQ suggested approximately 22-24 *Ye*SctQ per injectisome (45,46), yet the precise stoichiometry of *Ye*SctQ-containing complexes remains unknown.

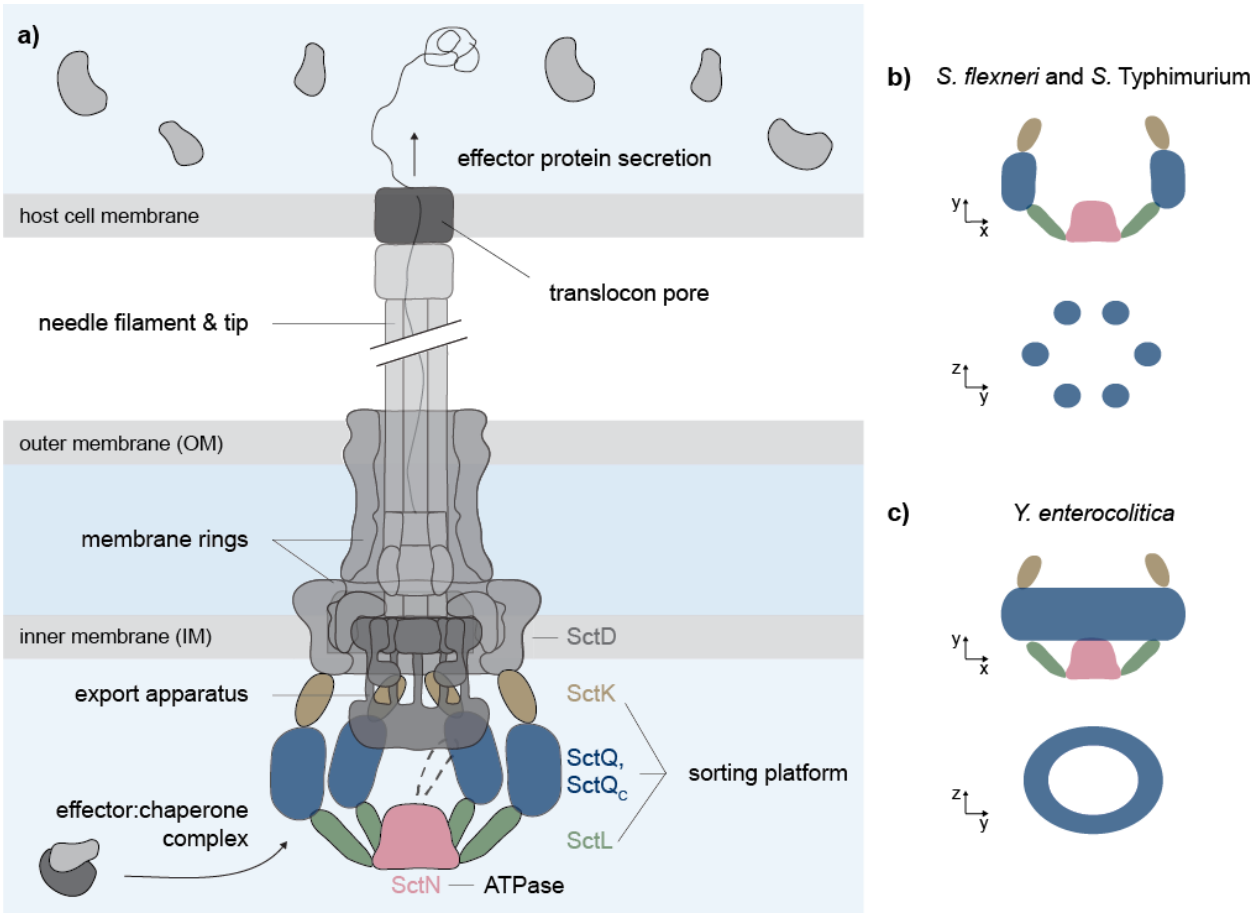

Figure 1.1: Graphical representation of the injectisome. a) The injectisome is composed of multiple subcomplexes and spans three membranes. Effectors are dissociated from their cognate chaperons prior to secretion through the hollow needle. b) Current model of the sorting platform and ATPase in *S.* Typhimurium and *S. flexneri*. Axial scanning of their respective cryo-ET subtomogram averages suggests 6 distinct puncta. c) Current model of the sorting platform and ATPase in Y. *enterocolitica*. Axial scanning of cryo-ET subtomogram averages suggests excess density and a ring-like structure in the region of *Ye*SctQ.

1.1.5 Binding properties and functional regulation of the ATPase, SctN

The functional relevance of the ATPase, SctN, as well as its molecular organization with fully-assembled injectisomes, is generally known and ubiquitous among bacterial species. For each injectisome, the homo-hexameric ATPase is cradled by 6 spoke-like dimers of SctL. SctN hexamerization activates its ATPase activity and thus establishes a proton motive force (PMF) across the bacterial inner membrane, where the energy derived from ATP hydrolysis allows for the secretion of substrates through the needle complex (47,48).

Despite these similarities, however, the binding properties of cytosolic (i.e., injectisome-unbound) SctN differs among bacterial species. In *S.* Typhimurium, it was shown that *Se*SctN required the presence of *Se*SctK, *Se*SctQ, and *Se*SctD in order to interact with *Se*SctL. Removal of any of these proteins abolished high-molecular weight species of *Se*SctN, supporting the notion that a fully assembled sorting platform is required for *Se*SctN recruitment and hexamerization (43). Conversely, in *S. flexneri* and *Y. enterocolitica*, SctN readily binds SctL. In vitro characterization of isolated *Sf*SctN:*Sf*SctL complexes in *S. flexneri* revealed that dimers of *Sf*SctL differentially regulate *Sf*SctN ATPase activity depending on its oligomeric state. Specifically, while *Sf*SctL increases the ATPase activity of monomeric *Sf*SctN, it decreases its activity for higher-order *Sf*SctN oligomers, suggesting functionally distinct roles for these freely diffusing complexes (44,49). Similarly, isolated *Ye*SctN:*Ye*SctL complexes in *Y. enterocolitica* revealed that increasing concentrations of *Ye*SctL correlates with decreasing ATPase activity of freely diffusing *Ye*SctN (50). SctL thus serves as an ATPase regulator protein in *S. flexneri* and

Y. enterocolitica, while such regulatory properties are seemingly not required in *S.* Typhimurium.

1.1.6 Functional roles of the sorting platform and ATPase

A key feature of injectisomes for all the aforementioned bacterial species is their ability to selectively secrete substrates in a defined order during type III secretion-dependent assembly and effector protein secretion – this feature is due, in part, to and the reason for the term "sorting platform." Specifically, a 2011 study in *S.* Typhimurium revealed that, depending on the presence or absence of specific substrates necessary for injectisome assembly and effector protein secretion, a mDa-sized complex composed of *Se*SctK, *Se*SctQ, and *Se*SctL was sequentially bound to (or "loaded" with) different secretion substrates complexed with their cognate chaperones (51). In particular, deletion of the translocases, or middle-secretion substrates, resulted in binding of chaperone:effector complexes, or late-secretion substrates. Upon deletion of the needle length regulator protein, which is an early-secretion substrate, neither the translocases nor the effector proteins associate to this complex. It was therefore concluded that *Se*SctK, *Se*SctQ, and *Se*SctL bind secretion substrates in a hierarchal manner. Despite these findings, they have not been replicated in other gastrointestinal pathogens. Additionally, the aforementioned differences in the binding properties and corresponding structures of the sorting platform and ATPase in different gastrointestinal pathogens leaves many questions unanswered.

Similar to the FliM/FliN pair found in the flagellar-associated T3SS, SctQ is expressed alongside its alternatively expressed C-terminal fragment, $SctQ_C$. In *Y.*

enterocolitica, *Ye*SctQ$_C$ is a core structural component of the injectisome and is required for secretion of effectors (45,52). The crystal structure of *Ye*SctQ$_C$ revealed that it forms a homodimer that associates with full-length *Ye*SctQ (52). In *S. flexneri*, *Sf*SctQ$_C$ is also required for secretion and forms a 2*Sf*SctQ$_C$:*Sf*SctQ heterotrimer (44,53). In *S.* Typhimurium, however, full-length *Se*SctQ alone (i.e., without the presence of *Se*SctQ$_C$) is stable and permits secretion of effectors, albeit with reduced efficiency (54).

In order to help clarify the functional roles of the sorting platform and ATPase, live-cell microscopy in combination with in vitro interaction data in *Y. enterocolitica* has been leveraged. Specifically, secretion of effector proteins in *Y. enterocolitica* can be turned OFF or ON in the presence or absence of extracellular calcium, respectively. Fluorescence recovery after photobleaching (FRAP) experiments revealed that *Ye*SctQ dynamically binds and unbinds with injectisomes under secretion-ON conditions. Interestingly, under secretion-OFF conditions, the *Ye*SctQ exchange rate is reduced by approximately a factor of two (45). Additional studies showed that *Ye*SctQ, *Ye*SctK, *Ye*SctL, and *Ye*SctN have substantial cytosolic populations in addition to injectisome-bound populations (45,46). Under secretion-OFF conditions, the interaction strengths among these proteins increases and the diffusion rates of the corresponding complexes decreases (46). Collectively, these findings suggest that the binding properties among cytosolic and injectisome-bound complexes containing *Ye*SctK, *Ye*SctQ, *Ye*SctL, and *Ye*SctN are sensitive to the secretion-state of the injectisome. This feature allows for faster exchange of cytosolic subcomplexes with injectisomes during secretion, which is potentially correlated with the delivery of chaperone:effector secretion substrates by these subcomplexes (**Fig. 1.2**). These findings may help explain the reason for the excess *Ye*SctQ

and *Ye*SctQ$_C$ protein density (see section 1.1.4), as it may be a required feature to ensure rapid delivery of secretion substrates.

The *Y. enterocolitica* injectisome additionally switches between secretion-ON and -OFF states under neutral and low pH conditions, respectively (55). Gastrointestinal pathogens experience low pH environments during infection, which negatively impacts cell viability. It was found that, under these conditions, the *Y. enterocolitica* sorting platform and ATPase dissociate from injectisomes, halting secretion. Secretion resumes shortly after restoration of neutral pH conditions. This reversible mechanism likely regulates the injectisome's energy consumption, where consumption is minimal under environmentally harsh conditions.

Despite these findings, it remains to be determined precisely how the cytosolic population of proteins that compose the sorting platform and ATPase in Y. enterocolitica 1) associate with one another in the cytosol and 2) contribute to the functionality and regulation of secretion. Specifically, determining the cytosolic subcomplexes formed in *Y. enterocolitica* will begin to uncover the steps that contribute to the assembly of the sorting platform and ATPase. Results shown in *S.* Typhimurium suggest that *Se*SctK first binds with injectisomes, followed by a *Se*SctQ:*Se*SctL complex and, finally, the ATPase, *Se*SctN (43). Conversely, live-cell microscopy results in *Y. enterocolitica* revealed that all the proteins composing the sorting platform and ATPase are required for injectisome binding (45,46,52). This suggests that these proteins first associate with one another in the cytosol prior to their association with injectisomes. Establishing the cytosolic subcomplexes that compose the sorting platform and ATPase is ultimately necessary in order to determine their functional roles.

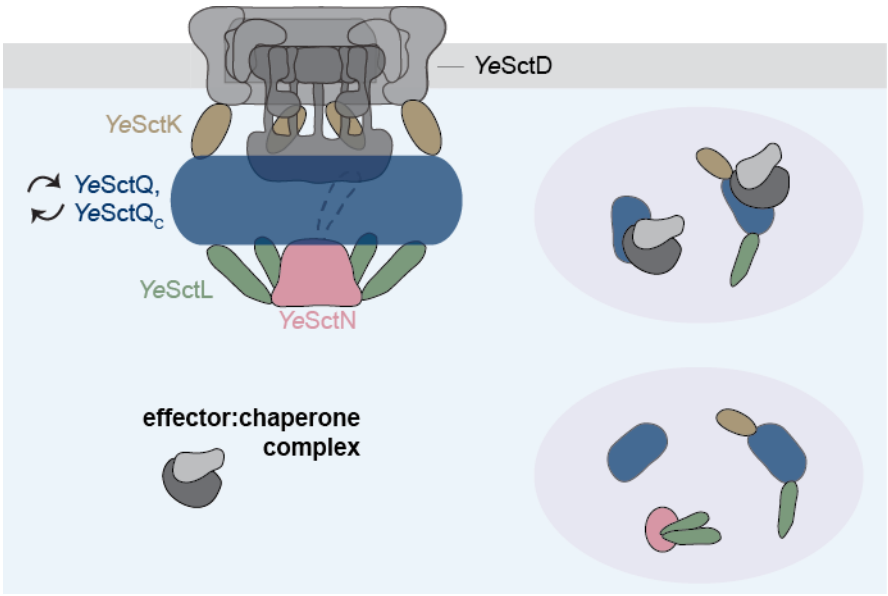

Figure 1.2: Potential mechanism of substrate recruitment to the *Y. enterocolitica* injectisome. Live-cell fluorescence microscopy data reveals the exchange rate of *Ye*SctQ with injectisomes increases upon activation of effector protein secretion. In vitro interaction data reveals interaction between *Ye*SctQ with chaperone:effector complexes. Additional in vitro interaction data reveals a network of interactions among *Ye*SctK, *Ye*SctQ, *Ye*SctL, and/or *Ye*SctN.

1.2 Dissertation Outline

Live-cell fluorescence imaging has been pivotal in determining functional roles of the proteins that compose the sorting platform and ATPase. The work presented in this dissertation leverages advancements in super-resolution fluorescence imaging to uncover the roles of these proteins. In Chapter 2, single-molecule localization microscopy (SMLM) will be introduced. Chapter 3 focuses on the previously established (56) and newly updated data analysis pipeline utilized to process single-molecule localizations and quantify the

diffusive states of freely-diffusing proteins. Single-molecule tracking analysis that reveals cytosolic subcomplexes formed by *Ye*SctQ, *Ye*SctL, and *Ye*SctN in *Y. enterocolitica* will be presented in Chapter 4. Long-exposure, 2D single-molecule tracking data reveals the binding times of *Ye*SctK, *Ye*SctQ, *Ye*SctL, and *Ye*SctN with injectisomes, which will be discussed in Chapter 5. Finally, Chapter 6 will focus on the significance of these findings as well as future directions in aimed at uncovering the complex nature of the sorting platform and ATPase.

Chapter 2
SUPER-RESOLUTION FLUORESCENCE IMAGING

**** Excerpts from this chapter are taken from Prindle, J. R., de Cuba, O.I.C., & Gahlmann, A. *Single-molecule tracking to determine the abundances and stoichiometries of freely-diffusing protein complexes in living cells: Past applications and future prospects* (J. Chem. Phys., 2023, Vol. 159, No. 7)**

2.1 Super-resolution fluorescence imaging

Fluorescence microscopy allows for the visualization of biomolecules within their native biological context. The inherent resolution constraints of conventional diffraction-limited fluorescence imaging restrict the amount of detail that can be captured from a fluorescence image. Specifically, the diffraction limit represents the distance in which two fluorescent emitters can no longer be distinguished from one another. This is due to the fact that light waves diffract as they pass through a lens, causing the signal from individual molecules, otherwise known as the point-spread-function (PSF), to spread out and overlap. The diffraction limit, described by the Airy pattern, therefore sets a fundamental constraint on the achievable resolution in optical microscopy. The diffraction-limit of light was first described by Ernest Abbe with the following equation,

$$d_{x,y} = \frac{\lambda}{2 * n * sin\theta} \tag{2.1}$$

where d is the diameter of the PSF, λ is the wavelength of the emitted light, n is the refractive index of the medium the lens is in, and θ is the maximum half-angle of the cone of light entering the lens. The numerical aperture (NA) is equal to $n * sin\theta$, simplifying the Equation 2.1 to:

$$d_{x,y} = \frac{\lambda}{2 * NA} \tag{2.2}$$

In practical terms, the diffraction-limit defines the smallest resolvable details in the specimen being imaged, which is approximately 300 nm. Therefore, it is not possible extract desirable information from a biological structure of interest with features smaller than this limit, as the resulting image would appear blurred.

In order to surpass the resolution constraints of optical microscopy, various 'super-resolution' techniques were devised, culminating in the award of the Nobel Prize in Chemistry in 2014. Super-resolution fluorescence microscopy can be broadly classified into two main approaches, each relying on distinct mechanisms: point-spread-function (PSF)-based localization and modulation-enhanced localization. For PSF-based localization methods, such as PALM (57,58) and STORM (59), the diffraction-limited image of a point source (i.e. the PSF) is analyzed to estimate the emitter position (56,60,61). In modulation-enhanced localization methods (62), such as MINFLUX (63–66), pulsed interleaved MINFLUX (p-MINFLUX) (67), single-molecule confocal laser tracking combined with fluorescence correlation spectroscopy (SMCT–FCS) (68), Orbital scanning (69–71), 3D-DyPLoT(72,73), 3D-SMART (74), and TSUNAMI (75), spatially structured illumination patterns result in fluorescence intensity modulations that are analyzed to localize individual fluorophores. The temporal resolution of modulation-enhanced localization can reach 0.1 ms, which is about two orders of magnitude higher than what is typically achieved in PSF-based localization methods. The increased time resolution of modulation-enhanced localization methods derives from a more efficient use of the information delivered by each photon (63). However, in modulation-enhanced methods, individual fluorescent molecules are imaged sequentially in time, whereas PSF-based localization methods permits imaging of many fluorescent molecules concurrently.

2.1.1 PSF engineering

The conventional 2D PSF provides the lateral *xy* positions of any fluorescent emitter. Scanning above and below the focal plane for the conventional 2D PSF results in

a spatially-symmetric Airy pattern, and therefore, any information about the axial *z* position of the fluorescent emitter is lost. PSF engineering is achieved by placing a phase mask in the back focal plane of the microscope, where this phase mask introduces aberrations into the PSF signal. These aberrations therefore encode 3D information about the 2D PSF shape (76).

While several approaches towards engineering the 2D PSF have emerged, only a subset of these approaches will be mentioned here. The earliest approach leveraged the astigmatism imaging method, which incorporates a cylindrical lens into the back focal plane of the imaging pathway. The cylindrical lens generates two marginally distinct focal planes along the *x* and *y* directions, and as a result, the ellipticity and orientation of the fluorophores PSF changes as the *z*-position changes (77). Shortly after the introduction of the astigmatic PSF, the double-helix PSF (DHPSF) was introduced. Briefly, a phase mask placed in the back focal plane of the microscope splits a single 2D PSF into two distinct lobes. As the *z*-position of the emitter changes with respect to the focal plane, these lobes rotate around one another. Given that the angle of the lobes relative to one another are asymmetric above and below the focal plane, the *z*-position of the molecule can be extracted (78). Finally, the DHPSF was improved upon through the introduction of the tetrapod PSF, which exhibits a larger axial range compared to the DHPSF (79). All of the experiments presented in this dissertation were performed in bacteria. Given that the axial range of the DHPSF is approximately 1.5 μm, while bacteria are typically 0.2 – 1 μm in width, the DHPSF is suitable for achieving 3D super-resolution capabilities throughout the entire bacterial cell.

2.1.2 Labelling target molecules with fluorescent emitters

The work described in this dissertation utilizes PSF-based localization, which relies on controlling the fluorescent emitter concentration in space and time. In PSF-based experiments, a majority of the fluorescent emitters are in a fluorescence OFF state. Therefore, only a small subset of fluorophores are in a fluorescence ON state at any given time. Sparse fluorescence signal avoids overlapping signals from nearby molecules and allows for the localization of individual emitters. Therefore, this technique is known as single-molecule localization microscopy (SMLM). The number of photons collected from an individual fluorescent emitter is correlated with the precision of the localization, σ,

$$\sigma \approx \frac{s}{\sqrt{N}} \tag{2.3}$$

where s is the standard deviation of the PSF and N is the number of photons collected above the background.

For any super-resolution fluorescence imaging experiment, the choice of fluorophore must be appropriate for the desired application. Given that fluorescence-labeling has been covered extensively (80,81) and is not the focus of this dissertation, it will only be briefly mentioned here. Generally, fluorescent labels can be broadly placed into two categories: fluorescent proteins or fluorescent dyes. Fluorescent proteins are genetically encodable, allowing for 100% labelling efficiency of the target molecule. Fluorescent protein labelling can be achieved *in trans* through the use of expression plasmids or through genetic incorporation into chromosomal DNA. However, fluorescent proteins are large (typically ~ 30 kDa) and are approximately 10 times less bright than fluorescent dyes. Given that the number of collected photons correlates with the precision

of the localization, fluorescent dyes therefore provide more precise localizations. Despite this advantage, fluorescent dyes have their own drawbacks. Specifically, nonspecific labeling of fluorescent dyes results in fluorescence background. For any single-molecule imaging experiment, the choice of fluorophore must ultimately be appropriate for the desired application.

2.2 Instrumentation

In order to perform 3D super-resolution fluorescence imaging, a custom-built inverted fluorescence microscope was previously constructed (**Fig 2.1**). Utilizing a custom-built microscope is advantageous, as the instrument is modular, allowing for the addition or switching out of optics necessary for the needs of the fluorescence imaging experiment. The microscope consists of excitation and emission pathways for the collection of fluorescence signal, as well as a phase contrast pathway for imaging of bacterial cell shapes. While this microscope has been described in detail previously, the following sections will walk the reader through the details of each pathway.

2.2.1 Fluorescence imaging

Fluorescence microscopy necessitates the utilization of a precise wavelength of light to effectively excite fluorophores within a specimen. This wavelength corresponds to the energy required to excite electrons in the fluorescent emitter to higher energy states, enabling the subsequent emission of fluorescence upon their return to lower energy levels. Our custom-built, inverted fluorescence microscope contains four lasers with different

excitation wavelengths, and therefore, four excitation pathways. A 514 nm laser (Coherent Genesis MX514 MTM) and a 561nm laser (Coherent Genesis MX561 MTM) is used for excitation of fluorescent emitters, while a 405 nm laser (Coherent OBIS 405) and a 488 nm laser (Genesis MX488-1000 STM) are used to 'activate' photo-activatable fluorescent emitters prior to excitation with the 561 nm laser. In each excitation pathway, the initial laser beam undergoes expansion through a two-lens telescope, resulting in the formation of a collimated beam with dimensions larger than the original input. The excitation beam is then circularly polarized after passing through a wavelength-appropriate zero order quarter wave plate. There is an additional bandpass filter in the 514 nm excitation pathway (Chroma ET510/10bp) to limit the excitation wavelength range. A set of dichroic mirrors (Chroma T470lpxr and Chroma T525lpxr) is then utilized, where the three excitation pathways converge into the same pathway. The excitation laser is then directed towards another dichroic mirror (Chroma ZT405-440/514/561rpc-UF1). This allows the excitation laser to be reflected into the objective of the microscope (UPLSAPO 60X 1.4 NA), which focuses the light onto the sample.

Photons from fluorescent emitters within the sample are collected by the objective lens. The objective lens used in the microscope is infinity corrected. In this design, the lens is designed so that parallel rays of light coming from an object at an infinite distance are focused at the back focal plane of the lens. This means that the focal plane is effectively moved to infinity. In order to capture the image plan, a tube lens is therefore used to collect and transmit emission light. After passing through the tube lens, the emitted light is then passed through two achromatic doublet lenses, which limits the effects of chromatic and spherical aberrations. These lenses are set up in a 4f configuration. The first lens performs

a Fourier transform of the image, which is multiplied by the DHPSF transfer function through a custom DHPSF phase mask. The second lens then converts the Fourier transform of the image back into the real image. The fluorescence signal finally reaches the detector and appears as a DHPSF, allowing for the 3D super-resolution capabilities (see section 2.1.1).

The emission signal also passes through a series of filters. A 514 nm long-pass filter (Semrock LP02-514RU-25) and 561 nm notch filter (Semrock NF03-561E-25) are used to limit the amount of scattered excitation light from the 514 nm and 561 nm excitation beams, respectively, entering the emission pathway. Additionally, a 700 nm short pass filter (Chroma ET700SP-2P8) is used to limit any additional light outside the range of the fluorescence signal from entering the emission pathway.

In order to switch between the 561 nm and 514 nm emission pathways, a dichroic beam-splitter (Chroma T560lpxr-uf3) placed after the first 4f lens directs the respective pathways to separate cameras. The resulting fluorescence signal is recorded on a scientific Complimentary Metal-Oxide Semiconductor (sCMOS) detector (Hamamatsu ORCA-Flash 4.0 V2). The advantages of sCMOS detectors over Charged-Coupled Device (CCD) detectors are plentiful. Specifically, sCMOS detectors exhibit higher quantum efficiency across a broader range of wavelengths, making them more sensitive in detecting light and resulting in better signal-to-noise ratios. Additionally, sCMOS detectors offer faster readout speeds as well as larger fields-of-view compared to CCD detectors.

2.2.2 Phase contrast imaging

After collecting fluorescence signal from a field-of-view, a phase contrast image of the bacterial cell shapes is acquired. Phase contrast imaging permits imaging of biological structures without a fluorescent label. By collecting a phase contrast image of bacterial cells after collection of single-molecule fluorescence data, individual molecules can be assigned to the bacterial cell they originate from.

The sample is illuminated with a red light-emitting diode (LED) that is positioned on the illumination tower above the microscope stage. After passing through a set of lenses, the red LED is then passed through an annulus ring, producing a ring of light. This ring is then passed through a condenser lens, which focuses the ring onto the sample stage. Importantly, light passing through, in this case, the cell membrane, will be scattered. Light passing through the membrane is phase shifted by -90° and is scattered in all directions. Similar to the fluorescence pathway, the light is then collected by the objective and passed through the tube lens. However, for collection of phase contrast data, a flip mirror situated after the tube lens is moved to the "up" position, and the light is reflected into a separate pathway. The light then passes through an additional 4f system, where a phase ring is placed in the Fourier plane between the two 4f lenses. As light passes through the phase ring, it is phase shifted by +90°. As a result, light that has not passed through the biological sample is phase shifted 180° with respect to the light that was scattered. The scattered and transmitted light then destructively interfere with one another, where this interference enhances the contrast between the biological sample and the transparent mounting substrate. The final phase contrast image is then visualized on the detection camera (Aptina MT9P031).

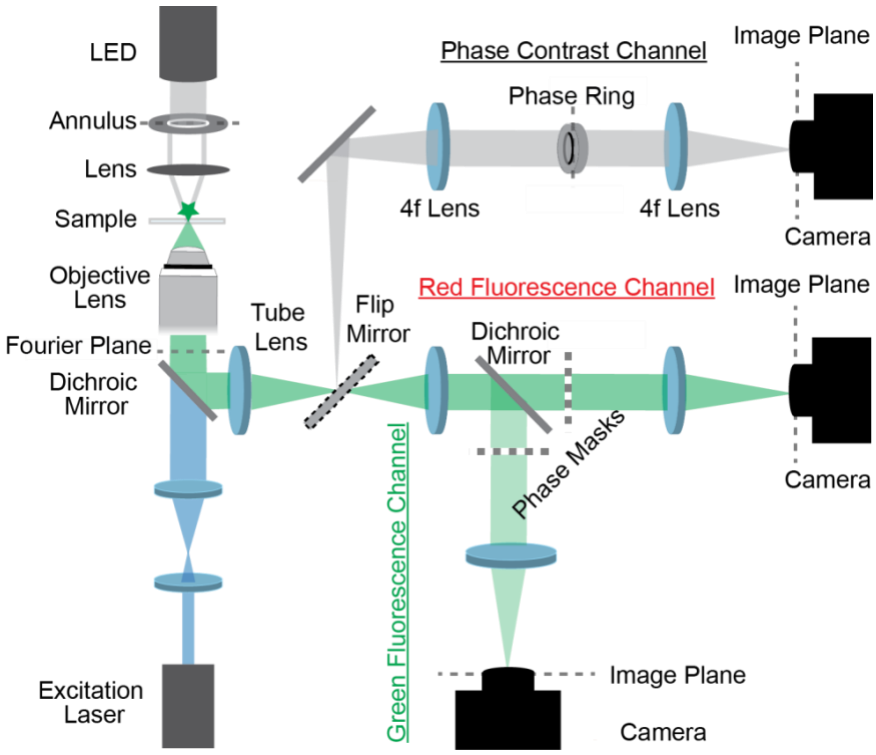

Figure 2.1. Schematic of microscope layout including excitation, emission, and phase contrast pathways. A detailed description of optical elements can be found in accompanying text. Special thanks to Dr. Alecia Achimovich for producing this figure.

Chapter 3
PROCESSING AND ANALYSIS OF SINGLE-MOLECULE TRAJECTORIES

**** Excerpts from this chapter are taken from Prindle, J. R., de Cuba, O.I.C., & Gahlmann, A. *Single-molecule tracking to determine the abundances and stoichiometries of freely-diffusing protein complexes in living cells: Past applications and future prospects* (J. Chem. Phys., 2023, Vol. 159, No. 7)**

3.1 Processing Experimental Single-molecule Localization Data

In order to extract the 3D positions of each single-molecule, the corresponding DHPSF signals must be processed and filtered. Given that the processing software used in this dissertation has been described in detail previously, only a brief explanation will be provided here. However, an additional filtering step was added to the existing software, which is described at the end of this section.

3.1.1 PSF fitting

Raw, single-molecule fluorescence data is processed using a modified version of the easy-DHPSF MATLAB software (56,82,83). Potential DHPSF signals are found based on a template image matching step. Specifically, a series of DHPSF images of a bright fluorescent bead is used to create a series of template images for different *z* positions. These template images are then used to match and select DHPSFs within the full experimental image. After template matching, single-molecule fluorescence data is fit with a modified double-Gaussian model. Specifically, a Maximum Likelihood Estimator (MLE) was implemented for the double-Gaussian fitting, which has been previously shown to be a more robust estimator of single-molecule positions compared to least-squares (LS). The center position between the two, fit lobes corresponds to the lateral *xy* position of the single-molecule, where the angle of the lobes corresponds the axial *z* position. If necessary, the fit localizations can be further filtered based on specific quality metrics, such as lobe distance, lobe diameter, and the number of photons collected for each localization.

Lateral and axial drift is a common problem encountered when collecting fluorescence imaging data. In order to correct for such drift, a bright fluorescent bead is

used as a fiducial marker for each single-molecule fluorescence movie. For each frame, the fiducial is fit to a double-Gaussian model, where the lateral and axial positions are determined. Obtaining the known positions of the fiducial therefore allows for drift correction of single-molecule localizations.

3.1.2 Cell registration and localization filtering

Single-molecule localizations are assigned to the bacterial cells they originate from, where localizations outside of cells were not considered for further analysis. Specifically, phase contrast images of bacterial cells were segmented using the OUFTI software (84). Abnormal cells, such as ones that are abnormally long or currently undergoing cell division, were discarded. Meshes of bacterial cell outlines were then input into a previously described MATLAB script. Specifically, the alignment of outlines with fluorescence data involves a two-step 2D affine transformation utilizing the 'cp2tform' function in MATLAB. In the initial step, five control point pairs were manually chosen for the segmented cell outlines and the corresponding single-molecule localizations. After generating the initial transformation, cells containing fewer than 10 localizations were discarded. The center of mass for all remaining cell outlines and the corresponding single-molecule localizations were then used to generate a more extensive set of control point pairs for calculating the final transformation function. The inclusion of a substantial set of control points ensures the robustness of the transformation, preventing biases introduced by cells with few localizations or those partially outside the field-of-view.

An additional filtering step was added to the cell-segmented single-molecule localizations. Specifically, localizations that are beyond or far from the axial (z) boundaries

of the cell were not included for further analysis (**Fig 3.1ab**). After registering single-molecule fluorescence data to their respective bacterial cell outlines, *z*-position quality scores of all cell-segmented single-molecule localizations were computed as

$$QS_i = \exp(-a * |z_i - \langle z \rangle|) \quad (3.1)$$

where a = 0.0008 and $\langle z \rangle$ is the average z-position of all emitters (**Fig 3.1c**). This step is necessary, as the following processing step links the single-molecule localizations into trajectories. Therefore, linking of molecules that are outside the axial (*z*) boundaries of the cell is avoided.

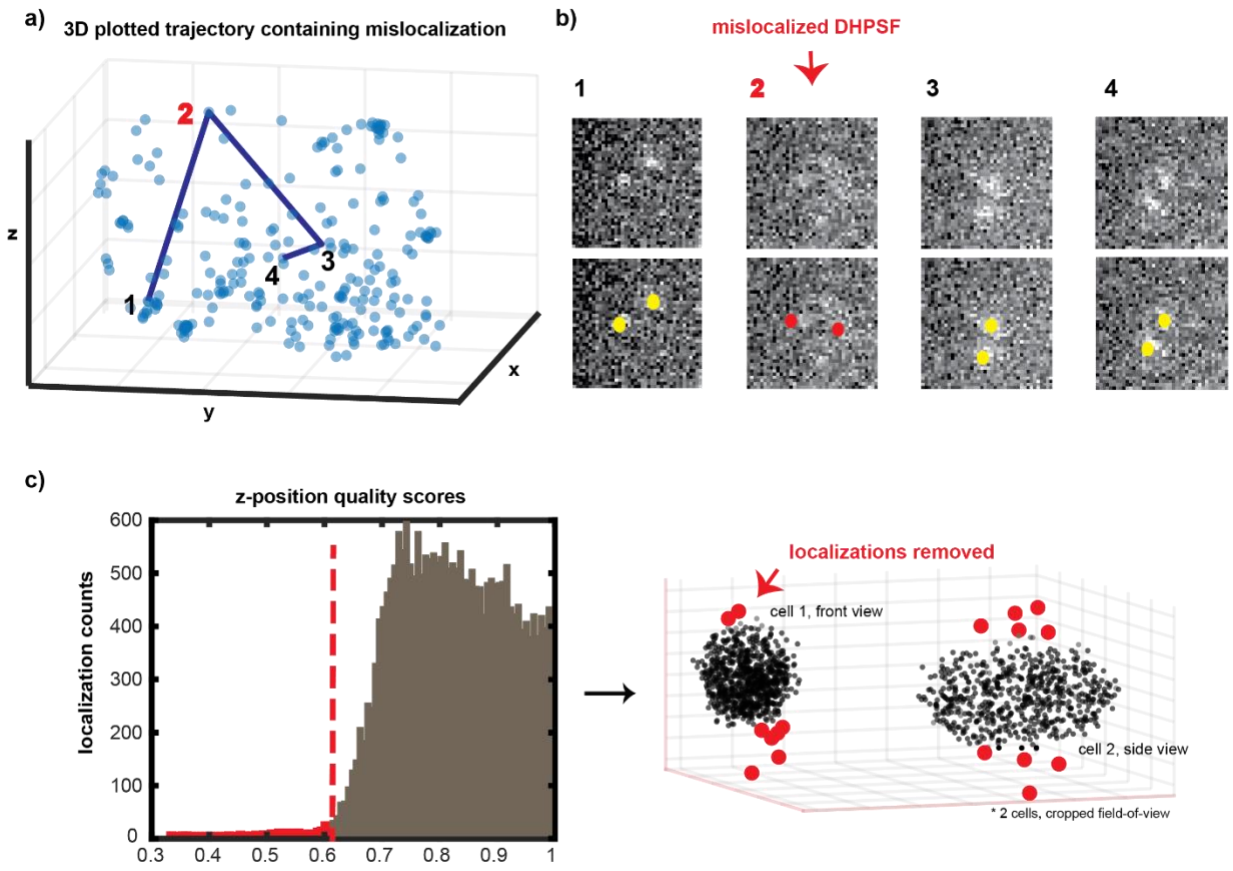

Figure 3.1: a) 3D scatterplot of *Ye*SctQ localizations in a single bacterial cell. Localizations of a putative single-molecule trajectory are connected by blue lines. However, the second localization corresponds to an incorrect localization. b) Individual camera frames giving rise to the connected localizations shown in panel a. The second frame shows only a dim fluorescence signal due to fluorophore blinking, which leads to mislocalization of the emitter. c) z-position quality scores of all localizations detected in $N \sim 40$ cells across the field-of-view. *z*-position quality score of the *i*-th emitter is computed as $QS_i = exp(-a * |z_i - \langle z \rangle|)$, where $a = 0.0008$ and $\langle z \rangle$ is the average z-position of all emitters.

3.2 Single-molecule tracking

Historically, single-molecule trajectories have been analyzed by computing the mean squared displacements (MSDs) at different time lags

$$MSD_n = \frac{1}{N-n}\sum_{i=1}^{N-n}(|\vec{r}_{i+n} - \vec{r}_i|)^2 \qquad n = 1, \ldots . N-1 \tag{3.2}$$

where N is the number of localizations in the trajectory and $\vec{r}_i$ are single-molecule positions within that trajectory, typically sampled at a constant time interval Δt. For a randomly diffusing Brownian particle, the MSD is proportional to the translational diffusion coefficient D, i.e.

$$MSD_n = 2mD_n\Delta t \tag{3.3}$$

where m is the dimensionality of the acquired single-molecule localizations ($d = 2$ for 2D tracking and $d = 3$ for 3D tracking). In any experiment, there are two main sources of localization errors that affect single-molecule position measurements. Static localization errors limit the precision of single-molecule position measurements, most notably because of the finite number of detected fluorescence photons (60,61) (**Fig 3.2a**). In addition, localizations of moving emitters suffer from dynamic localization errors, because the fluorescence signal is motion-blurred (**Fig 3.2b**). Dynamic localization errors limit both the precision and accuracy of single-molecule localizations. The combined localization error σ modifies the MSD vs. time relationship to $MSD_n(n;\Delta t) = 2dDn\Delta t + 2d\sigma^2$ (85). The straightforward relationship between the experimental observable, MSD_n, and the quantity interest, D, has made MSD analysis an immensely popular approach in the biophysical life sciences. However, many (>1000) trajectory points are needed to reliably

estimate (within 10% error) the diffusion coefficient of a single-molecule acquired with experimentally-realistic localization errors (85). Experimentally acquired single-molecule trajectories typically contain only around 10-1000 displacements, which does not meet the above threshold. If the tracked particles undergo unconfined Brownian motion that is governed by a single diffusion coefficient value, >1000 trajectory points could be obtained by pooling a sufficient number of short trajectories. In such a case, MSD analysis would provide an accurate estimation of the diffusion coefficient. However, the assumption of homogeneous, unconfined Brownian motion is unlikely to hold for protein diffusion in living cells. MSD analysis is therefore not suitable to determine diffusion coefficient(s) of single molecules in living cells.

To overcome the drawbacks of MSD analysis, the field has developed a variety of alternative approaches. These approaches are based on curve-fitting (56,83) or Bayesian inference methods (86–89) to analyze pooled trajectory data. In the following section, the curve-fitting approach used to analyze trajectory data in this dissertation will be outlined. The full processing and analysis pipeline is depicted in Figure 3.4.

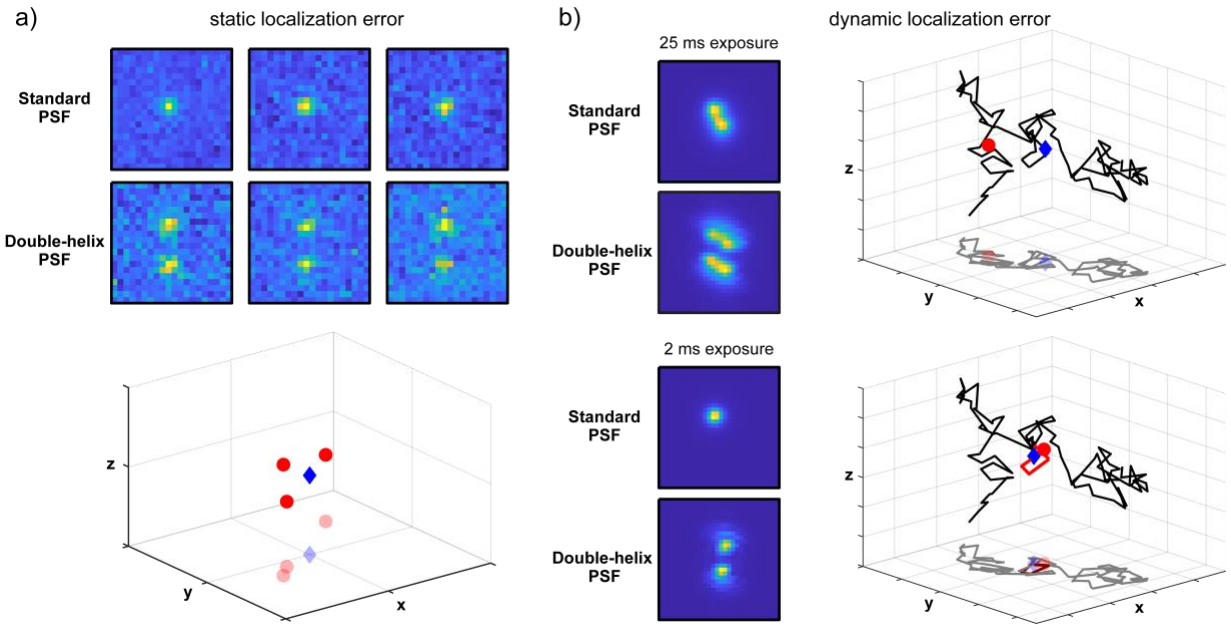

Figure 3.2. Experimental considerations affecting single-molecule tracking results. Special thanks to Olivia de Cuba for generating the data used for and helping create this figure. (**a**) For a stationary fluorophore at a fixed position in space (blue diamond), different PSF images will be obtained at different times due to uncorrelated Gaussian camera read noise and Poisson shot noise. (Two sets of simulated images are shown for the standard PSF and the double-helix PSF, respectively). Because of these random noise contributions, the estimated emitter positions (red circles) will differ from the actual position of the emitter. (**b**) For a freely-diffusing fluorophore, motion-blurred PSF images will be obtained. (Two sets of simulated images are shown for the standard PSF and the double-helix PSF, respectively. Gaussian camera read noise and Poisson shot noise are omitted here for clarity). Because of motion blur, the estimated emitter position (red circle) will differ from the center-of-mass position of the emitter during the exposure/illumination time (blue diamond). By reducing the exposure/illumination time from 25 ms to 2 ms (referred to as stroboscopic illumination), the amount of motion blur can be reduced.

3.2.1 The apparent diffusion coefficient

Filtered 3D single-molecule localizations in subsequent frames were linked into trajectories using a distance threshold of 2.2 μm. Trajectories with fewer than 4 localizations were not considered for further analysis. Additionally, if 2 or more localizations were present in the same cell at the same time, the trajectories corresponding to each localization were not considered for further analysis. After computing the MSD of each trajectory, the apparent diffusion coefficient, D^*, could then be computed as

$$D^* = \frac{MSD}{2 * m * \Delta t} \tag{3.4}$$

where m is the dimensionality of the data and Δt is the camera exposure time. For tracking experiments performed in this dissertation, m = 3 and Δt = 25 ms. The calculated diffusion coefficient from single-molecule MSD_n calculations (i.e. averaged over a single-molecule trajectory) should be regarded as the apparent single-molecule diffusion coefficient D^* to distinguish it from the actual diffusion coefficient D that governs the motion of the molecule. Estimates of D^* differ from D because of the drawbacks of MSD analysis, namely short trajectory lengths and localization errors due to limited photon budgets and motion blurring, but also confinement of trajectories within small volumes, such as bacterial cells (56,83,90,91), cellular organelles (92), and phase-separated condensates (93,94). For each tracking experiment presented in this dissertation, thousands of single-molecule trajectories were acquired and, therefore, thousands of D^* values were computed for a tracked molecule of interest.

3.2.2 Generating simulated model functions

A previously developed approach to account for the complex interdependencies between trajectory length, localization errors, and confinement effects contained in experimental D^* distributions (**Fig. 3.3a**) by simulating D^* distributions (**Fig. 3.3b**) and then using them to fit the experimental D^* distributions was modified and further optimized (56,95). For bacterial cell imaging, realistically simulated D^* distributions was achieved through a numerical forward convolution approach that begins with Monte Carlo sampling of Brownian motion within a confining volume matching the size and shape of the bacterial cells being imaged. Fine time steps of 100 ns are chosen to generate a list of positions visited by a diffusing molecule during the chosen camera and/or laser exposure time (e.g. 25 ms). The fine trajectory is split up into 50 segments, and the center-of-mass within each segment is computed. The center-of-mass positions are then convolved with the microscope's PSF to generate 50 subframes, which are summed to generate a motion-blurred image of the diffusing molecule. Next, the signal intensity is scaled to match the experimentally observed signal intensities, and the resulting image is modulated by the addition of background intensity as well as by the addition of Gaussian camera read noise and Poisson shot noise that is inherent to the photon detection process. Simulated localizations and trajectories are obtained through the same localization and tracking algorithms as used for experimentally acquired data, and an apparent diffusion coefficient D^* is calculated for each simulated trajectory. This process is repeated for N = 5000 trajectories to obtain a well-sampled D^* distribution corresponding to a single "true" diffusion coefficient, D.

The simulation library used to analyze empirical D^* distributions was modified. The previous simulation library contained 64 D^* distributions, which corresponded to D values ranging from 0.05 to 20 $\mu m^2/s$. This library was updated after simulating additional D^* distributions with D values ranging from 0.01 to 0.1 $\mu m^2/s$ with a step size of 0.01 as well as 0.1 to 0.5 $\mu m^2/s$ with a step size of 0.05. Additionally, D^* distributions ranging from 16 to 20 $\mu m^2/s$ were removed from the library. The final library contains 70 diffusion coefficient values ranging from 0.01 to 15 $\mu m^2/s$.

In order to analyze empirical D^* distributions (**Fig. 3.3a**) with the simulated D^* distributions, where the "true" diffusion coefficient, D, is known, the above library of simulated D^* distributions (**Fig. 3.3b**) was used to generate a 2D gridded interpolant function. The empirical cumulative distribution function (eCDF) is first computed for each simulated distribution. After initializing variables and setting query points, the eCDFs are padded, extending them for subsequent interpolation. The interpolation process is executed separately for the first 25 simulated curves and the remaining curves, with distinct interpolation methods employed for optimal accuracy. Specifically, the first 25 curves of the library were interpolated using a B-spline (order 3), while the remaining curves were interpolated using a cubic smoothing spline. A second interpolation along the diffusion coefficient dimension using a cubic smoothing spline with the addition of a smoothing parameter was then performed. The smoothing parameter controls the trade-off between fitting the data closely and achieving a smooth interpolation. For the first 200 data points, the smoothing weights are adjusted to have higher roughness at the beginning of the curve and lower roughness at the end of the curve. Beyond the first 200 points, the smoothing weights are set to have moderate roughness throughout the entire range of diffusion

coefficient values. Each of the 1D interpolations then serve as an input to compute the 2D interpolation using the 'griddedInterpolant' MATLAB function. Finally, a matrix of curves computed from the 2D gridded interpolant is generated for subsequent analysis of empirical D^* distributions.

3.2.3 Linear fitting – Diffusion coefficient spectrum

After obtaining the simulated model function that directly accounts for the complex interdependencies between trajectory length, localization errors, and confinement effects contained in experimental D^* distributions, curve-fitting of experimental D^* distributions is now possible. Major revisions were made to the previously established curve-fitting method (56). Specifically, the curve-fitting approach of D^* distributions utilized in this dissertation follows a two-step process: 1) linear fitting using the trust-region method, which results in a spectrum of D values, and 2) nonlinear fitting using the particle swarm optimization method, which refines the number of parameters selected in the linear fitting process into a defined number of diffusive states and their relative population fractions.

Trust region methods are optimization algorithms used to solve unconstrained or constrained optimization problems (96). Specifically, trust-region methods iteratively optimize a model of the objective function within a region around the current iterate, also known as the trust region. The trust region is updated dynamically based on the performance of the model compared to the true objective function. Within each iterate, the trust region is either accepted or updated. If the model accurately represents the objective function, the algorithm accepts the step and moves towards the minimum. However, if the model deviates significantly, the size of the trust region is adjusted to restrict the step size,

preventing too large of steps that may lead to divergence or overshooting of the minimum. Ultimately, trust region methods provide a robust and computationally efficient approach to optimization.

For the software developed in this dissertation, the objective function is defined by the 'lsqlin' MATLAB function, which is used for constrained least squares optimization. The objective function in this context is the sum of squared differences between the predicted values and the actual data points. Specifically, 'lsqlin' seeks to minimize the residual sum of squares (RSS)

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3.5}$$

where y_i is the observed value for the i-th data point, $\hat{y}_i$ is the predicted value for the i-th data point obtained from the model, and n is the total number of data points. After interpolation of the observed data, the software sets up the optimization problem using the 'trust-region-reflective' algorithm within the 'lsqlin' MATLAB function.

The software utilizes a bootstrapping approach, which involves repeatedly sampling subsets from the original dataset and estimating the parameters of interest for each samples subset. During each iteration, the sampled data are then fit based on the above minimization problem, and 100 total bootstrapping iterations are performed. Bootstrapping therefore provides 100 different spectra that were obtained for each sampled subset of the data, allowing for the calculation of a mean bootstrapped spectrum as well as its standard deviation. The mean bootstrapped spectrum ultimately contains the fitted linear coefficients that most accurately represent the observed data. These linear coefficients are

then plotted as a spectrum of D values that manifest for a Brownian diffuser confined within the cytosol of rod-shaped bacterial cells (**Fig. 3.3c**).

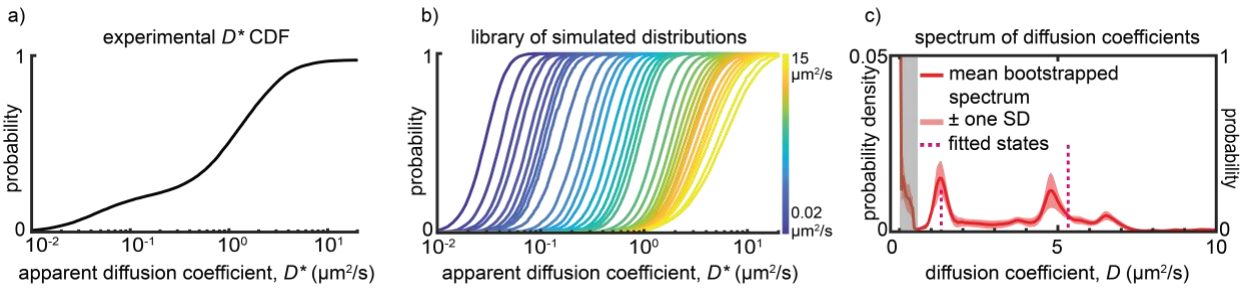

Figure 3.3. (**a**) Experimentally measured D^* distribution of eYFP-labeled SctQ in *Yersinia enterocolitica*. (**b**) Simulated distributions of apparent diffusion coefficients based on Monte Carlo simulations of confined Brownian motion in rod-shaped bacterial cells. These simulations account for both random and systematic measurement errors encountered in single-molecule tracking measurements. (**c**) The diffusion coefficient spectrum (red) of eYFP-*Ye*SctQ shows two prominent peaks centered at D = 1.3 μm^2/s and D = 4.8 μm^2/s, as well as a smaller peak centered at D = 6.5 μm^2/s. Bootstrapping analysis provides the standard deviation of the spectrum (red, shaded). A two diffusive state model (pink line, dashed) was selected by 5-fold cross-validation. The different diffusive states are due to *Ye*SctQ participating in different hetero-oligomeric protein complexes in the bacterial cytoplasm. SctQ also dynamically binds and unbinds from the membrane-imbedded type III secretion system. Thus, a large stationary population (shaded gray area) is observed as well.

3.2.4 Nonlinear fitting – Diffusive state analysis

The peaks and integrated peak areas in the diffusion coefficient spectrum provide the initial values for curve fitting using a more constrained model function

$$CDF_{simulated} = \sum_{i=1}^{K} a_i CDF_{simulated,i}(D_i) \tag{3.6}$$

$$\sum_{i=1}^{K} a_i = 1 \tag{3.7}$$

where $K = 1,2,3\ldots$ enumerate the different diffusive states that are resolved in the diffusion coefficient spectrum and a_i are the nonlinear coefficients. Determining the optimal K-state model has historically been challenging, because optimization routines can get stuck in local minima corresponding to vastly different parameter values that, in some cases, are heavily influenced by the initial parameter values.

To determine the optimal fit parameters for each K-state model, particle swarm optimization was used. In particle swarm optimization, the algorithm iteratively updates a population of potential solutions, which are referred to as particles, with each particle representing a potential set of parameters that describe the observed data (97). At each iteration, the particles adjust their positions, or parameter values, based on their own historical best positions and the global best positions found by any particle in the swarm. This adjustment is determined by the objective function 'linearComb' in MATLAB, which computes the mean squared error (MSE) between the predicted values and the observed data for each set of parameters. The MSE is computed as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3.8}$$

The parameters of the nonlinear fit model are extracted from the particle positions, and the model's predicted values are computed based on these parameters. Ultimately, the result of the particle swarm optimization routine includes optimal parameter sets, or the diffusion coefficient(s), *D,* and its respective population fraction(s), for different *K*-state models.

An independent approach to select the optimal *K*-state model that does not overfit the data with too many diffusive states is *m*-fold cross-validation (98). In this approach, the data is divided into *m* subsets. One subset is considered the "validation" data while the others are considered "training" data. A model is then fit to the training data and evaluated on the validation data. Generally, as more parameters are added to the model, the cross-validated errors decrease, reach a minimum, and then increase again due to overfitting. When fitting experimental D^* CDFs, the optimal *K*-state model can thus be selected based on two criteria: the *m*-fold cross-validation error is comparable to or lower than those of other models, and the model fit parameters agree with the peaks observed in the diffusion coefficient spectrum. This 2-step procedure addresses the problem of model selection commonly encountered in frequentist (i.e. least squares- or maximum likelihood-based) curve fitting approaches of single-molecule tracking data.

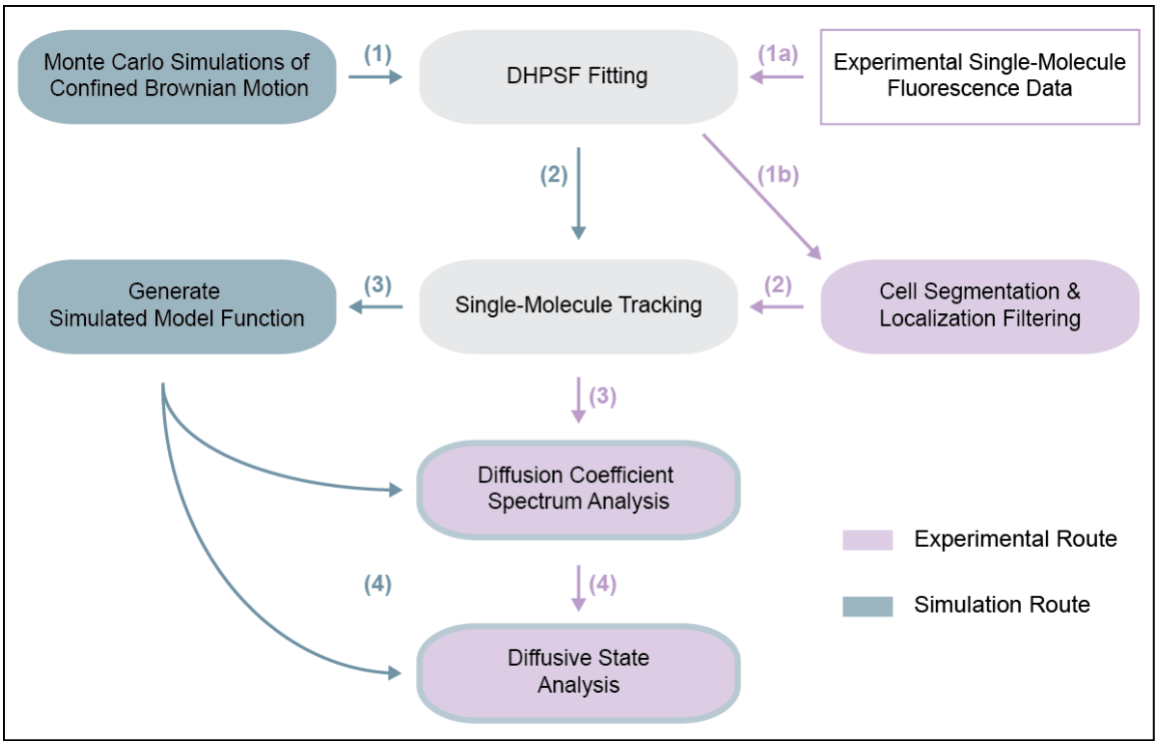

Figure 3.4: Data processing pipeline for experimental and simulation data of spatially-confined, motion-blurred DHPSFs. Connected lines represent inputs from a previous processing module into the proceeding module, (**pink**) Experimental processing route. 1a) Experimental, single-molecule fluorescence data initially undergoes DHPSF fitting (gray). 1b) Successful DHPSF fits are filtered from cell-segmented meshes (*xy*) and axial (*z*) position quality scores. 2) Cell- and axial position-filtered DHPSFs undergo nearest neighbor tracking (gray), forming single-molecule trajectories with respective apparent diffusion coefficients. 3) Diffusion coefficient spectrum analysis and 4) diffusive state analysis of the corresponding apparent diffusion coefficient distributions. (**blue**) Simulation processing route. 1) Spatially-confined, motion-blurred DHPSFs from Monte Carlo simulations are fitted (gray) in the same manner as experimental data. 2) Succesful DHPSF fits undergo nearest-neighbor tracking (gray) in the same manner as the processing route. 3) A library of simulated apparent diffusion coefficients are used to generate the 2D

gridded-interpolant fitting function. 4) The simulated model function is used to fit experimental apparent diffusion coefficient distributions within the diffusion coefficient spectrum analysis module and diffusive state analysis module.

Chapter 4
DISTINCT CYTOSOLIC COMPLEXES CONTAINING THE TYPE III SECRETION SYSTEM ATPASE RESOLVED BY THREE-DIMENSIONAL SINGLE-MOLECULE TRACKING IN LIVE *YERSINIA ENTEROCOLITICA*

****Chapter adapted from Prindle, J. R., Wang, Y., Rocha, J. M., Diepold, A., & Gahlmann, A. *Distinct Cytosolic Complexes Containing the Type III Secretion System ATPase Resolved by Three-Dimensional Single-Molecule Tracking in Live Yersinia Enterocolitica* (ASM Microbiology Spectrum, 2022, Vol. 10, No. 6)**

JR Prindle created bacterial strains required for imaging, designed and performed experiments, and generated software necessary for analysis of single-molecule tracking data. Y Wang generated the software necessary for *k*-fold cross-validation. JM Rocha helped collect single-molecule tracking data.

4.1 ABSTRACT

The membrane-embedded injectisome, the structural component of the virulence-associated type III secretion system (T3SS), is used by gram-negative bacterial pathogens to inject species-specific effector proteins into eukaryotic host cells. The cytosolic injectisome proteins are required for export of effectors and display both stationary, injectisome-bound populations as well as freely-diffusing cytosolic populations. How the cytosolic injectisome proteins interact with each other in the cytosol and associate with membrane-embedded injectisomes remains unclear. Here, we utilize 3D single-molecule tracking to resolve distinct cytosolic complexes of injectisome proteins in living *Yersinia enterocolitica* cells. Tracking of the eYFP-labeled ATPase, *Ye*SctN, and its regulator, *Ye*SctL, reveals that these proteins form a cytosolic complex with each other and then further with *Ye*SctQ. *Ye*SctNL and *Ye*SctNLQ complexes can be observed both in wild type cells and in Δ*sctD* mutants, which cannot assemble injectisomes. In Δ*sctQ* mutants, the relative abundance of the *Ye*SctNL complex is considerably increased. These data indicate that distinct cytosolic complexes of injectisome proteins can form prior to injectisome binding, which has important implications for how injectisomes are functionally regulated.

4.2 IMPORTANCE

Injectisomes are membrane-embedded, multiprotein assemblies used by bacterial pathogens to inject virulent effector proteins into eukaryotic host cells. Protein secretion is regulated by cytosolic proteins that dynamically bind and unbind at injectisomes. However, how these regulatory proteins interact with each other remains unknown. By measuring the

diffusion rates of single molecules in living cells, we show that cytosolic injectisome proteins form distinct oligomeric complexes with each other prior to binding to injectisomes. We additionally identify the molecular compositions of these complexes and quantify their relative abundances. Quantifying to what extent cytosolic proteins exist as part of larger complexes in living cells has important implications for deciphering the complexity of biomolecular mechanisms. The results and methods reported here are thus relevant for advancing our understanding of how injectisomes, and related multiprotein assemblies, such as bacterial flagellar motors, are functionally regulated.

4.3 INTRODUCTION

Bacterial type III secretion systems (T3SS) are used by gram-negative bacteria to assemble flagellar motors for cell motility and injectisomes that translocate virulence factors, called effector proteins, into eukaryotic host cells. Injectisomes are expressed by many bacterial pathogens, including *E. coli*, *Salmonella*, *Pseudomonas*, *Shigella*, and *Yersinia*, which are responsible for widespread human disease both historically and currently. Flagellar motors and injectisomes are large, multicomponent protein structures that span two and sometimes three cell membranes, as well as the bacterial cell wall (12). Assembly of the injectisome is achieved through hierarchal secretion of early, intermediate, and late secretion substrates (3,4,12,99–102). Early secretion substrates form an extracellular needle that extends away from the bacterial cell envelope. Upon contact with a host cell, middle secretion substrates are secreted through the needle to form a translocon pore in the eukaryotic cell membrane (12,99–101). Late secretion substrates, also called effector proteins, are then translocated into the host cell cytosol (3,4,102). While there are

species-specific differences among the effector proteins, the structural proteins of injectisomes are highly conserved (103).

Molecular-resolution structures of the fully-assembled cytosolic injectisome complex remain elusive. The cytosolic injectisome proteins SctN, L, Q, and K, associate with the injectisome loosely and transiently, so that structural imaging of fully-assembled injectisomes is only possible *in situ* through the use of cryo-electron tomography (cryo-ET) (104,38,105–107). Injectisomes in *Salmonella* and *Shigella* minicell mutants were found to not exhibit continuous densities representative of a cytosolic ring (C-ring) (38,104), a feature that has been observed consistently for bacterial flagellar motors (100,108). Instead, the sub-tomogram averages suggested six distinct pods composed of SctK, Q, and L that seem to cradle the hexameric ATPase (38,104). More recently, cryo-focused ion beam milling followed by cryo-ET enabled the visualization of *Y. enterocolitica* injectisomes inside the phagosomes of infected human myeloid cells (40). In contrast to the six distinct cytosolic pods observed in *Shigella* and *Salmonella* minicell mutants, the *Y. enterocolitica* injectisome contained more continuous, ring-like densities likely composed of SctD, K, and Q, and a 6-fold symmetric, cradle-like structure likely composed of SctL.

The cytosolic injectisome proteins have substantial freely-diffusing populations. *In vivo* fluorescence recovery after photobleaching (FRAP) experiments revealed a continual exchange (on a 1-minute time scale) between freely-diffusing and injectisome-bound states for *Ye*SctQ (*Ye* prefix indicates *Yersinia enterocolitica*). Interestingly, the *Ye*SctQ exchange rate at the injectisome increased by a factor of two upon chemical activation of secretion (45), suggesting that structural dynamism may be important for functional

regulation of type 3 secretion. Fluorescence correlation spectroscopy (FCS) further revealed the presence of freely-diffusing and injectisome-bound populations for *Ye*SctK, L, and N (46). In that study, the *average* diffusion rate for each cytosolic injectisome protein shifted upon chemical activation of secretion. These results further point to a dynamic and adaptive network of cytosolic interactions that may be required for injectisome binding. Earlier work from our lab identified two distinct *Ye*SctQ-containing cytosolic protein complexes in living cells (83). However, the full compositions of these and additional SctQ-independent cytosolic complexes are yet to be determined.

Efforts at reconstituting cytosolic injectisome protein complexes *in vitro* have begun to elucidate how these proteins may interact with one another. Native mass spectrometry (MS) experiments showed the ability of *Se*SctQ (*Se* prefix indicates *Salmonella enterica*) to bind *Se*SctL and N with various stoichiometries. Importantly however, a *Se*SctL$_2$:*Se*SctN heterotrimer was consistently observed in any complex that included these two proteins (109). Dynamic light scattering was used to determine the relative size of cytosolic injectisome proteins and their respective complexes (44). Key differences were observed for reconstituted complexes containing *Sf*SctK and *Sf*SctQ and complexes containing *Sf*SctQ and *Sf*SctL (*Sf* prefix indicates *Shigella flexneri*). Specifically, the complex of *Sf*SctK and *Sf*SctQ is globular and compact (hydrodynamic diameter = 10.2 nm), whereas the complex of *Sf*SctQ and *Sf*SctL is much larger (hydrodynamic diameter = 18.5 nm). These differences were attributed to the ability of *Sf*SctQ to adopt different conformational states depending on its binding partner.

To gain further insights into how the cytosolic injectisome protein complexes interact with each other in living *Y. enterocolitica*, we use 3D single-molecule localization

and tracking microscopy to resolve the different diffusive states of *Ye*SctQ, L, and N in different genetic backgrounds. We show that deletion of the injectisome protein *Ye*SctD, which presumably binds *Ye*SctK(40), abrogates the clustered, membrane proximal localization of all cytosolic injectisome proteins and increases the abundance of the same cytosolic complexes that are also present in wild-type cells. We further find that deletion of *Ye*SctQ increases the abundance of a complex containing *Ye*SctN and L and eliminates the diffusive state assigned to the larger complex containing *Ye*SctN, L, and Q. Comparing the diffusion rates of all tracked proteins allows us to distinguish between oligomerization models that differ in terms of protein complex stoichiometry. Our combined results narrow down the possible stoichiometries of cytosolic injectisome protein complexes that can form prior to injectisome binding.

4.4 RESULTS

4.4.1 Diffusion of monomeric eYFP under secretion-active conditions

Tracking of the cytosolic injectisome proteins under secretion-active conditions requires that cells are exposed to a $25^\circ \rightarrow 37^\circ$ C temperature jump and that calcium ions are removed from the growth medium through chelation (110–114). Concurrent with the removal of Ca^{2+} ions, $MgCl_2$ and glycerol are added (see methods). The altered growth medium composition results in a different osmotic environment for the cells, which could alter the density of the cytosol. To quantify this effect, we tracked monomeric eYFP under secretion-active conditions in living *Y. enterocolitica* cells. The pooled single-molecule trajectories provide highly sampled distributions of molecular motion behaviors, quantified using the apparent diffusion coefficient (D^*). A simulated library of D^* distributions

(**Fig. 4.1a**) enables linear fitting of experimentally acquired *D** distributions to obtain spectra of intracellular Brownian diffusion coefficients that manifest for a given protein. Peaks in these spectra can be analyzed to determine the diffusion rates of different diffusive states, and the areas underneath the peaks determine their relative abundances. These parameters can then be further refined by nonlinear fitting of a defined number of diffusive states (see Methods). Transforming the raw data (cumulative distribution functions (CDFs) of apparent diffusion coefficients) into diffusion coefficient spectra addresses a key challenge in analyzing single-molecule tracking data, namely choosing a suitable fitting model among many with different numbers of parameters (diffusive states and their relative abundances) (88,115).

For monomeric eYFP under secretion-active conditions, we observe a diffusion coefficient spectrum exhibiting a single, well-defined peak at D = 9.1 μm^2/s that produces a good fit to the experimentally acquired *D** distribution (**Fig. 4.1bc**). Non-linear fitting using a single diffusive state at D = 9.2 μm^2/s results in an equally good fit (**Fig. 4.1bc**). The narrowness of the sole spectral peak shows that the motion of eYFP is homogeneous in the *Y. enterocolitica* cytosol and well described by a single diffusive state with a Brownian diffusion coefficient at 9.2 μm^2/s. This property establishes eYFP as a suitable probe for measuring the diffusion coefficients for eYFP-labeled proteins. We note that a D = 9.2 μm^2/s diffusive state is slower than the single diffusive state at 11.9 μm^2/s, which we measured previously under standard, i.e. non-secreting, growth conditions (83,116). This observation is consistent with an overall osmotic upshift that increases the biomolecular density of the cytosol and thus limits the mobility of eYFP. Controlled osmotic upshifts

have been previously shown to slow diffusion of fluorescent and fluorescently-labeled proteins (117,118).

Under secretion-active conditions, we also observe a small (7%) population of slow (D < 0.5 μm^2/s) eYFP, indicating that even small proteins can occasionally get trapped, presumably in crowded pockets of the cytosol (117). We verified that trajectories corresponding to slow moving eYFP are indeed randomly localized to the cytosol and do not show a preference for subcellular locations, such as the membrane or the cell poles (**Fig. 4.2a**). Notably, the slow population is not detectable when tracking eYFP under standard growth conditions.

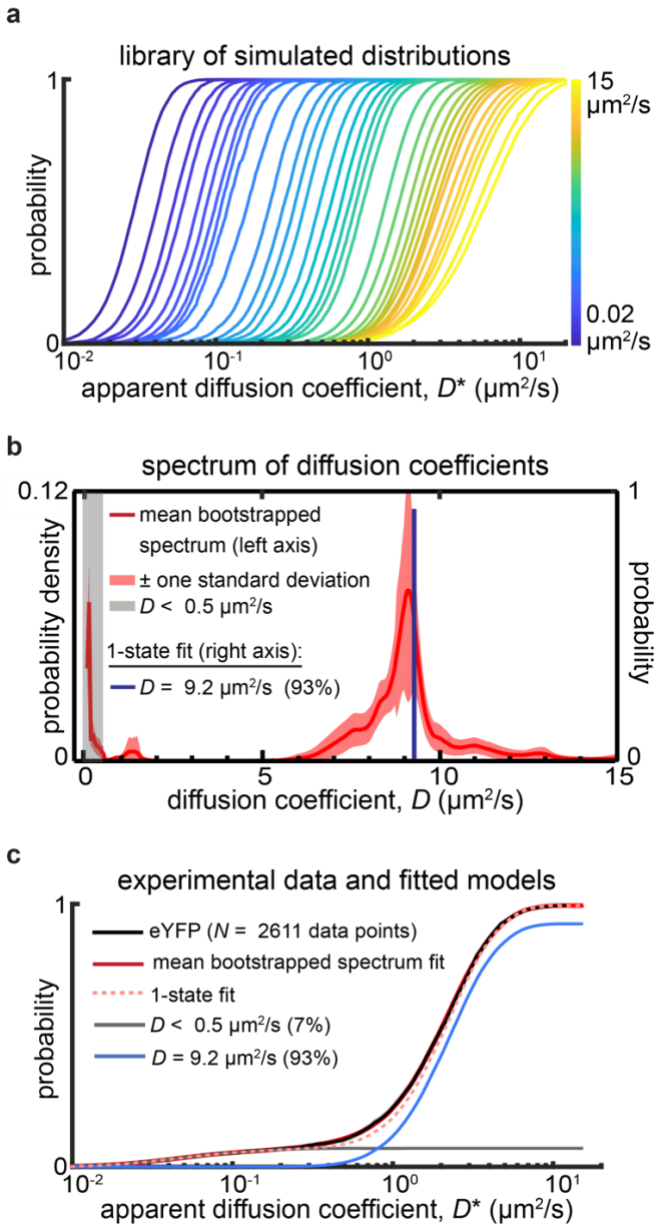

Figure 4.1: a) Simulated distributions of apparent diffusion coefficients based on Monte Carlo simulations of confined Brownian motion in rod-shaped bacterial cells. These simulations account for both random and systematic measurement errors encountered in single-molecule tracking measurements (see Methods). (**b** and **c**) The experimentally measured apparent diffusion coefficient distribution of freely diffusing eYFP molecules is fit well using a tightly peaked distribution of diffusion coefficients centered at D =

9.2 μm^2/s (red curve in panel b) or using a single diffusive state with D = 9.2 μm^2/s (blue line in panel c). The excellent agreement between theory and experiment supports that the assumption of bacterial cell-confined Brownian diffusion is valid.

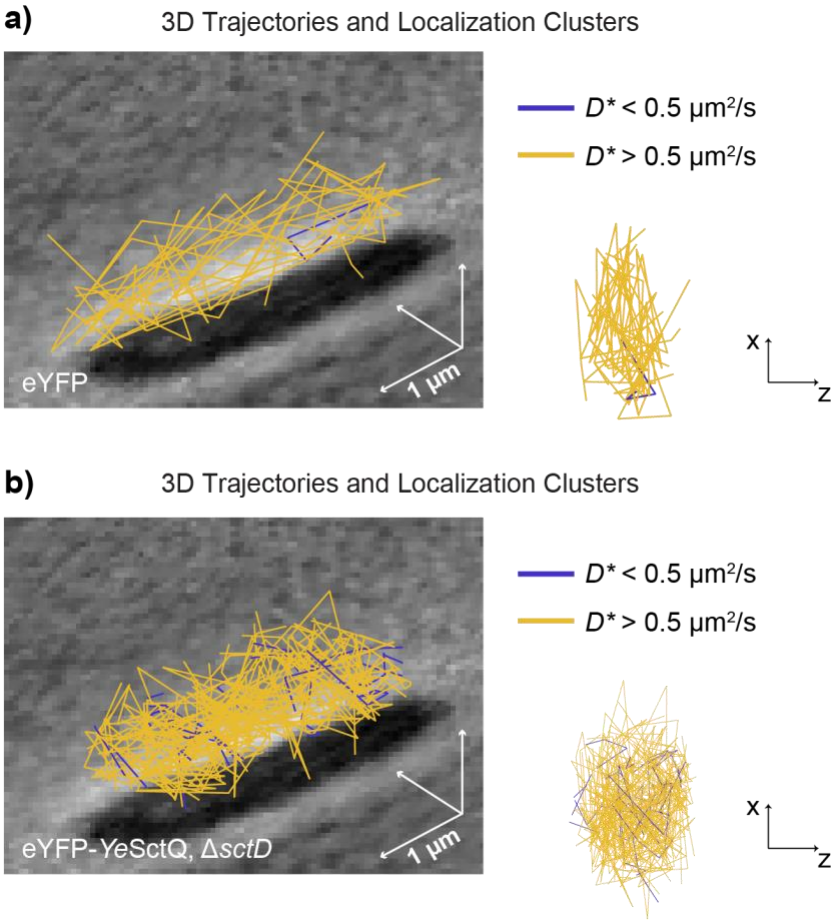

Figure 4.2: a) 3D trajectories for eYFP under secretion active conditions. b) Trajectories classified as with D* < 0.5 μm2 /s are not localized to the cell membrane. b) 3D trajectories for YeSctQ ΔsctD under secretion active conditions. Trajectories classified as with D* < 0.5 μm2 /s are not localized to the cell membrane.

4.4.2 Diffusive state assignment of *Ye*SctQ indicates two distinct protein complexes

Our previously published *Ye*SctQ tracking results in *Y. enterocolitica* under secretion-active conditions suggested the presence of three diffusive states at $D \sim 1.0$, 4.0, and 15 μm^2/s, with the $D \sim 15$ μm^2/s diffusive state constituting a 20% population fraction of all tracked proteins. Such a result is however inconsistent with the monomeric eYFP tracking result, as an eYFP-labeled protein should not be able to diffuse faster than 9.2 μm^2/s under secretion-active conditions. We therefore re-examined the raw *Ye*SctQ single-molecule trajectories and found that uncharacteristically fast apparent diffusion coefficients are often due to large displacements resulting from mislocalization of fluorescence signals beyond the axial (*z*-) boundaries of the cell (**Fig. 3.1**). Removal of these localizations and the associated trajectories removed the fast $D = 15$ μm^2/s diffusive state. Importantly, the diffusive states below 9.2 μm^2/s are largely unaffected by this filtering step.

In contrast to freely diffusing eYFP (**Fig. 4.1**), the diffusion coefficient spectrum of *Ye*SctQ shows two prominent peaks at $D \sim 1.3$ and 4.8 μm^2/s and a small peak at 6.5 μm^2/s in addition to an injectisome-bound, stationary population at $D = 0$ μm^2/s (**Fig. 4.3a**). eYFP-*Ye*SctQ fusion proteins were expressed as a genomic replacement under the control of its native promotor in actively secreting cells, and the fusion protein was stable and not degraded (83). These results are roughly consistent with our previous study (83), which identified three cytosolic diffusive states for *Ye*SctQ at $D \sim 1.1$, 4.0, and 13.9 μm^2/s (the fastest state was assigned to the eYFP-*Ye*SctQ monomer). The distinct peaks in the newly analyzed diffusion coefficient spectrum provide the initial diffusion rates and populations fractions for non-linear fitting of the experimental D^* CDF of *Ye*SctQ. Still, to test whether

overfitting the data with too many different diffusive states (i.e. a non-linear model with too many free parameters) could be an issue, we initialized 1-, 2-, and 3-state fits by combining population fractions and initial guesses for diffusion coefficients from the spectrum (see Methods). Non-linear fitting using a 2-state model produced diffusive states at D ~ 1.3, and 5.3 μm^2/s, which agrees with the prominent peaks in the diffusion coefficient spectrum (**Fig. 4.3a**) although the 5.3 μm^2/s diffusive state appears to be right shifted by the small spectral density at 6.5 μm^2/s. Non-linear fitting using a 3-state model produced diffusive states at D ~ 1.2, 4.4, and 7.0 μm^2/s, however the population fraction of the D ~ 7.0 μm^2/s diffusive state appears to be vastly overestimated (**Fig. 4.4a**). Additionally, 5-fold cross-validation analysis favors a 2-state over a 3-state model (**Fig. 4.4b**). These results suggest that the majority of *Ye*SctQ is part of two predominant cytosolic complexes diffusing at D ~ 1.3 and 4.8 μm^2/s, while the monomeric fraction of this injectisome protein remains small by comparison.

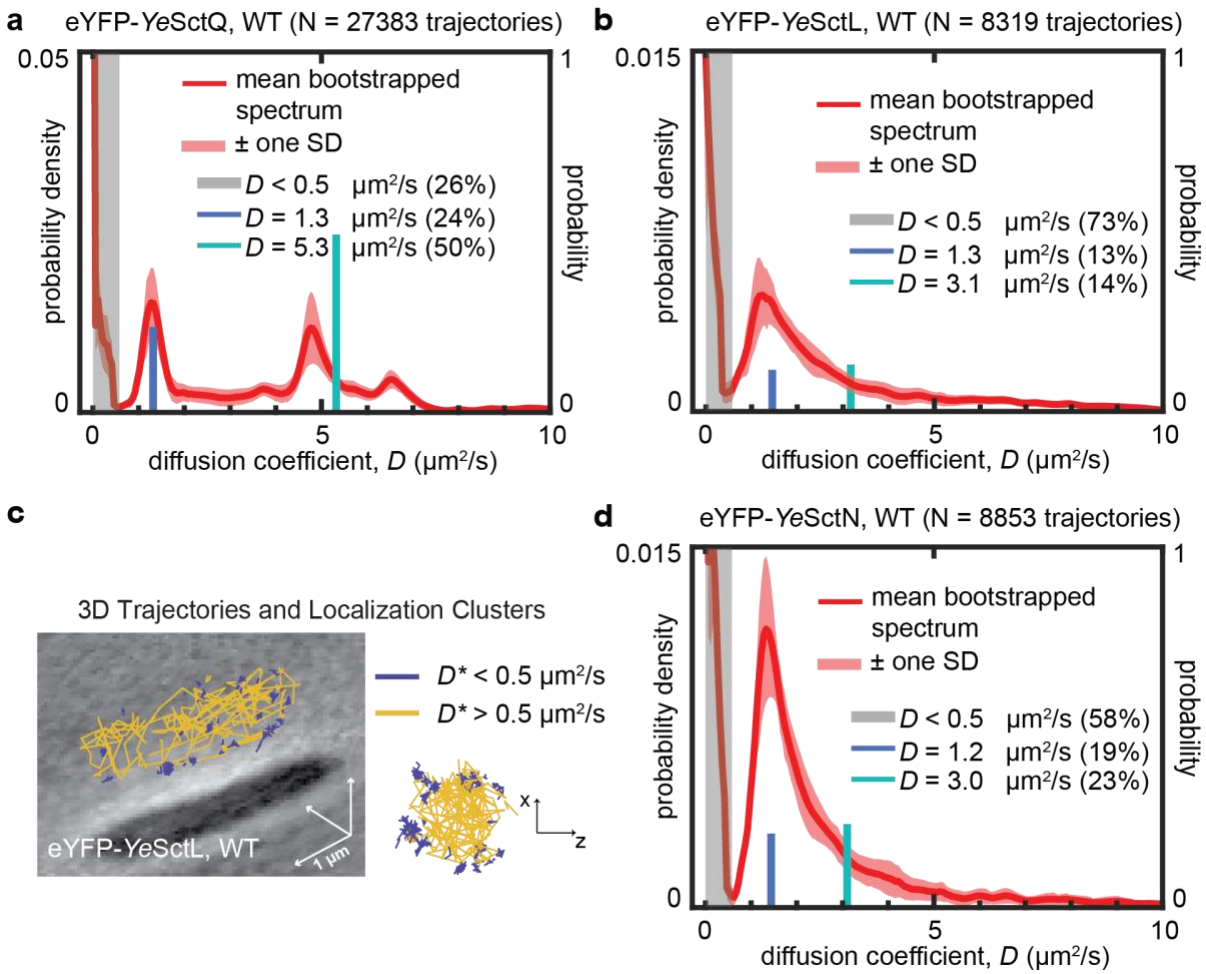

Figure 4.3: (a) Diffusion coefficient spectrum for *Ye*SctQ in wild-type cells. The experimentally measured distribution is best fitted with two states at *D* values of ~1.3 μm²/s and ~5.3 μm²/s (Fig. 4.4). (b) Diffusion coefficient spectrum for *Ye*SctL in wild-type cells. The experimentally measured distribution is best fitted with two states at *D* values of ~1.3 μm²/s and ~3.1 μm²/s (Fig. 4.7). (c) 3D trajectories for *Ye*SctL with *D** values of <0.5 μm²/s exhibit clustering at the membrane indicative of association with injectisomes. (d) Diffusion coefficient spectrum for *Ye*SctN in wild-type cells. The experimentally measured distribution is best fitted with two diffusive states at *D* values of ~1.2 μm²/s and ~3.0 μm²/s (Fig. 4.8).

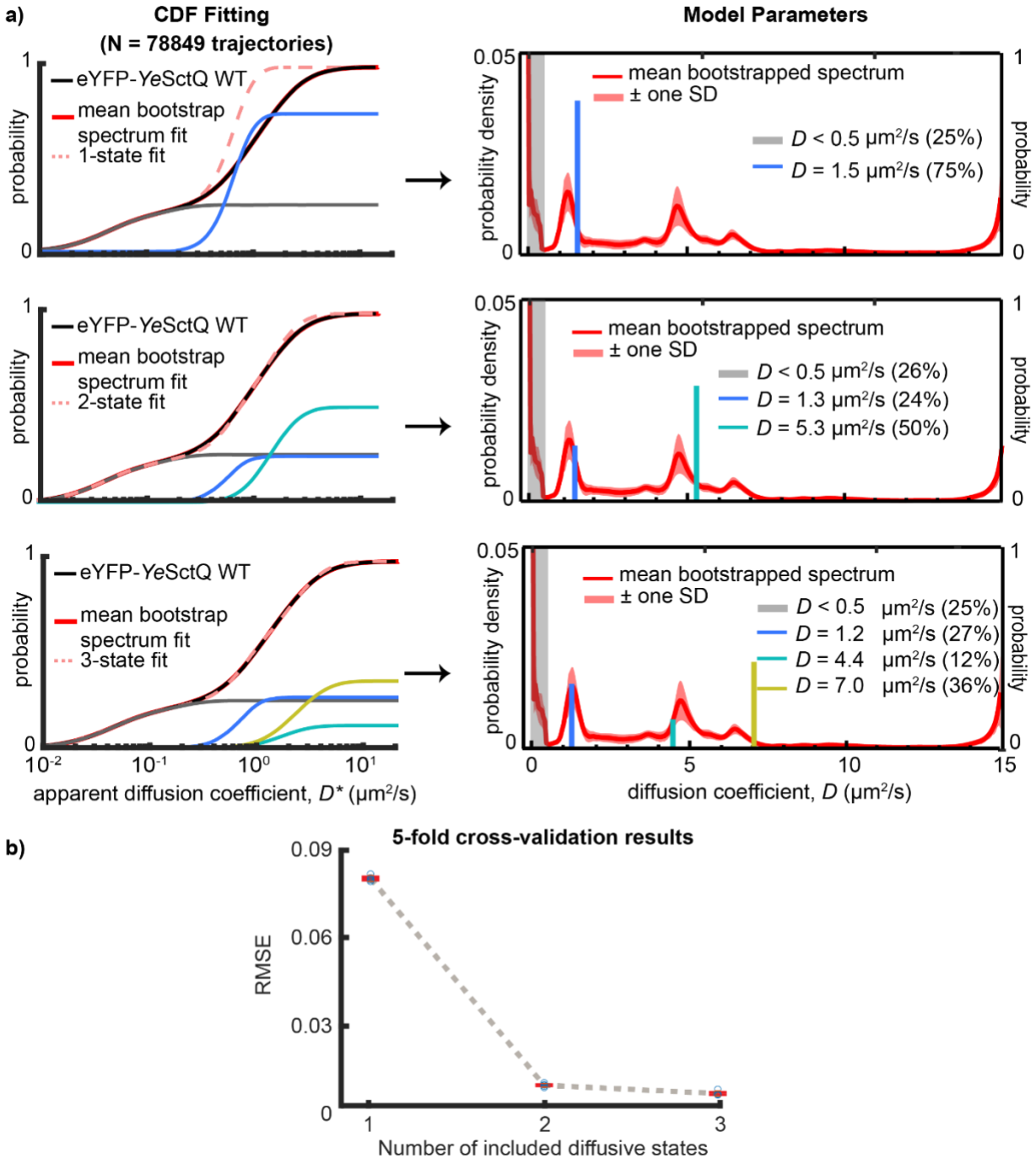

Figure 4.4: a) Fitting model comparison for eYFP-YeSctQ single-molecule tracking data. Individual rows correspond to 1-, 2-, and 3-state models, respectively. A 2-state model converged on D ~ 1.3 and 5.3 μm2 /s, which agrees well with peaks in the diffusion coefficient spectrum. b) The 2-state model is also supported by 5-fold cross-validation

analysis, which shows that a 2-state model does not overfit the data. A 3-state model could also have been chosen based on the cross-validation analysis alone. However, the 3-state model includes diffusive states at D ~ 7.0 µm2 /s with an overly large population fraction of 36%, which does not match the diffusion coefficient spectrum.

4.4.3 *Ye*SctL shares one diffusive state with *Ye*SctQ

Having reproduced our previously obtained results in the diffusion coefficient spectra, we next analyzed the intracellular diffusive behaviors of eYFP-labeled *Ye*SctL. We had previously tracked *Ye*SctL with PAmCherry1 but chose to switch the fluorescent label to eYFP for the present study due to its homogeneous diffusion behavior in *Y. enterocolitica* (**Fig. 4.1ab**). Again, we expressed eYFP-*Ye*SctL as a genomic replacement under the control of its native promotor. The eYFP-*Ye*SctL fusion protein was functional and not degraded (**Fig. 4.5 & 4.6**). Examination of the *Ye*SctL diffusion coefficient spectrum reveals a prominent spectral peak at D ~ 1.3 with a small shoulder to the right (**Fig. 4.3b**). The 3-state model for *Ye*SctL produced diffusive states at D ~ 1.3 (population fraction = 17%), 3.2 (population fraction = 3%), and 7.4 µm^2/s (population fraction = 7%). Given that the diffusion coefficient spectrum clearly indicates a prominent peak at D ~ 1.3 µm^2/s and does not include a peak at D ~ 0.9 or 7.4 µm^2/s, we turned to the 2-state model, whose fitted values agree better with the spectral peak positions. The choice of the 2-state model is also supported by 5-fold cross-validation analysis (**Fig. 4.7**).

The 2-state model provides two diffusive states for *Ye*SctL at D ~ 1.3 and 3.1 µm^2/s with corresponding population fractions of 13% and 14%, respectively (**Fig. 4.3b, Fig. 4.7**). These mobile states are in addition to an abundant (population fraction = 73%)

stationary, injectisome-bound population with $D < 0.5$ μm^2/s. Plotting the 3D trajectories using a $D^* = 0.5$ μm^2/s threshold clearly shows membrane-associated clusters representing the injectisome-bound population (**Fig. 4.3c**). The $D \sim 1.3$ μm^2/s diffusive state is consistent with our previous result obtained with PAmCherry-*Ye*SctL, based on which we concluded that both *Ye*SctQ and *Ye*SctL diffuse at the same slow rate as part of the same complex (83). The SctLQ interaction is also supported by a number of previous studies (44,46,109). On the other hand, the $D \sim 3.1$ μm^2/s *Ye*SctL diffusive state has not been previously observed. The presence of this additional state suggests another homo- or hetero-oligomeric complex containing *Ye*SctL. However, such a complex does not contain *Ye*SctQ, because a $D \sim 3.1$ μm^2/s diffusive state is not observed in the eYFP-*Ye*SctQ data.

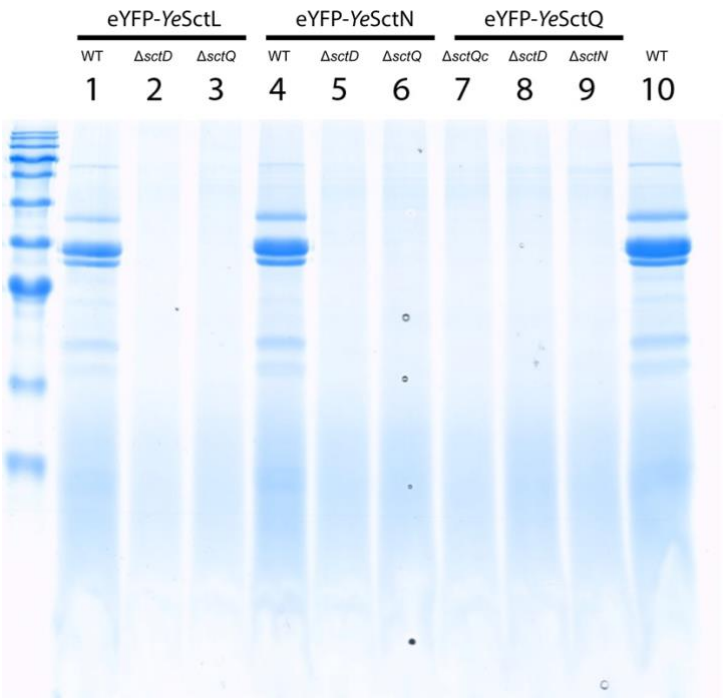

Figure 4.5: Secretion profile for mutant strains expressing eYFP-labeled cytosolic injectisome proteins in place of the unlabeled protein. eYFP-*Ye*SctL and eYFP-*Ye*SctN both show near WT levels of secretion in the WT background. The deletion of any one (cytosolic) injectisome protein abolishes secretion, as observed previously36. Also of note is the loss of secretion for eYFP-*Ye*SctQ in the Δ*Ye*SctQ$_C$ background, further confirming the functional relevance for *Ye*SctQ$_C$ for secretion. Data collected and analyzed by the lab of Dr. Andreas Diepold.

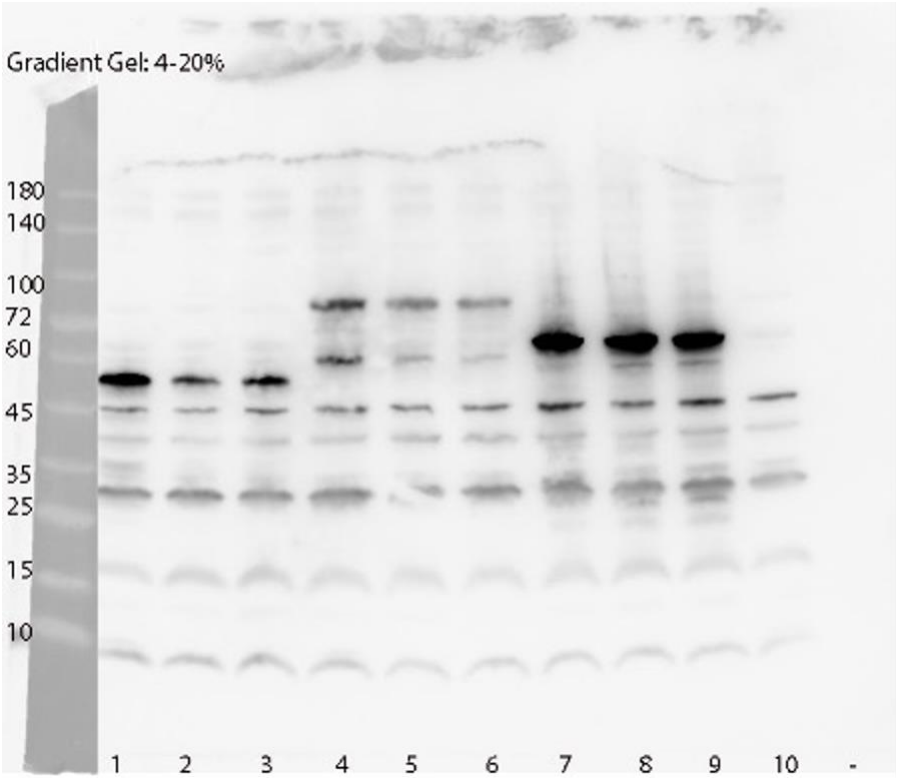

Figure 4.6: Expression levels and stability of the used fusion proteins. All fusions proteins seem to be expressed to a similar degree (YscQ > YscL > YscN, as expected) at the expected MW. No visible degradation is evident for eYFP-YeSctL, some degradation for eYFP-*Ye*SctN (~ 60 kDa, ~20- 40% of intensity of full-length band, and very weak degradation for eYFP-*Ye*SctQ (weak band at around 55-60 kDa). Data collected and analyzed by the lab of Dr. Andreas Diepold.

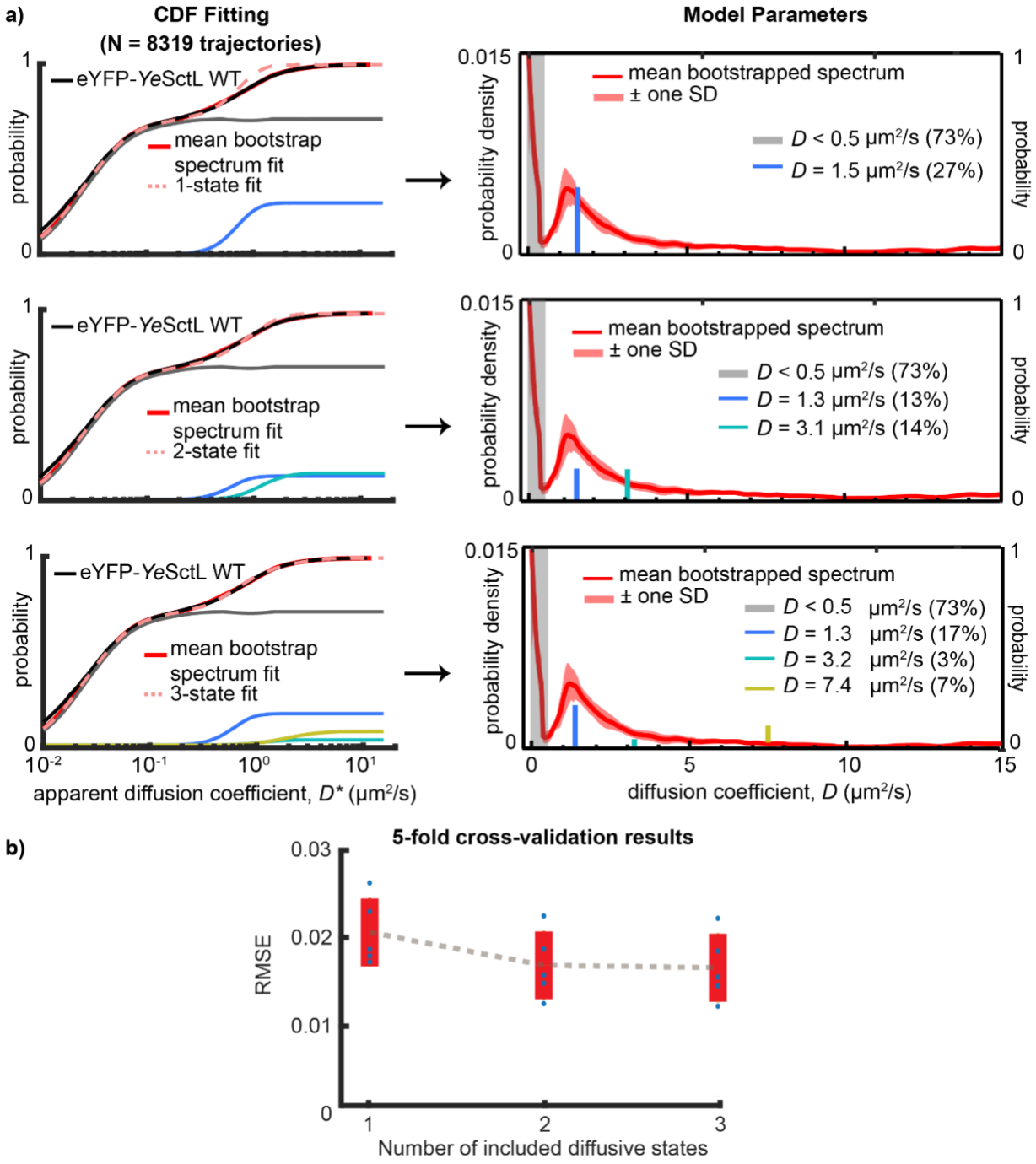

Figure 4.7: a) Fitting model comparison for eYFP-*Ye*SctL single-molecule tracking data. Individual rows correspond to 1-, 2-, and 3-state models, respectively. A 2-state model converged on D ~ 1.3 and 3.1μm^2/s, which agrees well with peaks in the diffusion coefficient spectrum. b) The 2-state model is also supported by 5-fold cross-validation

analysis, which shows that a 2-state model does not overfit the data. A 3-state model could also have been chosen based on the cross-validation analysis. However, the 3-state model includes diffusive states at $D \sim 7.4\ \mu m^2/s$, which does not match the diffusion coefficient spectrum.

4.4.4 *Ye*SctN and *Ye*SctL share two distinct diffusive states in *Y. enterocolitica*

There are numerous reports documenting interactions between SctN and SctL in both flagellar and virulence-associated T3SSs (119,120,50,121). SctL binding prevents SctN hexamerization prior to injectisome binding and thereby negatively regulates SctN's ATPase activity, which is the one of the most well-established functional roles of the cytosolic injectisome proteins (49,50,119,122–124). To detect the SctNL interaction in living cells, we expressed eYFP-*Ye*SctN as a genomic replacement under the control of its native promotor and acquired *Ye*SctN single-molecule trajectories in actively secreting cells. The eYFP-*Ye*SctN fusion protein was functional and only minimally degraded (**Fig. 4.5 & 4.6**). The resulting diffusion coefficient spectrum shows a very prominent spectral peak at $D \sim 1.3\ \mu m^2/s$ with, similar to the *Ye*SctL diffusion coefficient spectrum, a small shoulder to the right (**Fig. 4.3d**). Non-linear fitting with a 1-state model produces a diffusive state at $D \sim 1.7\ \mu m^2/s$. A 2-state models produces diffusive states at $D \sim 1.2$ and $3.0\ \mu m^2/s$ with corresponding population fractions of 19% and 23% (**Fig. 4.3d**, **Fig. 4.8a**), a result that better matches the diffusion coefficient spectrum. The 2-state model is also supported by cross-validation analysis (**Fig. 4.8b**). The robust emergence of a $D \sim 1.3\ \mu m^2/s$ diffusive state, a state that is also observed in YeSctL and YeSctQ measurements (in both the 2- and 3-state models) suggests the formation of a hetero-oligomeric *Ye*SctNLQ

complex. The 2-state model produces an intermediate diffusive state at $D \sim 3.0$ μm^2/s, and such a diffusive state was also observed for *Ye*SctL ($D \sim 3.1$ μm^2/s). We thus tentatively assign this diffusive state to a SctNL complex. While the SctNL interaction is supported by a number of previous studies, the population fraction of this state is low for both *Ye*SctN and *Ye*SctL in wild-type cells. We therefore concluded that establishing the existence of this state needed further experimental support.

We argued above that the oligomeric complex that gives rise to the $D \sim 3.0$ μm^2/s diffusive state does not contain *Ye*SctQ. It follows that deletion of *sctQ* should not affect the formation of that complex. To test this hypothesis, we acquired single-molecule trajectories for *Ye*SctN and *Ye*SctL in a Δ*sctQ* mutant background. Remarkably, the CDFs of apparent diffusion coefficients for these two proteins (i.e. the raw data) are essentially congruent in the Δ*sctQ* background, strongly indicating co-diffusive behavior (**Fig. 4.9a**). Plotting the 3D trajectories for eYFP-*Ye*SctN in the Δ*sctQ* background using a $D^* = 0.5$ μm^2/s threshold shows an absence of membrane-associated clusters (**Fig. 4.9b**). The same phenomenon is observed for eYFP-*Ye*SctL in the Δ*sctQ* background (data not shown), indicating that injectisome association is no longer possible in the absence of SctQ (46). The spectra of diffusion coefficients of these two proteins show the same trend: each spectrum contains a prominent peak at $D \sim 2.8$ μm^2/s. Notably, a peak at $D \sim 1.3$ μm^2/s, which we previously assigned to a *Ye*SctNLQ complex, is also absent. A 1-state non-linear fit converged on $D \sim 2.7$ μm^2/s in each case with corresponding population fractions of 93% for *Ye*SctL and 92% for *Ye*SctN, respectively (**Fig. 4.10, Fig. 4.11**).

Together, these results support the conclusion that *Ye*SctL and *Ye*SctN are able to form two distinct cytosolic complexes. One of these complexes contains *Ye*SctQ and the

other does not. Under our experimental conditions, the *Ye*SctNLQ complex diffuses at $D \sim 1.3\ \mu m^2/s$, while the *Ye*SctNL complex diffuses at $D \sim 2.8\ \mu m^2/s$. In wild-type cells, the presence of *Ye*SctQ leads to a smaller abundance of the *Ye*SctNL complex in favor of the *Ye*SctNLQ complex as well as injectisome-bound *Ye*SctL and *Ye*SctN.

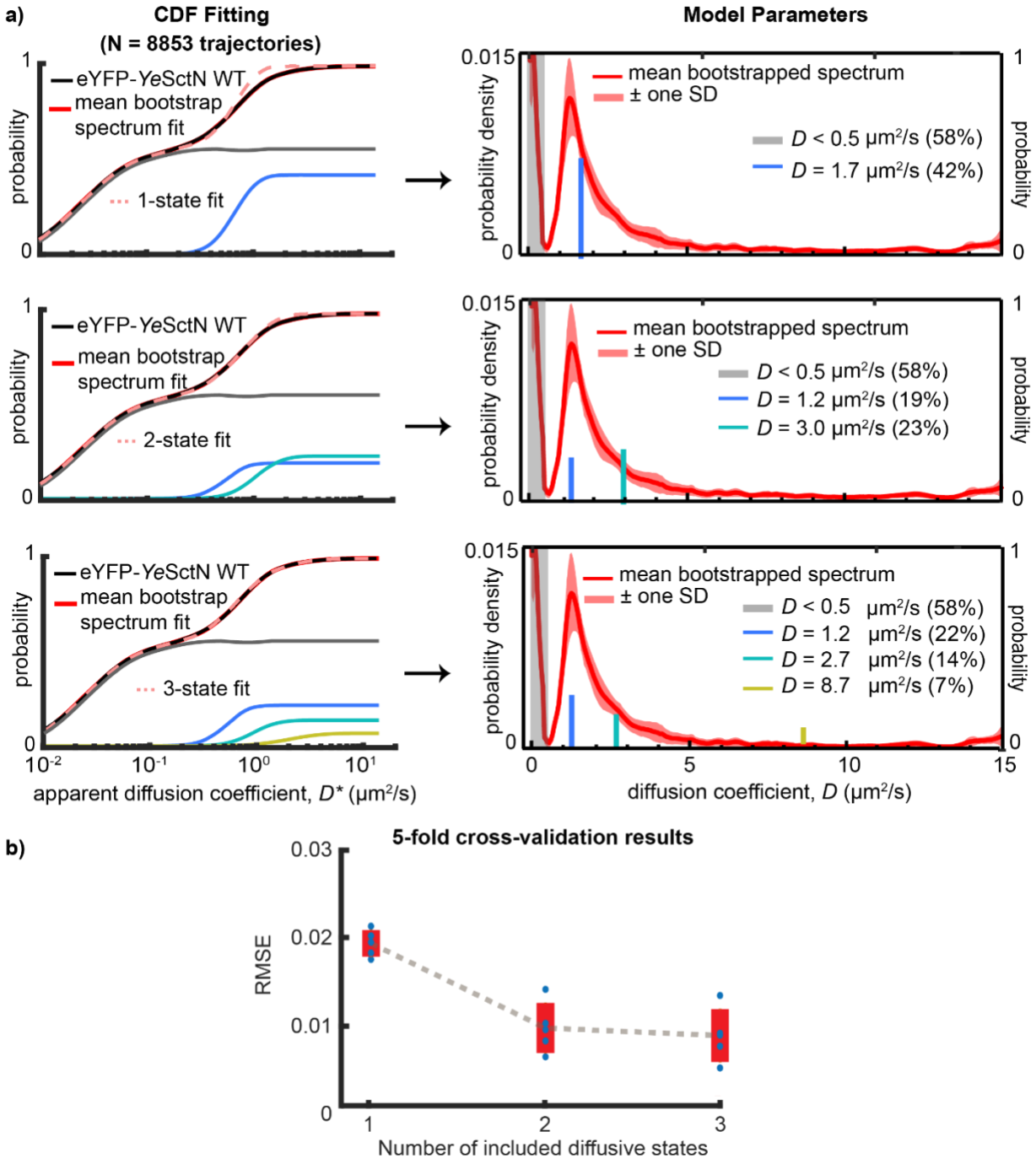

Figure 4.8: a) Fitting model comparison for eYFP-*Ye*SctN single-molecule tracking data. Individual rows correspond to 1-, 2-, and 3-state models, respectively. A 2-state model converged on D ~ 1.2 and 3.0 μm^2/s, which agrees well with peaks in the diffusion coefficient spectrum. b) The 2-state model is also supported by 5-fold cross-validation

analysis, which shows that a 2-state model does not overfit the data. A 3-state model could also have been chosen based on the cross-validation analysis. However, the 3-state model includes diffusive states at $D \sim 8.7$ μm²/s, which does not match the diffusion coefficient spectrum.

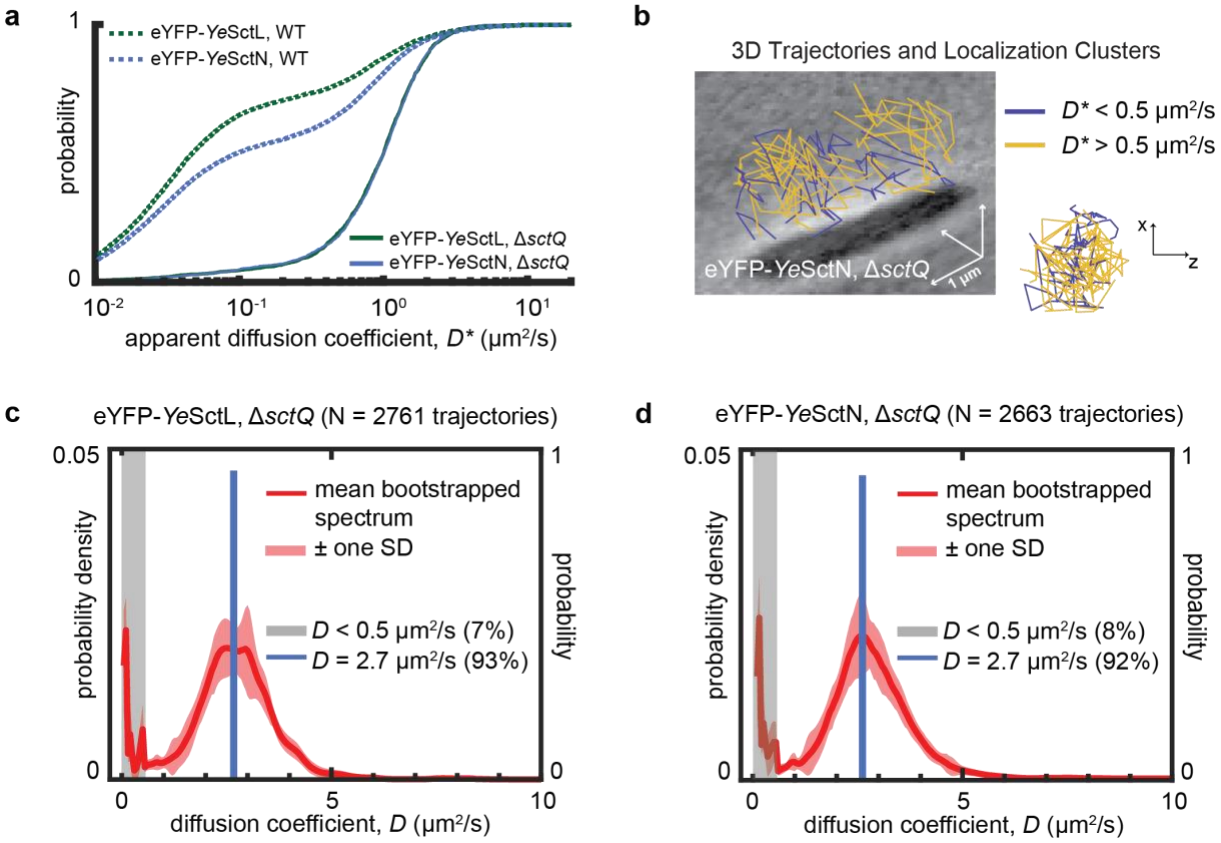

Figure 4.9: a) Experimentally measured CDFs of *Ye*SctL and *Ye*SctN apparent diffusion coefficients in wild type cells and Δ*sctQ* mutants. The *Ye*SctL and *Ye*SctN distributions overlay very well in the Δ*sctQ* background. b) 3D trajectories for *Ye*SctN in the Δ*sctQ* background show an absence of membrane-associated clusters for $D^* < 0.5$ μm²/s. c) and d) Diffusion coefficient spectra for *Ye*SctL and *Ye*SctN in the Δ*sctQ* background. In each

case, the experimentally measured distribution is best fit with one state at D = 2.7 μm^2/s (Figure 4.10 & 4.11).

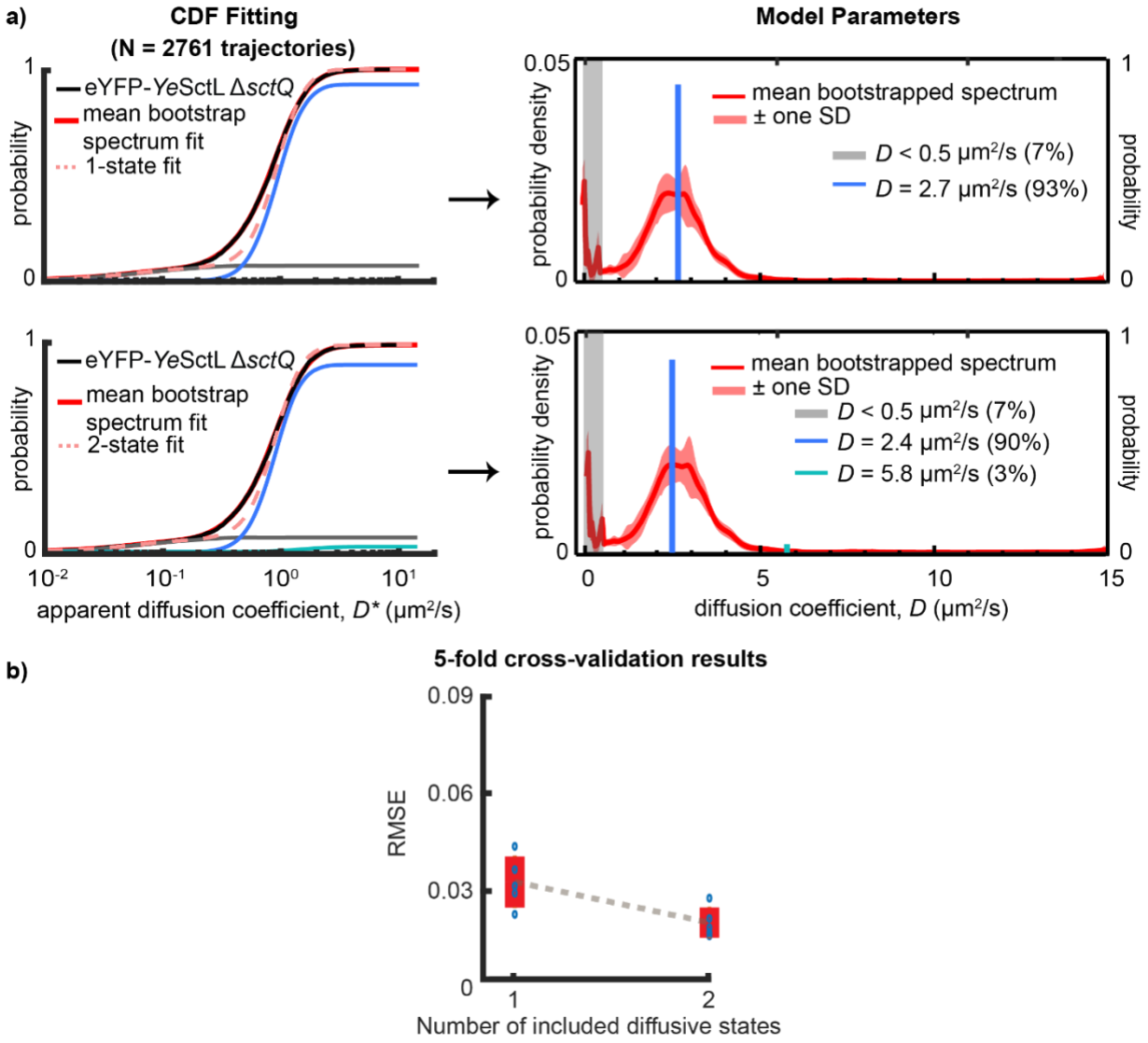

Figure 4.10: a) Fitting model comparison for eYFP-*Ye*SctL Δ*sctQ* single-molecule tracking data. Individual rows correspond to 1-and 2-state models, respectively. A 1-state model converged on D ~ 2.7 μm^2/s, which agrees well with peaks in the diffusion coefficient spectrum. b) The 1-state model is also supported by 5-fold cross-validation

analysis, which shows that a 1-state model does not overfit the data. A 2-state model could also have been chosen based on the cross-validation analysis. However, the 2-state model includes a diffusive state at $D \sim 5.8\ \mu m^2/s$ with a very small population fraction (3%) that does not agree with the diffusion coefficient spectrum. Thus, the 1-state model is favored over the 2-state model.

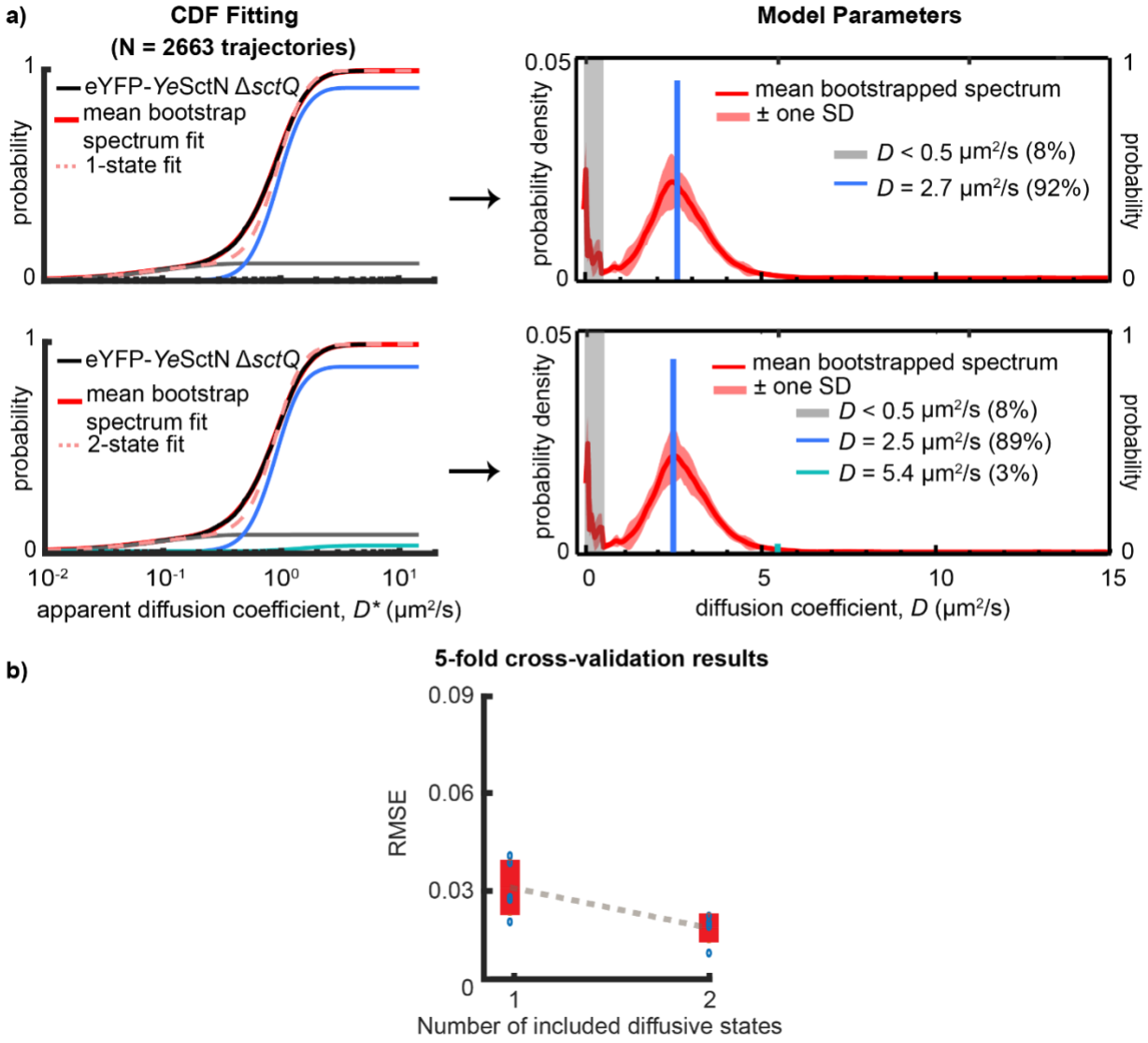

Figure 4.11: a) Fitting model comparison for eYFP-*Ye*SctN Δ*sctQ* single-molecule tracking data. Individual rows correspond to 1-and 2-state models, respectively. A 1-state model converged on D ~ 2.7 μm^2/s, which agrees well with peaks in the diffusion coefficient spectrum. b) The 1-state model is also supported by 5-fold cross-validation analysis, which shows that a 1-state model does not overfit the data. A 2-state model could also have been chosen based on the cross-validation analysis. However, the 2-state model includes a diffusive state at D ~ 5.4 μm^2/s with a very small population fraction (3%) that

does not agree with the diffusion coefficient spectrum. Thus, the 1-state model is favored over the 2-state model.

4.4.5 Complex formation among *Ye*SctQ, L, and N increases in the absence of injectisomes

SctD is the inner-membrane ring protein of the injectisome. SctD is required for proper injectisome assembly (19) and subsequent effector protein secretion. SctD provides a binding interface for SctK, which in turn allows SctQ, L, and N to bind to the injectisome (125). Thus, SctQ, L, and N are completely cytosolic in the Δ*sctD* background. To test whether the same diffusive states manifest in the absence of fully-assembled injectisomes, we acquired single-molecule trajectories in the Δ*sctD* mutant background for *Ye*SctQ, L, and N.

The *Ye*SctQ Δ*sctD* spectrum is very similar to the spectrum observed in wild-type cells. Two clear spectral peaks are evident at $D \sim 1.3$, and 4.5 $\mu m^2/s$ (**Fig. 4.12a, Fig. 4.13**). A 2-state non-linear fit produces diffusive states at $D \sim 1.2$ and 4.5 $\mu m^2/s$, a result that is in excellent agreement with the diffusive state observed in wild-type cells. These data show that the same complexes formed in wild-type cells can also form in the absence of injectisomes. Consistent with the loss of injectisome binding, the combined population fractions of the three distinct cytosolic diffusive states increase. The $D < 0.5$ $\mu m^2/s$ state(s) decreased from 24% in wild-type cells to 10% in the Δ*sctD* mutant. We note however that the 10% of slow diffusing molecules in the Δ*sctD* mutant do not localize to the membrane (**Fig. 4.2b**).

The *Ye*SctL and *Ye*SctN Δ*sctD* spectra both show a broad peak extending from *D* ~ 1 to 4 μm^2/s. A 2-state fit converges to diffusive states at *D* ~ 1.2 and 2.6 μm^2/s for *Ye*SctL and *D* ~ 1.0 and 2.5 μm^2/s for *Ye*SctN. Both of these results agree well with the diffusion coefficient spectra (**Fig. 4.12bc, Fig. 4.14, Fig. 4.15)** and are also in close agreement with the diffusive states observed for each protein in wild-type and Δ*sctQ* cells. Taken together, these results establish that the same complexes formed in wild-type cells can also form, in increased abundance, in the absence of injectisomes. In other words, the presence of injectisomes is not necessary for cytosolic complex formation.

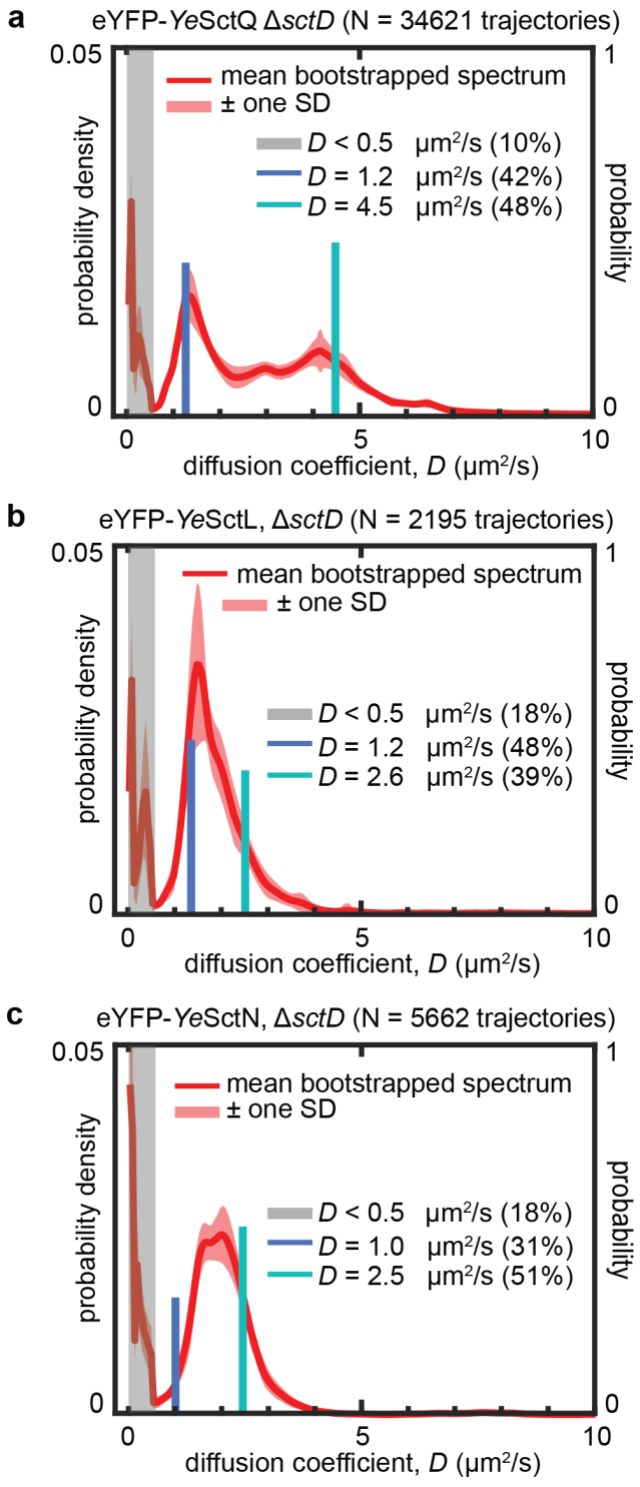

Figure 4.12: Removal of *Ye*SctD results in an increased abundance of native cytosolic complexes. a) *Ye*SctQ, b) *Ye*SctL, and c) *Ye*SctN all share a diffusive state around $D \sim 1.3$

μm^2/s, consistent with results obtained in wild type cells. Also consistent with wild type results is the presence of a diffusive state at D ~ 4.5 μm^2/s for *Ye*SctQ and the presence of a shared diffusive state for *Ye*SctL and *Ye*SctN at D ~ 2.5 μm^2/s.

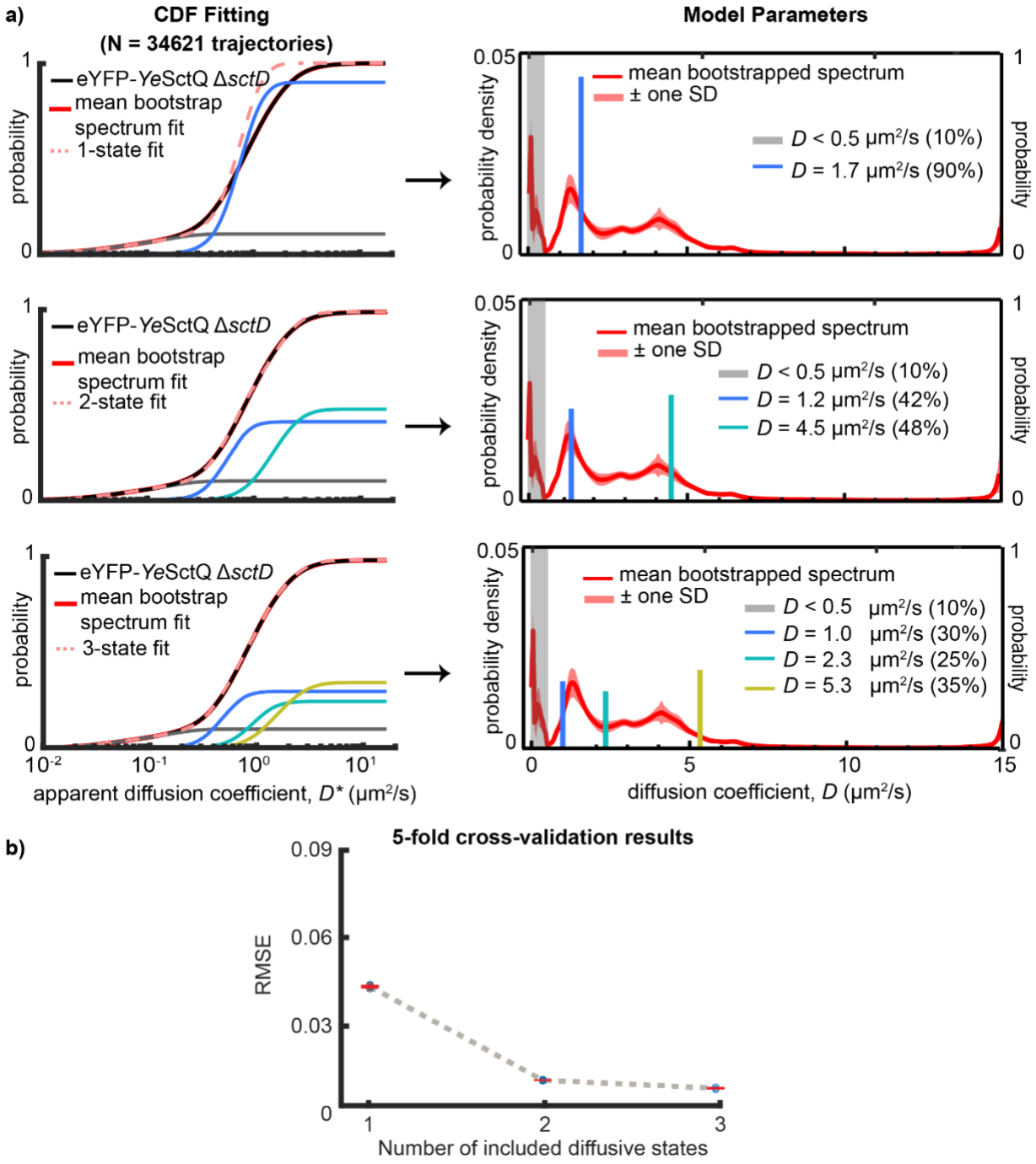

Figure 4.13: a) Fitting model comparison for eYFP-*Ye*SctQ Δ*sctD* single-molecule tracking data. Individual rows correspond to 1-, 2- and 3-state models, respectively. A 2-state model converged on D ~ 1.2 and 4.5 μm^2/s, which agrees well with peaks in the diffusion coefficient spectrum. b) The 2-state model is also supported by 5-fold cross-

validation analysis, which shows that a 2-state model does not overfit the data. A 3-state model could also have been chosen based on the cross-validation analysis. However, the 3-state model includes a diffusive state at $D \sim 2.3$ μm^2/s that does not agree with the diffusion coefficient spectrum. Thus, the 2-state model is favored over the 3-state model.

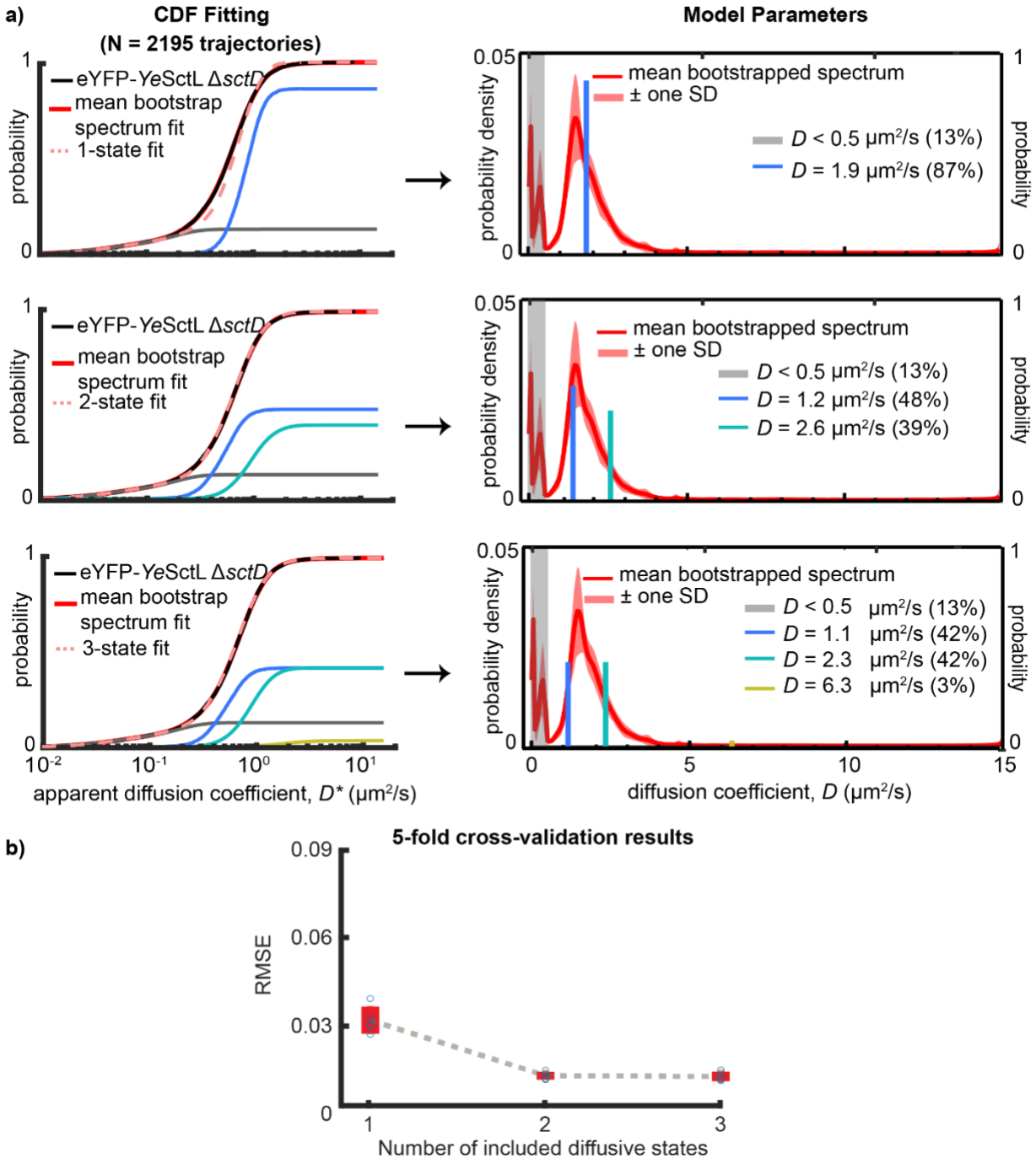

Figure 4.14: a) Fitting model comparison for eYFP-*Ye*SctL Δ*sctD* single-molecule tracking data. Individual rows correspond to 1-, 2-, and 3-state models, respectively. A 2-state model converged on $D \sim 1.2$ and 2.6 μm^2/s, which agrees well with peaks and spectral densities in the diffusion coefficient spectrum. b) The 2-state model is also supported by

5-fold cross-validation analysis, which shows that a 2-state model does not overfit the data. A 3-state model could have also been chosen based on the cross-validation analysis. However, the 3-state model includes a diffusive state at $D \sim 6.3\ \mu m^2/s$ with a very small population fraction (3%) that does not agree with the diffusion coefficient spectrum. Thus, the 2-state model is favored over the 3-state model.

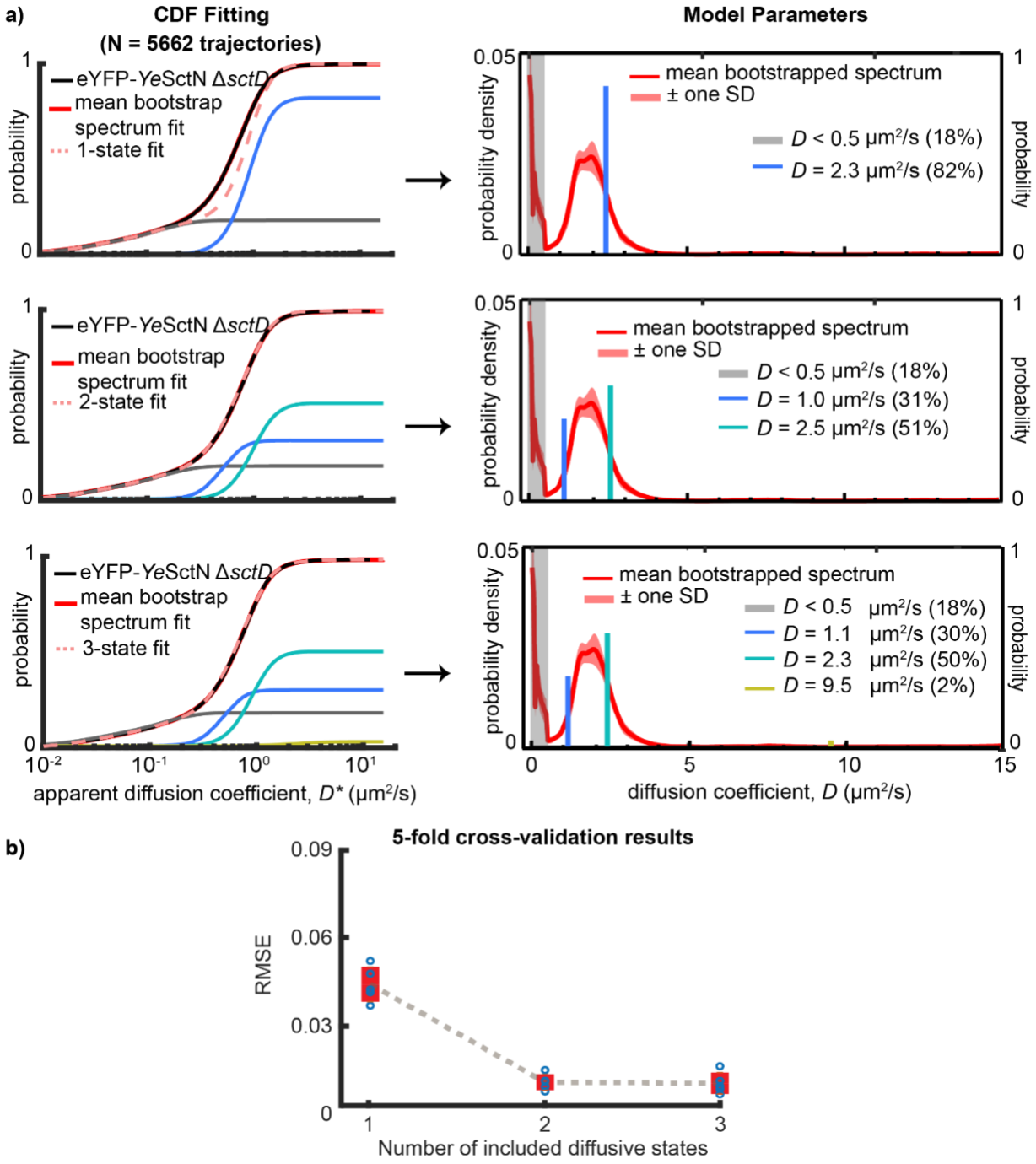

Figure 4.15: a) Fitting model comparison for eYFP-*Ye*SctN Δ*sctD* single-molecule tracking data. Individual rows correspond to 1-, 2-, and 3-state models, respectively. A 2-state model converged on D ~ 1.0 and 2.5 μm^2/s, which agrees well with peaks in the diffusion coefficient spectrum. b) The 2-state model is also supported by 5-fold cross-

validation analysis, which shows that a 2-state model does not overfit the data. A 3-state model could have also been chosen based on the cross-validation analysis. However, the 3-state model includes a diffusive state at $D \sim 9.5$ μm^2/s with a very small population fraction (2%) that does not agree with the diffusion coefficient spectrum. Thus, the 2-state model is favored over the 3-state model.

4.5 DISCUSSION

Determining the functional role(s) of the cytosolic injectisome proteins has been challenging, because our understanding of their interactions at the injectisome and in the cytosol remains limited. Here, we present live-cell single-molecule tracking data that suggest the formation of distinct, freely-diffusing protein complexes. One complex containing *Ye*SctN and *Ye*SctL, but not *Ye*SctQ, diffuses at $D \sim 2.8$ μm^2/s. A larger complex containing *Ye*SctN, L, and Q diffuses at $D \sim 1.3$ μm^2/s. These complexes form robustly in wild-type and mutant cells that are unable to fully assemble T3SS injectisomes (**Fig. 4.16a**).

A *Ye*SctNL complex is consistent with a proposed regulatory mechanism of *Ye*SctL. SctL is thought to negatively regulate ATPase activity in both bacterial flagellar motors and injectisomes by preventing SctN hexamerization in the cytosol(50). Recent cryo-EM data on purified proteins show that the extreme N-terminus of FliI (the flagellar homologue of SctN) contains a charged region to which FliH (the flagellar homologue of SctL) can bind(126). Abolishing this interaction through point mutations results in a higher propensity of ATPase hexamerization *in vitro*, as well as decreased cell growth and cell motility phenotypes. Further, the SctNL (FliIH) complex has been consistently observed

as a heterotrimer using *in vitro* biochemical and biophysical approaches (53,105,109,119,123,124,127), with a dimer of SctL (FliH) binding to the ATPase monomer SctN (FliI). Based on these results, we posit that the $D \sim 2.8\ \mu m^2/s$ diffusive state that we observe both wild-type and Δ*sctQ* mutants and the $D \sim 2.5\ \mu m^2/s$ diffusive state that we observe in Δ*sctD* mutants is likely due to a *Ye*SctN$_1$:*Ye*SctL$_2$ heterotrimer. In wild-type cells, 19% of *Ye*SctN and 13% of *Ye*SctL contribute to this heterotrimer population. In Δ*sctQ* and Δ*sctD* mutants, the abundance of this heterotrimer increases substantially (93% and 92% for *Ye*SctN and *Ye*SctL, respectively, for Δ*sctQ* and 31% and 48% for *Ye*SctN and *Ye*SctL, respectively, for Δ*sctD*).

A *Ye*SctNLQ complex is consistent with cytosolic pod complexes containing SctN, L, Q, and K. Sub-tomogram averages in *Shigella* and *Salmonella* suggest six distinct pods composed of SctK, Q, and L that seem to cradle the hexameric ATPase at the injectisome(38,104). Subsequent work has focused on reconstituting such pod complexes *in vitro*. *Se*SctNLQ(109) and *Sf*SctLQK complexes(44) were successfully isolated, and the estimated size and shapes of these complexes were qualitatively consistent with cryo-ET pod densities. Based on these results, we speculate that the $D \sim 1.3\ \mu m^2/s$ diffusive state we observe in wild-type cells and in Δ*sctD* mutants is likely not just a SctNLQ complex, but in fact a SctNLQK heterooligomer. This claim is substantiated by the loss of the $D \sim 1.3\ \mu m^2/s$ diffusive state when we track eYFP-*Ye*SctL and eYFP-*Ye*SctN proteins in the Δ*sctQ* mutant. Whether a $D \sim 1.3\ \mu m^2/s$ diffusive state also manifests for SctK in wild-type cells and in Δ*sctD* mutants remains to be determined.

The molecular composition of the $D \sim 4.8\ \mu m^2/s$ diffusive state observed for eYFP-*Ye*SctQ could be consistent with a *Ye*SctQK complex. In our previous work(83), we

assigned this intermediate diffusive state to a *Ye*SctQ homo-oligomer that also contains *Ye*SctQ$_C$, the alternatively expressed C-terminal fragment of *Ye*SctQ. This assignment was made based on tracking eYFP-*Ye*SctQ proteins in a Δ*pYV* mutant, which does not express any other T3SS proteins. In that mutant, ~80% of the tracked *Ye*SctQ proteins diffused at $D \sim 3.6$ μm^2/s, which is notably smaller than the $D \sim 4.8$ μm^2/s diffusive state that we observe in wild type cells. This discrepancy indicates that the $D \sim 3.6$ and 4.8 μm^2/s diffusive states likely originate from different complexes. The spectrum of eYFP-*Ye*SctQ diffusion coefficients in a Δ*pYV* background indeed show a prominent spectral peak ranging from $D \sim 3$ to 4.5 μm^2/s, but little spectral density at $D \sim 4.8$ μm^2/s (**Fig. 4.17**). *In vitro* light-scattering measurements on *Sf*SctQ-containing complexes(44) suggested the ability of *Sf*SctQ to adopt different conformational states depending on its interacting partner. A small, globular complex (hydrodynamic diameter = 10.2 nm) was observed upon co-expression and co-purification of *Sf*SctQ and K, while a larger complex (hydrodynamic diameter = 13.4 nm) was observed for co-purified *Sf*SctQ and Q$_C$. Based on these results, we speculate that the $D \sim 4.8$ μm^2/s diffusive state in wild type cells and in Δ*sctD* mutants is due to a *Ye*SctQK complex. In wild-type cells, we observe very little spectral density in the $D = 3.5$ - 4 μm^2/s range, which indicates that a *Ye*SctQQ$_C$ complex is not abundantly present. On the other hand, if injectisomes are not assembled, as in the Δ*sctD* mutants, we do observe spectral density in the $D = 3.5$ - 4 μm^2/s range. These results indicate that the presence of injectisomes and thus the ability of *Ye*SctQ to bind to and unbind from injectisomes affects the type of complexes that *Ye*SctQ forms in abundance. Determining the compositions and stoichiometries of *Ye*SctQ-containing complexes in specific deletion mutants will be the subject of future work.

Complex formation among cytosolic injectisome proteins may provide the foundation for functional regulation of type 3 secretion. Specifically, a pool of freely diffusing complexes of injectisome proteins that are available for injectisomes binding would provide a mechanism for timely reactivation of secretion. Recent work has shown that an extracellular pH drop from 7 to 4 results in disassociation of *Ye*SctN, L, Q, and K from injectisomes, an effect that correlated with loss of secretion(128). Notably, the intracellular pH decreased only slightly (pH=7→6.3) in acidic extracellular environments. Reverting the external pH from 4 to 7 reverses these effects. Injectisome binding is restored within ~10 min and effector protein secretion resumes shortly thereafter. The results reported here lend further support for a secretion reactivation model relying on cytosolic complexes of injectisome proteins: in Δ*sctQ* and in Δ*sctD* mutants, in which injectisome binding is not possible, we observe clear signatures of the same diffusive states in wild-type cells that we assign to distinct *Ye*SctNLQ, *Ye*SctNL, and *Ye*SctQK complexes.

Identifying which cytosolic injectisome proteins assemble into freely-diffusing complexes in a native biological context can help determine how these proteins bind and unbind dynamically to and from injectisomes. While we have identified the compositions of three prominent complexes, the exact stoichiometries and functions of these complexes remains to be determined. Previous work by others has shown that the cytosolic diffusion rate of proteins scales with molecular weight, but not according to the Stokes-Einstein equation, which stipulates that $D \propto (\text{Molecular Weight})^{-1/3}$. Instead, a $D \propto (\text{Molecular Weight})^{-2/3}$ scaling has been observed by others independently for different proteins (129–132). Here, we obtain a $D \propto (\text{Molecular Weight})^{-0.71}$ scaling when plotting the measured diffusion coefficients against the molecular weights of monomeric eYFP and the protein

complexes discussed above (likely stoichiometries were estimated based on current biophysical models of the cytosolic injectisome protein complexes) **(Fig. 4.16b**) (44,121). This analysis allows us to distinguish between oligomerization models that differ in terms of protein complex stoichiometry. For example, earlier studies suggested that SctK, Q, Q_C, L, and possibly N, form a high molecular weight (>1MDa) complex, termed the sorting platform, that sequentially interacts with different secretion substrates and their chaperones(51,133). However, efforts aimed at reconstituting sorting platform complexes *in vitro* have not been reported. The analysis in Figure 5b suggests that the majority of cytosolic injectisome proteins exist as part of smaller complexes with molecular weights of less than 500 kDa. Our combined results thus narrow down the possible stoichiometries of cytosolic injectisome protein complexes that can form prior to injectisome binding. Future work will need to determine the molecular structures of the complexes identified here and how their individual abundances change in response to environmental signals that induce T3SS assembly and activate secretion. However, as shown in this study, careful attention must be paid to the effect of the osmotic environment on cytosolic protein diffusion rates. Such efforts will help determine how the cytosolic injectisome proteins interact in the cytosol and when bound to injectisomes and how their dynamic exchange at injectisomes contributes to functional regulation of secretion.

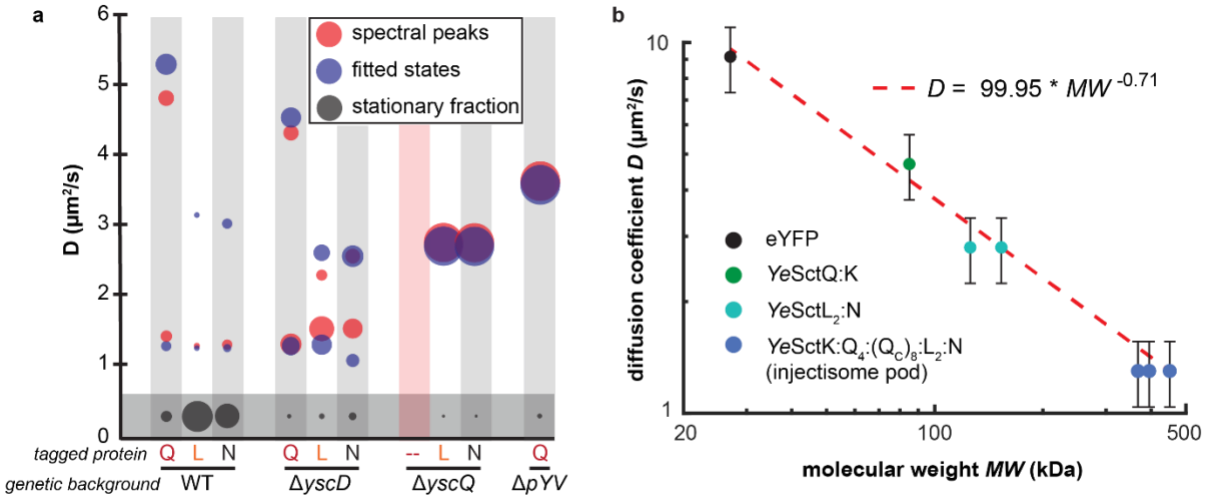

Figure 4.16: a) Comparison of spectral peaks and fitted diffusive states for each indicated eYFP-labeled injectisome protein tracked in different genetic backgrounds. The relative population fraction for each peak/diffusive state is represented by the size of the circle. Spectra for eYFP-*Ye*SctN and eYFP-*Ye*SctL in the Δ*sctD* mutant do not contain prominent spectral peaks corresponding to the fitted $D \sim 2.8$ μm^2/s diffusive state, whereas prominent spectral peaks are resolved in the Δ*sctQ* mutant. The fitted diffusive states for $D < 0.5$ μm^2/s in the Δ*sctD*, Δ*sctQ*, and Δ*pYV* mutants do not localize to the membrane. b) Measured diffusion coefficients plotted against the molecular weights of the indicated protein complexes. The straight line corresponds to the least squares fit using the model shown.

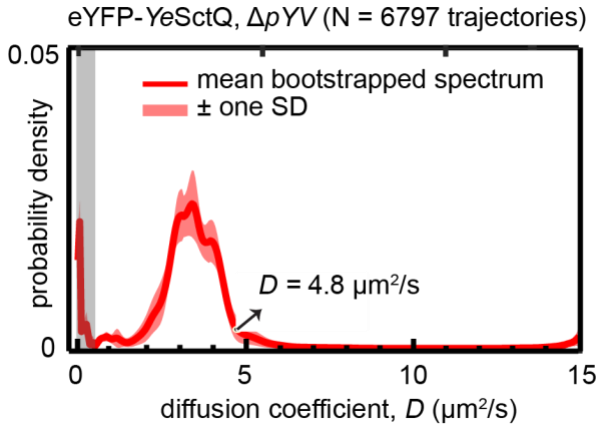

Figure 4.17: Diffusion coefficient spectrum for eYFP-*Ye*SctQ in the Δ*pYV* mutant showing spectral density between $D \sim 3$ to 5 $\mu m^2/s$. Δ*pYV* mutants do not express any other T3SS proteins.

4.6 MATERIALS AND METHODS

4.6.1 Bacterial Strains

Yersinia enterocolitica strains expressing fluorescent fusion proteins were generated by allelic exchange as previously described (19,134). Mutator plasmids containing 250-500bp flanking regions, the coding sequence for eYFP, and a 13 amino acid flexible linker region between the fluorescent and target protein were introduced into *E. coli* SM10 λpir for conjugation with *Y. enterocolitica* pIML421asd. After sucrose counter-selection for the second allelic exchange event, colonies were grown overnight in BHI media (Sigma Aldrich, St. Louis, Missouri) with nalidixic acid (Sigma Aldrich) [35 µg/mL] and 2,6-diaminopimelic acid (Chem Impex International, Wood Dale, Illinois) [80 µg/mL]. PCR screening was then performed to confirm target insertion and constructs were confirmed by sequencing (GeneWiz, South Plainfield, New Jersey).

4.6.2 Cell Culture

Yersinia enterocolitica strains were inoculated from freezer stock one day prior to imaging and grown overnight in BHI media containing nalidixic acid (NaI) [35 µg/mL] and 2,6-diaminopimelic acid (dap) [80 µg/mL] at 28°C with shaking. On the day of imaging, 250 µL of overnight culture were transferred into fresh BHI media containing NaI, and dap and grown at 28°C with shaking for one hour. Glycerol [4 mg/mL], $MgCl_2$ [20 mM] and EDTA [5 mM] were then added to culture medium and transferred to a 37°C water bath with shaking for 3 hours to induce expression of the *yop* regulon and ensure secretion activation, respectively. Cells were pelleted by centrifugation at 5000x g for 3 minutes and resuspended three times into fresh M2G (4.9 mM Na_2HPO_4, 3.1 mM KH_2PO_4, 7.5 mM NH_4Cl, 0.5 mM $MgSO_4$, 10 µM $FeSO_4$ (EDTA chelate; Sigma), 0.5 mM

$CaCl_2$) with 0.2% glucose as the sole carbon source. After the three washes, the remaining pellet was re-suspended in M2G, dap, $MgCl_2$, glycerol, and EDTA/$CaCl_2$. Cells were finally plated on 1.5 – 2% agarose pads in M2G containing dap, glycerol, and $MgCl_2$.

4.6.3 Secretion assay and protein analysis

Cultures were inoculated to an optical density at 600 nm of 0.12 from stationary overnight cultures. After 1.5 h of growth at 28°C, induction of the *yop* regulon was performed by shifting the culture to 37°C. Cultures were further incubated at 37°C for 3 hours. Bacteria from 2 ml culture were collected (15,000 g, 10 min, 4°C). 1.8 ml of supernatant was mixed with 0.2 ml trichloroacetic acid (final concentration 10%) and incubated over night at 4°C to precipitate proteins within the supernatant. Proteins were collected (15,000 g, 15 min, 4°C) and washed twice with ice-cold acetone (15,000 g, 5 min, 4°C). The pellet was dried at room temperature for 1 h. The pellet was resuspended in SDS-PAGE loading buffer (SDS (2% w/v), Tris-HCl (0.1 M), glycerol (10% w/v), DTT (0.05 M), pH = 6.8) and 0.6 OD units (1 OD unit is the equivalent of 1 ml culture at an OD_{600} of 1) in 15 µl were used for the SDS-PAGE gel analysis. For Western Blot analysis of the total cellular sample, the collected bacterial pellet was prepared for Western blot analysis by directly normalizing to 0.3 OD units in 15 µl loading buffer. All samples were heated for 10 min at 99°C before loading. Proteins were separated by SDS-PAGE on 15% acrylamide gels. For visualization, the gels were stained with InstantBlue (Expedeon). For immunoblots, the proteins were blotted on a nitrocellulose membrane. Detection of the eYFP-tag was performed by using primary rabbit antibodies against GFP (1:5000) (Invitrogen A6455, lot #1853896) and secondary anti-rabbit antibodies conjugated with

horseradish peroxidase (1:10,000) (Sigma A8275). The immunoblot was visualized using ECL chemiluminescence substrate (Pierce) on a LAS-4000 Luminescence Image Analyzer (Fujifilm).

4.6.4 Single-Molecule Super-Resolution Fluorescence Imaging

Image data were acquired on a custom-built inverted fluorescence microscope based on the RM21 platform (Mad City Labs, Inc, Madison, Wisconsin), as previously described (83,116). Immersion oil was placed between the objective lens (UPLSAPO 100 1.4 NA) and the glass cover slip (VWR, Radnor, Pennsylvania, #1.5, 22 mm 22 mm). Single-molecule images were obtained by utilizing eYFP photoblinking(135). A 514 nm laser (Coherent, Santa Clara, California, Genesis MX514 MTM) was used for excitation of eYFP (B350 W cm2). Zero order quarter wave plates (Thorlabs, Newton, New Jersey, WPQ05M-405, WPQ05M-514, WPQ05M-561) were used to circularly polarize all excitation lasers, and the spectral profile of the 514 nm laser was filtered using a bandpass filter (Chroma, Bellows Falls, Vermont, ET510/10 bp). Fluorescence emission from both eYFP was passed through a shared filter set (Semrock, Rochester, New York, LP02-514RU-25, Semrock NF03-561E-25, and Chroma ET700SP-2P8). The emission path contains a wavelength specific dielectric phase mask (Double Helix, LLC, Boulder, Colorado) that is placed in the Fourier plane of the microscope to generate a double-helix point-spread-function(78,136) (DHPSF). The fluorescence signals is detected on a sCMOS cameras (Hamamatsu, Bridgewater, New Jersey, ORCA-Flash 4.0 V2). Up to 20 000 frames are collected per field-of-view with an exposure time of 25 ms. A flip-mirror in the emission pathway enables toggling the microscope between fluorescence imaging and

phase contrast imaging modes without having to change the objective lens of the microscope.

4.6.5 Data Processing

All fluorescence images were processed in MATLAB using a modified version of the easyDHPSF code (82,83,116). To extract 3D localizations, fluorescence intensity from single-molecule emitters was fit to a double-Gaussian PSF model with maximum likelihood estimation. A median filter with a time window of 10 frames was used for background subtraction.

For each field-of-view, cell outlines were generated based on the phase contrast images using the open-source software OUFTI (84). Single-molecule localizations were then overlayed and aligned with the cell outlines. To ensure that single-molecule localizations align well with the corresponding cell, the cell outlines are registered to the fluorescence data by a two-step 2D affine transformation using the 'cp2tform' function in MATLAB. Five control points were manually selected based on the position of the cell poles of single-molecule localization data, which generated an initial transformation that allow for the removal of any cell containing less than 10 localizations. The center of mass for all the remining cells were then used to create a second, larger set of control pairs to compute the final transformation function. Only localizations within cell outlines were considered for further analysis.

4.6.6 Single-molecule Tracking Analysis

3D single-molecule localizations were filtered (**Figure 3.1**) and linked into trajectories with a distance threshold of less than 2.5 μm between subsequent localizations. If two or more localizations were present in the cell at the same time, the trajectory was not considered for further analysis to prevent the incorporation of two or more molecules into the same trajectory. Any trajectories containing less than 4 localizations were also not considered for further analysis.

Trajectory information was then used to calculate the Mean Square Displacement (MSD)

$$MSD = \frac{1}{N}\sum_{n=2}^{N}\left(x_n - x_{n-1}\right)^2$$

where N is the total number of time points and x_n is the 3D position of the molecule at time point n. The apparent diffusion coefficient, D^*, of a given molecule trajectory was then computed as

$$D^* = \frac{MSD}{2 \cdot m \cdot \Delta t}$$

where $m=3$ is the dimensionality of the data and Δt is the camera exposure time. Δt=25 ms for all experiments reported here.

4.6.7 Monte Carlo Simulations

To resolve the unconfined diffusion coefficients of distinct molecular complexes in living cells based on the experimentally measured distribution of apparent diffusion coefficients, we simulated confined Brownian motion trajectories inside a cylindrical

volume (radius = 0.4 μm, length = 5 μm). We added to our existing pool (116) of simulated diffusion coefficients to improve data fitting (see next section). Noisy, motion-blurred single-molecule images mimicking the raw experimental data were simulated for confined (in the cylindrical volume) single-molecules with defined Brownian diffusion coefficients. These images were then processed and linked into trajectories in the same manner as our experimental data. The resulting simulated CDFs account for confinement effects of the bacterial cell volume, signal integration over the camera exposure time, and experimentally calibrated signal-to-noise levels. Analyzing the simulation data in the same manner as experimental data ensures that static and dynamic localization errors (116) are accurately modeled for our data fitting routine.

4.6.8 Data Fitting

Experimental distributions of apparent diffusion coefficients were fit using a linear combination of simulated CDFs, where each CDF corresponds to a single diffusive state described by a single diffusion coefficient. The coefficients of the best fitting linear combination were determined using the Matlab function lsqlin() with the trust-region-reflective algorithm (The MathWorks, Inc, Natick, MA). The resulting linear coefficients can be displayed as a spectrum of diffusion coefficients that manifest for a tracked protein in living cells. To establish the robustness of individual peaks in the diffusion coefficient spectrum, we resampled the raw data using bootstrapping (N = 100). We then average these bootstrapped spectra and used the resulting linear coefficients to fit the experimental data (mean bootstrapped spectrum fit). The mean bootstrapped spectrum provides us with an initial estimate of the diffusion coefficients of prominent diffusive states and their relative

population fractions, estimated by determining the peak maxima and the area under the spectral peaks, respectively. The so-obtained parameters are then used as input parameters for non-linear fitting using the particleswarm() function in MATLAB. This approach allowed us to estimate the diffusion coefficients and population fractions of distinct diffusive states that manifest in living cells. In the case where the spectrum contains overlapping peaks or feature-less non-zero spectral density, we choose one or multiple diffusive states to initialize the non-linear fitting process. This approach results in the *n*-state (n = 1-4) fitting models shown in Supplementary Figs. 3-10, which are then evaluated using 5-fold cross-validation and qualitative agreement with the diffusion coefficient spectrum obtained by linear fitting.

Chapter 5
INJECTISOME BOUND-TIME ANALYSIS OF THE SORTING PLATFORM AND ATPASE

5.1 INTRODUCTION

If and how the proposed complexes containing the *Y. enterocolitica* sorting platform and ATPase (see Chapter 4) contribute to the functional mechanism of secretion remains unknown. A shuttling mechanism of effector proteins to the injectisome involving complexes containing *Ye*SctQ has been previously proposed (45,46). Specifically, dynamic binding and unbinding of *Ye*SctQ with injectisomes has been observed using fluorescence recovery after photobleaching (FRAP). After photobleaching of injectisome locations, the fluorescence intensity of PAmCherry-labeled *Ye*SctQ under secretion-ON conditions recovered 50% of its pre-bleach intensity ($t_{1/2}$) in ~68 ± 8 seconds. However, under non-secretion conditions, the $t_{1/2}$ value increases to 134 ± 16 seconds. In other words, the exchange of *Ye*SctQ increases by a factor of two under secretion-ON conditions (45).

The shuttling mechanism by *Ye*SctQ-containing complexes was more recently supported by additional experimental evidence (137). This study showed that the presence or absence of the effector protein YopO and its cognate chaperone SycO influenced the mobility of freely-diffusing *Ye*SctQ and *Ye*SctL. Specifically, in the presence of the YopO/SycO pair, decreased mobility of PAmCherry-labeled *Ye*SctQ and *Ye*SctL was observed in single-molecule tracking PALM experiments. This result was consistent with in vivo proximity labeling of *Ye*SctQ followed by purification and proteomics analysis, which showed *Ye*SctQ enriched with 6 of the 8 *Y. enterocolitica* effector proteins as well as *Ye*SctL (137). Taken together, the dynamic shuttling model of effector proteins by *Ye*SctQ-containing complexes is supported by 1) the dynamic binding and unbinding of *Ye*SctQ during secretion (45) and 2) the slowed diffusion of *Ye*SctQ in the presence of the

YopO/SycO pair (137). The increased abundance of *Ye*SctQ in the cytosolic fraction compared to the injectisome-bound fraction additionally supports this model (95,138).

Shuttling of effector proteins to the injectisome by *Ye*SctQ-containing complexes is an intriguing model, yet many questions remain. For instance, it is not known whether or not *Ye*SctQ exchanges with injectisomes as a monomer, in association with itself, in association with other proteins, or perhaps a combination of these different options. If *Ye*SctQ binds chaperone:effector complexes in a 1:1 ratio (i.e. one effector protein is delivered per *Ye*SctQ), then the rate of *Ye*SctQ exchange at injectisomes must be fast enough to match measured secretion rates. Previous data in *Salmonella* estimated that 7-60 effector proteins were secreted per second by an individual cell, with a range of 10-100 injectisomes present per cell (139–141). Taking the lowest value of each of these ranges provides the smallest achievable effector protein secretion rate for an individual injectisome, which results in 7 effectors secreted per second per 10 injectisomes, which corresponds to a secretion rate with a lower-bound of 1 effector protein every 1.43 seconds per injectisome.

In order to begin to test the effector shuttling model, quantitative analyses of *Ye*SctQ binding kinetics and protein secretion rates are useful. The existing data shows dynamic exchange of *Ye*SctQ through FRAP experiments. While these experiments indeed showed that PAmCherry-labeled *Ye*SctQ exchanges with bleached molecules during the fluorescence recovery time of individual injectisomes, the observed dynamics of *Ye*SctQ were slow, which does not correlate with the rapid secretion of effectors by *Salmonella*. The work presented in this chapter leverages a novel single-molecule binding kinetics analysis workflow in combination with a stochastic simulation algorithm to correlate

theoretical protein secretion rates with measured rates of dissociation ($k_{unbinding}$) of *Ye*SctQ from injectisomes. Results show *Ye*SctQ binding and unbinding with injectisomes during secretion. Re-analysis of previously published FRAP data shows significantly slower rates of dissociation compared to those obtained by single-molecule studies. Moreover, the *Ye*SctQ rate of injectisome unbinding is markedly faster than the measured unbinding rates of *Ye*SctL and *Ye*SctN. Stochastic simulations of injectisome binding kinetics support a model in which experimentally measured effector secretion rates are theoretically possible for the $k_{unbinding}$ rates obtained from this study, but does not when using the $k_{unbinding}$ value obtained from FRAP studies.

5.2 RESULTS

5.2.1 FRAP data suggests minimal *Ye*SctQ exchange with injectisomes

In order to directly compare the measured rates of *Ye*SctQ unbinding in *Y. enterocolitica*, the previously published FRAP data (45) was re-analyzed using custom-built fitting and FRAP simulation functions in MATLAB. The fitting function fits experimental FRAP data (normalized to a pre-bleach intensity of 1) to the exponential recovery function

$$I(t) = 1 - (e^{-k_{unbinding} * t}) \quad (5.1)$$

where $I(t)$ is the recovery in fluorescence intensity after bleaching and $k_{unbinding}$ is the dissociation rate constant, also known as k_{OFF}.. Equation 5.1 represents the cumulative probability that an event, which in this case is *Ye*SctQ unbinding from the injectisome, has occurred by time t. Using this equation operates under the limiting constraint of full

occupancy of *Ye*SctQ binding sites at the injectisome (46,106,142), and therefore, the *Ye*SctQ $k_{binding}$ rate, i.e. k_{ON} rate, does not influence the return in fluorescence intensity.

Data from three different FRAP experiments were fit with Eq. 5.1, and the resulting $k_{unbinding}$ values determined by each fit served as inputs into the FRAP simulation. The first condition, in which *Y. enterocolitica* was grown under secretion-OFF conditions, shows a fluorescence recovery time spanning ~350 seconds with a total fluorescence recovery of only ~60% (**Fig. 5.1a**). Fitting of the secretion-OFF data provides a $k_{unbinding}$ value of 0.003 s^{-1} meaning that, on average, *Ye*SctQ remains bound to the injectisome for 1/ $k_{unbinding}$ = 333 seconds. Fitting of secretion-ON data suggests faster rates of dissociation for *Ye*SctQ, with a fit $k_{unbinding}$ value of 0.010 s^{-1} (**Fig. 5.1b**). Under this condition, *Ye*SctQ remains bound to the injectisome for 1 / $k_{unbinding}$ = 94 seconds, on average, which is significantly faster than observations made under secretion-OFF conditions.

Diepold et al. also provide FRAP data for *Y. enterocolitica* when grown under a combination of the above two conditions. Specifically, they grew *Y. enterocolitica* under secretion-OFF conditions for 3 hours and, immediately before imaging, switched the growth medium to the secretion-ON condition. Fitting of the secretion-OFF to -ON data provides a $k_{unbinding}$ value of 0.013 s^{-1} meaning that, on average, *Ye*SctQ remains bound to the injectisome for 1 / $k_{unbinding}$ = 77 seconds (**Fig. 5.1c**). Therefore, the rate of dissociation of *Ye*SctQ is faster under the secretion-OFF to -ON condition compared to the secretion-ON condition. Collectively, re-analysis of these data suggests that the binding kinetics of *Ye*SctQ are sensitive to secretion-state of the injectisome, where *Ye*SctQ remains

bound to the injectisome for an extended period of time even in the secretion-ON state (77 seconds).

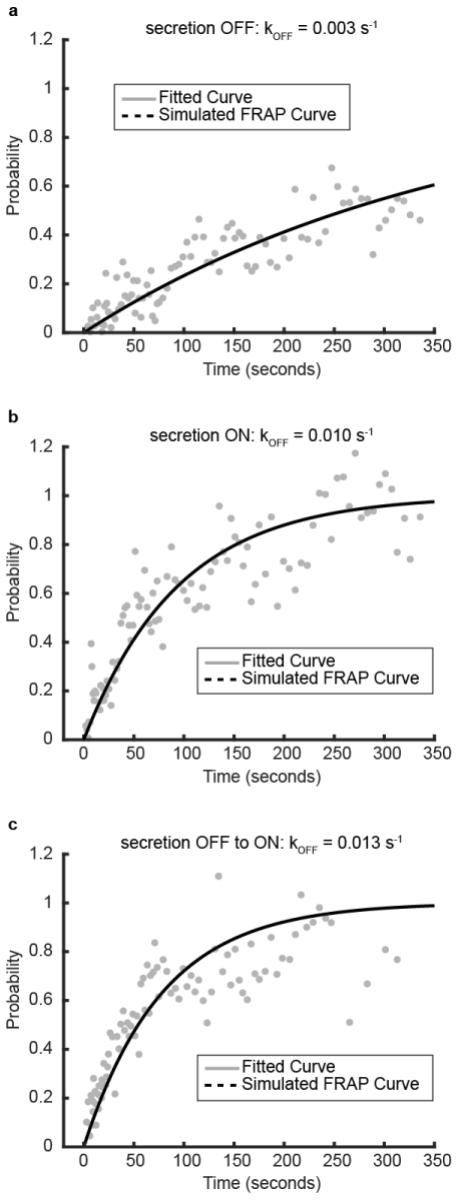

Figure 5.1: Distribution fits (black lines) of FRAP data (gray dots) for PAmCherry-labeled *Ye*SctQ in *Y. enterocolitica* when grown under a) secretion-OFF conditions, b) secretion-ON conditions, and c) secretion-OFF to -ON conditions.

5.2.2 Single-molecule bound-time analysis of *Ye*SctQ

As a complement to the above FRAP analysis, we measured bound-times of *Ye*SctQ with the injectisome by visualizing the corresponding single-molecule binding events of fluorescently-labeled *Ye*SctQ. This approach requires constitutive fluorescence of labeled-*Ye*SctQ molecules, as the entire duration of *Ye*SctQ binding times with injectisomes must be visualized. *Ye*SctQ is highly expressed in *Y. enterocolitica*, so that it is impossible to detect individual, injectisome-bound *Ye*SctQ labeled with constitutively-fluorescent proteins. However, labeling Halo-*Ye*SctQ with HaloTag ligands offers a solution to this issue. HaloTag covalently binds to synthetic fluorescent dyes (the ligands), which contain a specific reactive linker. Using a fluorescent dye allows for control over how many *Ye*SctQ proteins are labeled in a single-molecule imaging experiment.

Prior to imaging, *Y. enterocolitica* was grown under secretion-ON conditions. Halo-*Ye*SctQ was expressed from the native *pYV* virulence plasmid and was sparsely labeled with the fluorescent dye JFX549. Additionally, the laser intensity used within the imaging experiment was attenuated to ensure minimal photobleaching of JFX549. The camera exposure time was set to 0.5 s to observe only slow or stationary molecules. At this exposure time, fluorescent signals originating from fast-diffusing proteins are motion-blurred and become part of the background.

Individual *Ye*SctQ proteins readily bind to injectisomes, which can be seen in the raw data (**Fig. 5.2b**). For each protein binding event, a bound-time trace containing the number of photons collected per frame for an individual protein was generated (**Fig. 5.2c**). From the duration of these traces, the bound time distribution of single *Ye*SctQ at the injectisome was determined (**Fig. 5.2c**).

Fitting the bound-time distribution with an exponential decay allows for the determination of the *Ye*SctQ injectisome unbinding rate, $k_{unbinding}$, value. The exponential decay function is written in the form of

$$y = (A) * e^{-k_{unbinding} * t} \quad (5.2)$$

Unlike the data from the above FRAP experiments, where we operate under the limiting constraint of full occupancy of *Ye*SctQ binding sites at the injectisome (46,106,142), single-molecule bound-time data is not influenced by a given protein's binding kinetics, or $k_{binding}$ rate. Examining the bound-time data for *Ye*SctQ indeed shows an exponentially decaying relationship (**Fig. 5.2d**). Fitting this distribution with the above exponential decay function provides an excellent fit, with *Ye*SctQ $k_{unbinding}$ = 0.075 s^{-1}, which means that a single *Ye*SctQ remains bound to the injectisome for 1 / $k_{unbinding}$ = 13.3 seconds on average. The unbinding rate determined from this single-molecule study is therefore ~6X faster than the unbinding rate determined the above FRAP analysis. This difference has important implications for the shuttling model.

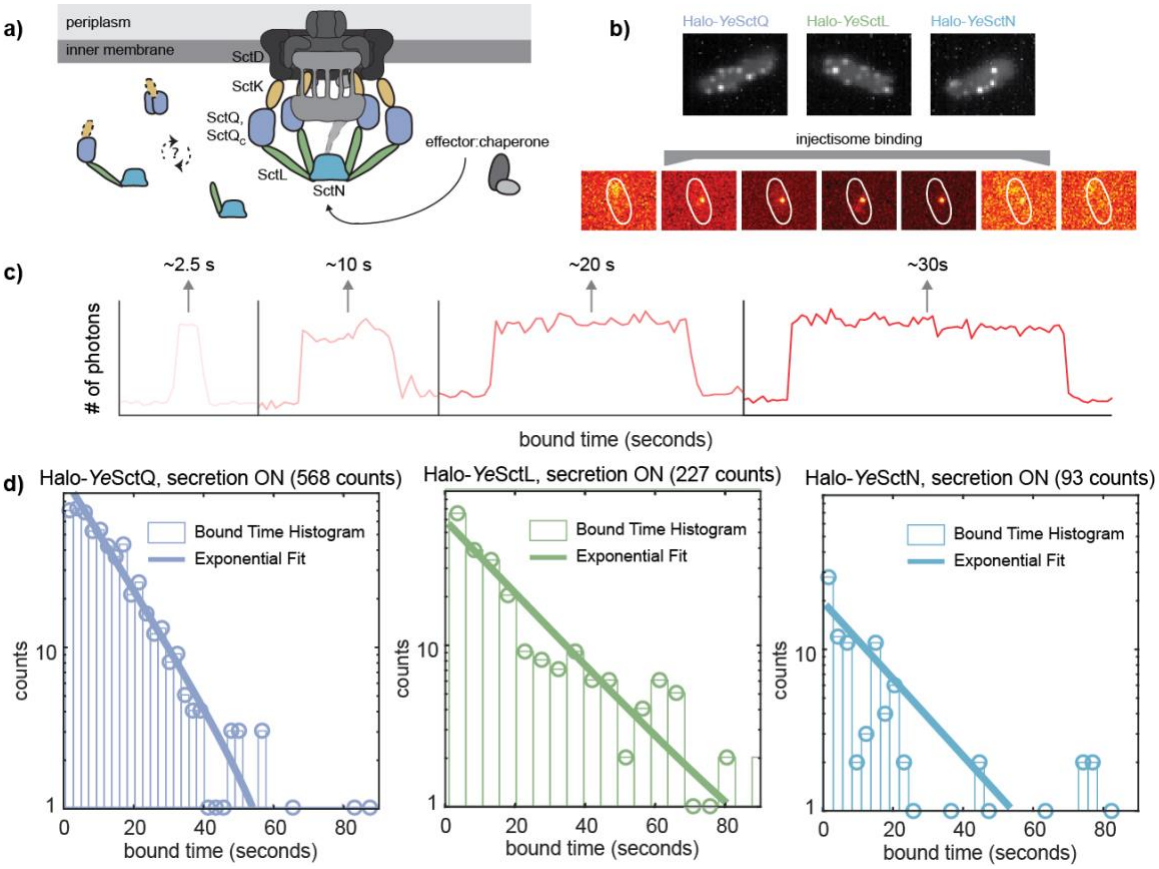

Figure 5.2: Single-molecule bound time analysis. a) *Ye*SctQ binds and unbinds with injectisomes, either as part of a complex or with *Ye*SctK, *Ye*SctL, and *Ye*SctN. Effector:chaperone complexes are additionally delivered to injectisomes, potentially by the specified *Ye*SctQ-containing complexes. b) (top) Average intensity images of JFX549-labeled Halo-*Ye*SctQ, -*Ye*SctL, and -*Ye*SctN reveal punctate foci at injectisome locations. (bottom) Example raw data of a *Ye*SctQ injectisome binding event. The individual *Ye*SctQ protein binds and unbinds with injectisome. The outline of the corresponding cell shows that unbound molecules remain localized within the cell but are severely blurred. c) Example bound-time traces from *Ye*SctQ shows reveals clear binding signatures of individual molecules for a range of binding times. d-f) Bound-time distribution fits for *Ye*SctQ, *Ye*SctL, and *Ye*SctN result in k_{OFF} values of 0.075 s^{-1}, 0.052 s^{-1}, and 0.055 s^{-1},

respectively. The number of counts reflects the difference in expression levels among the three proteins.

5.2.3 Simulating theoretical protein secretion rate distributions

I performed stochastic simulations based on the Gillespie algorithm (143,144) to quantify and test the effector protein shuttling model and to reconcile the difference between *Ye*SctQ unbinding rates reported from the above single-molecule bound-time analysis and the previously published FRAP analysis (**Fig. 5.3a**). Stochastic simulations for systems with known reaction rates can be used to track specific events over time, for example, molecular interactions within a cell. Our algorithm assumes that 24 binding sites are available for *Ye*SctQ at the injectisome (corresponding to the proposed copy number of *Ye*SctQ at a single injectisome (46)). The algorithm is structured to simulate a system where the number of available binding sites, $N_{BS} - N_{Bound}$, varies with time, and the available proteins for binding is present in excess to ensure that binding events are not limited by a small number of unbound proteins. To understand the system's dynamics at any given moment, we compute the 'total rate' of events, as defined by a master equation. This rate is the sum of two components: the rate at which proteins can bind to available sites and the rate at which proteins can unbind from the sites they currently occupy. Mathematically, this is represented by the equation:

$$total\ rate = k_{binding} * (N_{BS} - N_{Bound}) + k_{unbinding} * N_{Bound} \qquad (5.3)$$

where N_{BS} the total number of binding sites and N_{Bound} is the number of proteins currently bound to the complex. In order to simulate the time-evolution of binding and unbinding

events, the time to the next event is not fixed but is instead determined by a random process, reflective of the natural unpredictability of molecular interactions. At each time step of the simulation, the next event is predicted by δ_t, which is computed by

$$\delta_t = -\frac{\log(r_1)}{total\ rate} \tag{5.4}$$

with $r1$ being a uniformly distributed random number between 0 and 1. This expression reflects the exponential waiting time for the next event in a Poisson process. Once the time until the next event is determined, another random number, r_2, is generated, which reflects the probability of a binding event occurring at any given time. Specifically, if

$$r_2 < \frac{k_{binding} * (N_{BS} - N_{Bound})}{total\ rate} \tag{5.5}$$

then a binding event has occurred. Conversely, if r_2 is greater than the ratio defined in Eq. 5.5. expression, an unbinding event has occurred. If an unbinding event does occur, the algorithm finds the molecule that has spent the longest duration of time bound to the system and sets its binding time back to 0, indicating that it has unbound from the system. Additionally, N_{Bound} is decremented by 1, indicating a binding site has been made available. The random numbers, r_1 and r_2, allow us to simulate the stochastic behavior of the system, where the occurrence of each type of event (binding or unbinding) is governed by its respective probability. In the limit of very high $k_{binding}$ relative to $k_{unbinding}$, the $total\ rate$ for an individual injectisome approaches

$$k_{binding} * (N_{BS} - N_{Bound}) \approx k_{unbinding} * N_{Bound} \tag{5.6}$$

which represents the steady state of the system, i.e. the frequency of binding events is effectively equal to the frequency of unbinding events. The term $k_{binding} * (N_{BS} - N_{Bound})$ represents the rate at which new binding events can potentially occur. As N_{Bound} increases, the difference, $N_{BS} - N_{Bound}$, decreases, reducing the rate of new binding events. Therefore, as the system approaches a steady state, the number of available binding sites approaches zero (**Fig. 5.3b**). In this scenario, the above expression can be simplified and re-defined as

$$Rate\ of\ turnover \approx k_{unbinding} \quad (5.7)$$

In terms of the problem at hand, the rate of unbinding is equal to the turnover rate of *Ye*SctQ, which is governed by the *Ye*SctQ dissociation rate constant. A key assumption of this algorithm is that, for each binding event by a simulated molecule (i.e., *Ye*SctQ), a protein is secreted by the system (**Fig. 5.3a**). Further, only one *Ye*SctQ is required to secrete an effector protein. Given these assumptions, theoretical protein secretion rates could be determined.

Given that the *Ye*SctQ $k_{binding}$ value is unknown, a range of $k_{binding}$ values were considered in combination with the measured $k_{unbinding}$ values from the single-molecule bound-time and FRAP analyses. As the $k_{binding}$ grows larger, the number of occupied binding sites saturates and the rate of protein secretion becomes limited solely by the rate of *Ye*SctQ dissociation. With $k_{unbinding}$ set to 0.011 s^{-1}, the results of the simulation show that the theoretical protein secretion rate distribution asymptotes at ~0.2 effector proteins per second (**Fig. 5.3c**). Given the ideality of the simulated scenario (individual *Ye*SctQ proteins deliver effector:chaperone complexes with 100% efficiency), the calculated

secretion rate represents an upper bound. The reported $k_{unbinding}$ values from FRAP data are not consistent with an effector protein shuttling model, as the rate of secretion is limited to slower rates than experimentally recorded rates(139–141). Contrarily, simulation results with $k_{unbinding}$ set to 0.075 s^{-1} allows for faster secretion rates, up to 1.7 effector proteins per second (**Fig. 5.3c**), which is consistent with experimentally recorded rates (139–141). Thus, the results from the single-molecule bound-time analyses support the effector protein shuttling model. However, it remains unclear whether shuttling is done by *Ye*SctQ alone or by *Ye*SctQ-containing complexes.

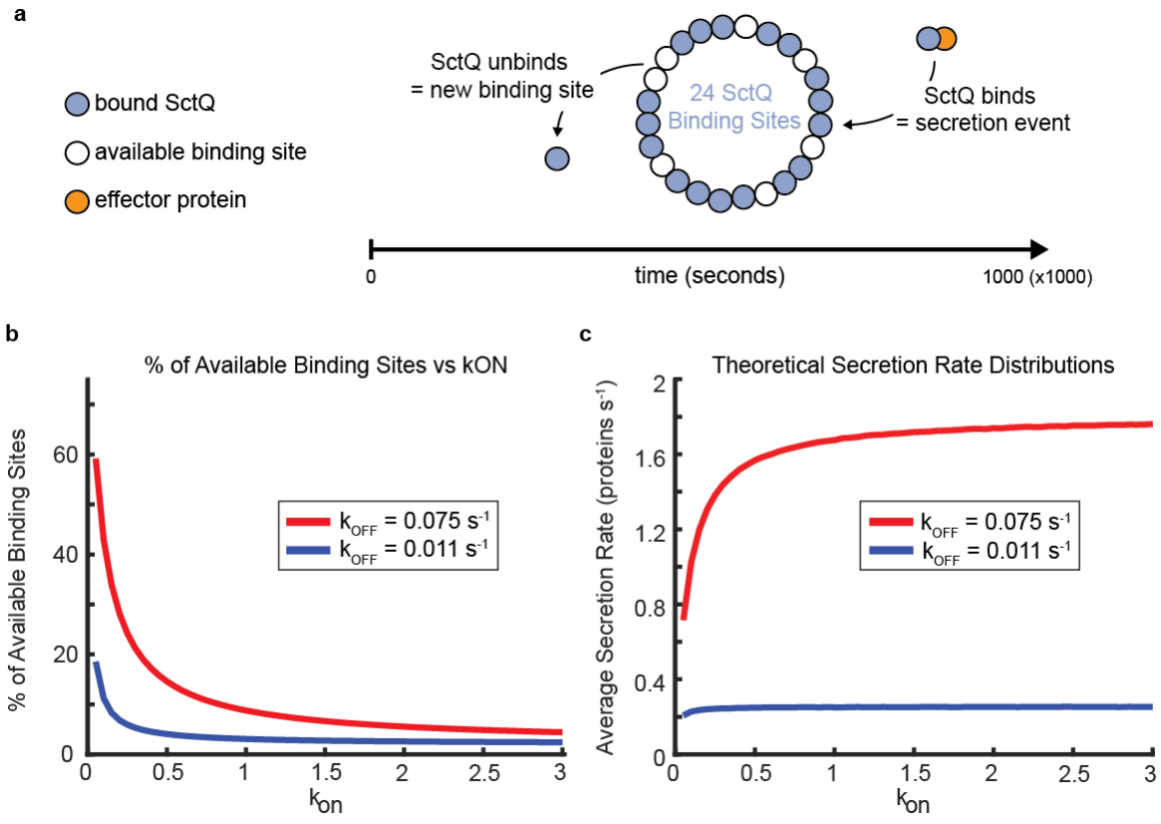

Figure 5.3: a) A model of effector protein secretion was developed, which simulates stochastic binding and unbinding events based on known $k_{binding}$ and $k_{unbinding}$ values.

For each simulation, a different $k_{binding}$ value is selected, and the simulation is performed for 1 million seconds. Each binding event corresponds to the secretion of an individual effector protein. b) Theoretical secretion rate distributions for simulations with $k_{unbinding}$ set to 0.011 s^{-1} and 0.075 s^{-1} and a range of $k_{binding}$ values. c) The percentage of available binding sites for *Ye*SctQ is saturated (i.e., approaches zero) as the $k_{binding}$ values grow larger. At a $k_{binding}$ value of 3, virtually 0% of the binding sites for *Ye*SctQ are available, which corresponds to the secretion rate being rate-limited by the $k_{unbinding}$ value of *Ye*SctQ.

5.2.4 Single-molecule bound-time analysis of *Ye*SctL and *Ye*SctN

If *Ye*SctQ binds and unbinds from injectisomes as part of a complex with *Ye*SctL and/or *Ye*SctN, then it stands to reason that *Ye*SctL and/or *Ye*SctN should have similar k_{OFF} values. Specifically, our previously published tracking results support a model in which *Ye*SctL and *Ye*SctN form two distinct complexes in the cytosol of *Y. enterocolitica* under secretion-ON conditions. While *Ye*SctL and *Ye*SctN interact independently of *Ye*SctQ, a majority of cytosolic *Ye*SctL and *Ye*SctN are complexed with *Ye*SctQ, as evidenced by a shared diffusive state among the three proteins (95). Therefore, if the $k_{unbinding}$ values for *Ye*SctL and *Ye*SctN differ from *Ye*SctQ, then these proteins may be exchanging with injectisomes as independent proteins or protein complexes.

Fitting of the *Ye*SctL and *Ye*SctN bound time distributions results in $k_{unbinding}$ values of 0.052 s^{-1} and 0.055 s^{-1}, respectively (**Fig. 5.2cd**). This correlates to an average injectisome-bound time of ~19.2 seconds for *Ye*SctL and ~18.2 seconds for *Ye*SctN. These

data therefore suggest that, on average, *Ye*SctL and *Ye*SctN remain bound to injectisomes longer than *Ye*SctQ. There are two main possibilities that may explain this scenario: 1) *Ye*SctQ, *Ye*SctL, and *Ye*SctN bind to injectisomes as part of a complex, but *Ye*SctQ unbinds independently of and faster than *Ye*SctL and *Ye*SctN or 2) *Ye*SctQ binds and unbinds from injectisomes independently of *Ye*SctL and *Ye*SctN. It is further possible, given the similar $k_{unbinding}$ values for *Ye*SctL and *Ye*SctN, that these proteins bind and unbind from injectisomes as part of the same complex, which is supported by the observed complex of *Ye*SctL and *Ye*SctN in our cytosolic tracking data (95). Further experiments are necessary to determine precisely which complexes are binding and unbinding from injectisomes. Such information will lend insight into the effector protein shuttling model, as each complex may have functionally distinct roles.

5.3 DISCUSSION

A model of *Ye*SctQ bound to effector:chaperone complexes in the cytosol of *Y. enterocolitica* as a means to deliver injectisome substrates for secretion has been proposed (137). However, in order to agree with experimentally measured protein secretion rates (139–141), the rate of effector protein delivery by *Ye*SctQ-containing complexes must therefore be rapid. The single-molecule bound-time analysis presented in this chapter is quantitatively consistent with an effector protein shuttle model by *Ye*SctQ containing complexes. If *Ye*SctQ shuttles effectors to the injectisome, then it is likely that *Ye*SctQ exchange occurs in dynamic equilibrium, where one *Ye*SctQ binds immediately after another unbinds. Given that *Ye*SctQ is expressed in excess abundance, a cytosolic pool of freely-diffusing *Ye*SctQ is therefore made available for injectisomes. In this model, the

secretion of effector proteins is therefore directly related to the rate of *Ye*SctQ dissociation from injectisomes.

Two assumptions made in this model need to be validated in order to confirm the values reported the theoretical rates of effector protein secretion. One component to the model that was not tested further is the efficiency of secretion. Specifically, the current model assumes that, for every *Ye*SctQ binding event with injectisomes, an effector protein is secreted. In other words, 100% of the available pool of *Ye*SctQ for injectisome binding is complexed with cognate effector:chaperone complexes. This assumption likely does not hold true in the context of living cells, and therefore, the effector secretion rate values provided here represent the fastest/highest theoretical secretion rates based on the measured *Ye*SctQ $k_{unbinding}$ values. If, for example, the system is only 50% efficient, then the average number of effector proteins secreted per second will decrease by 50%. In order to determine how the efficiency of effector protein delivery impacts secretion rate, one can simply compute

$$secretion\ rate = \frac{\eta * Number\ of\ binding\ events}{t_T} \tag{5.8}$$

where η represents the efficiency of effector protein delivery and t_T is the total simulation time. Additionally, this model assumed that individual *Ye*SctQ proteins deliver effector:chaperone complexes to the injectisome. If, for example, a complex containing 2 or 4 *Ye*SctQ were to deliver individual effector:chaperone complexes, then the effector secretion rate would decrease by 25% or 50%, respectively.

In order to verify the theoretical protein secretion rate estimates computed in this dissertation, quantitative protein secretion rate measurements in *Y. enterocolitica* under the

specific experimental conditions used in this dissertation are needed. While different methods have been introduced into the literature that allow for the quantification and visualization of protein secretion by injectisomes (140,145–148), accurate quantification of secretion of individual injectisomes is non-trivial (see Section 6.2.2). However, such measurements, together with accurate injectisome counts per cell, will help to further support or refute the effector protein shuttling model. Additionally, the above bound-time measurements should be repeated under the same conditions used in the FRAP analysis, e.g. secretion-OFF and secretion-OFF to -ON conditions, to confirm that the rate of *Ye*SctQ exchange is influence by the secretion-state of the injectisome. Taken together, these data will help further support or refute the effector protein shuttling model.

5.4 MATERIALS AND METHODS

5.4.1 Cell culture

Y. enterocolitica cultures were prepared for imaging under secreting conditions as described in Section 4.6.2. After 3 hours of growth, JFX549 was added to 1.5 mL of culture to a final concentration of 1 – 3 nM, and cells were incubated for an additional 30 minutes. Cells were then washed and plated for imaging as described in Section 4.6.2.

5.4.2 Super-resolution fluorescence imaging

Experiments were performed on a custom-built, dual-color inverted fluorescence microscope as described in Section 4.6.3. The laser intensity was attenuated significantly (lowest possible input laser power), and an optical density (OD) filter of 1.0 added to the imaging pathway. Additionally, for 2D tracking of stationary molecules, the DHPSF phase mask was removed. Up to 2000 frames were collected per field-of-view with an exposure time of 500 ms.

5.4.3 Data processing

Prior to PSF fitting, images from each field-of-view were background subtracted. PSFs were fit using ThunderSTORM in ImageJ (149) using the difference of Gaussians (DoG) filter option. A standard deviation intensity image was generated for each image stack, which clearly show injectisome locations as punctate foci. Injectisome locations from standard deviation images were additionally fit using ThunderSTORM, and these locations were used to filter single-molecule localizations within each, respective image stack. A nearest-neighbor tracking was then used to generate a list of trajectories, where

mobile trajectories were filtered from the data based on a distance threshold of 500 nm, i.e. trajectories with displacements larger than this distance were not considered for further analysis. For each stationary trajectory, the photon counts from the corresponding images were calculated, generating a bound-time trace for an injectisome-bound protein. A MATLAB script that displays the bound-time trace as well as a movie of the corresponding single-molecule binding event from the raw data was written, which allowed for manual annotation of the post-filtered traces. Traces were manually categorized as 1) stationary, 2) semi-mobile / slowly moving, 3) multiple molecules present, or 4) undetermined. Only traces from the first category then used for further analysis, which only contained individual protein binding events with injectisomes.

5.4.4 Data analysis

Bound-time distributions of each protein were fit to the Eq. 5.2. Fitting of the data was conducted utilizing the 'fitnlm' function within MATLAB, which incorporates a nonlinear model. To mitigate the influence of outliers, the 'RobustWgtFun' parameter was set to 'bisquare.'

5.4.5 Stochastic simulation algorithm

A stochastic simulation algorithm was written in MATLAB to explore the relationship between the binding rate constant ($k_{binding}$) and protein secretion rates in a biological system, which was modeled after the *Ye*SctQ effector shuttling model. Each simulation operates over a total time of 1000 seconds and assumes that, at the beginning

of the simulation, all binding sites by *Ye*SctQ are occupied. The script then performs 1000 simulations for each $k_{binding}$ value. During the simulation, the script tracks the binding and unbinding events of *Ye*SctQ at each timestep, and a protein is secreted by the system each time *Ye*SctQ binds. The average number of occupied binding sites at the injectisome and number of proteins secreted for each $k_{binding}$ value was then computed.

Chapter 6
CONCLUSIONS AND FUTURE DIRECTIONS

**** Excerpts from this chapter are taken from Prindle, J. R., de Cuba, O.I.C., & Gahlmann, A. *Single-molecule tracking to determine the abundances and stoichiometries of freely-diffusing protein complexes in living cells: Past applications and future prospects* (J. Chem. Phys., 2023, Vol. 159, No. 7)**

6.1 Significance

The injectisome is responsible for widespread human disease, both historically and currently. Determining the precise mechanisms that contribute to pathogenesis may aid in the development of novel drugs to combat its virulent properties. In *Y. enterocolitica*, the sorting platform and ATPase transiently associate with injectisomes during secretion, yet their functional roles remain largely unknown.

The work presented here leverages live-cell super-resolution microscopy approaches that provide an in-depth look into the binding properties of the sorting platform and ATPase, both in the cytosol and at the injectisome. A previously published diffusion analysis pipeline was updated to address the problem of model selection commonly encountered in frequentist (i.e. least squares- or maximum likelihood-based) curve fitting approaches of single-molecule tracking data. Specifically, the diffusion coefficient spectrum was introduced, which provides an easily-interpretable approach to analyzing complex diffusion data. Diffusion coefficient spectrum analysis of *Ye*SctQ, *Ye*SctL, and *Ye*SctN revealed that these proteins form distinct complexes in the cytosol. Further, bound-time analysis of these proteins with injectisomes revealed that, on average, *Ye*SctQ unbinds from injectisomes more rapidly than *Ye*SctL and *Ye*SctN.

Collectively, these data support a model in which distinct *Ye*SctQ-containing complexes shuttle effector proteins to the injectisome. According to this model, *Ye*SctQ associates with effector:chaperone protein complexes that are then delivered to the injectisome for secretion. On average, *Ye*SctQ remains bound to the injectisome for ~ 13.3 seconds, which correlates to a theoretical maximum secretion rate of 1.7 proteins per second. However, it remains unknown which of the *Ye*SctQ-containing complexes serves

as the shuttle. For instance, our data supports two, or perhaps three, *Ye*SctQ-containing complexes in the cytosol of actively secreting *Y. enterocolitica.* Validation of the effector protein shuttling model and determination of the *Ye*SctQ-containing complexes that deliver effector:chaperone complexes to the injectisome will be the subject of future research.

6.2 Future Directions

6.2.1 Deviations from Stokes-Einstein diffusion theory

As diffusion coefficient measurements of proteins become more common, it has become possible to test whether the $D \propto T/\left(\eta \cdot MW^{\frac{1}{3}}\right)$ scaling relation predicted by Stokes-Einstein diffusion theory applies to diffusion in intracellular aqueous environments. Deviations from Stokes-Einstein scaling should be expected, because the cytoplasm is not a homogeneous medium of uniform viscosity, but instead is a complex mixture of crowding agents, many of which (e.g. DNA, mRNA, or ribosomes) are substantially larger than the diffusing proteins and protein complexes of interest. Interestingly, currently available results show that, as molecular weight increases, the diffusion rates of proteins and protein complexes is slower than what is predicted by the Stokes-Einstein diffusion theory. One approach to explain this observation is to make the viscosity term dependent on the size of the diffusing molecule (150,151). Another approach is to adjust the $D \propto MW^{\beta}$ scaling exponent β. The results from early, independent experiments clearly show that $\beta <$ -1/3. (We note that this scaling only considers negatively charged proteins. Positively-charged proteins exhibit even slower diffusion on average due to their strong electrostatic interactions with negatively-charged, quasi-stationary biomolecules (152).) In a 2010

study (153), a scaling exponent β of -0.7 was determined using pulsed-FRAP measurements on fluorescently labelled glucose (NBD-glucose, 423 Da), green fluorescent protein (GFP, 26.9 kDa), and a large oligomeric protein assumed to consist primarily of four subunits of GFP-tagged β-galactosidase ($(\beta\text{-gal-GFP})_4$, ~582 kDa). When cells were osmotically upshifted (i.e., grown in media supplemented with 2 M NaCl), NBD-glucose remained fairly mobile, whereas the diffusion of GFP and $(\beta\text{-gal-GFP})_4$ became comparatively more impeded. In these osmotically upshifted cells, a scaling exponent β of -0.8 was determined – an effect attributed to increased biomolecular density in the cytoplasm. In a 2011 review (154), FRAP- and FCS-based diffusion coefficients measurements from multiple studies for different-sized proteins were combined and a scaling exponent β of -0.7 was determined.

More recent experiments agree with these findings (**Figure 6.1**). The FCS-measured diffusion rates for sfGFP fusion proteins with molecular weights ranging from 25 kDa to 165 kDa produced a scaling exponent β of -0.56 (131). The tested proteins are globular in shape and are not known to bind DNA or form homomultimers. Also in *E. coli*, single-molecule displacement mapping (155), was used to study diffusion of fluorescent fusion proteins with molecular weights ranging from 25 kDa to 300 kDa. The proteins were chosen based on the criteria that they do not interact with DNA or with any other proteins. The results showed a scaling exponent β of -0.54 (132). DNA-binding proteins with molecular weights ranging from ~ 70 kDa to ~ 500 kDa were studied using single-molecule tracking in *E. coli*, in which DNA was enzymatically degraded. These experiments produced a scaling exponent β of -0.75 (130). In contrast to the non-interacting proteins used in the above studies, 3D single-molecule tracking was used to measure the diffusion

coefficients of three soluble proteins of the type 3 secretion system (T3SS) in osmotically-upshifted *Y. enterocolitica*. T3SS proteins are known to interact with each other to form hetero-oligomeric complexes. When calculating the molecular weights based on the likely complex stoichiometries, a scaling exponent β of -0.71 was obtained (95).

Future work should be dedicated towards validating the above diffusion coefficient vs. molecular weight relationships. More recently, it has become possible to design proteins with specified folding and binding properties de novo (156). Previously testing the relationship between molecule size and its diffusion coefficient was made possible through the use of GFP-μNS particles (157). However, GFP-μNS particles are large (> 30 nm in diameter), and therefore, validating the Stokes-Einstein diffusion theory for small (< 10 nm in diameter) proteins is not possible. By having the ability to design proteins de novo, it may be possible to design probes that systematically test the diffusion coefficient vs. molecule weight relationship. The tested probes should ideally be globular in shape and non-interacting with other proteins. Tracking these probes and analyzing the corresponding diffusion data may therefore provide a range of diffusion coefficients for small proteins ranging from 0 – 500 kDa.

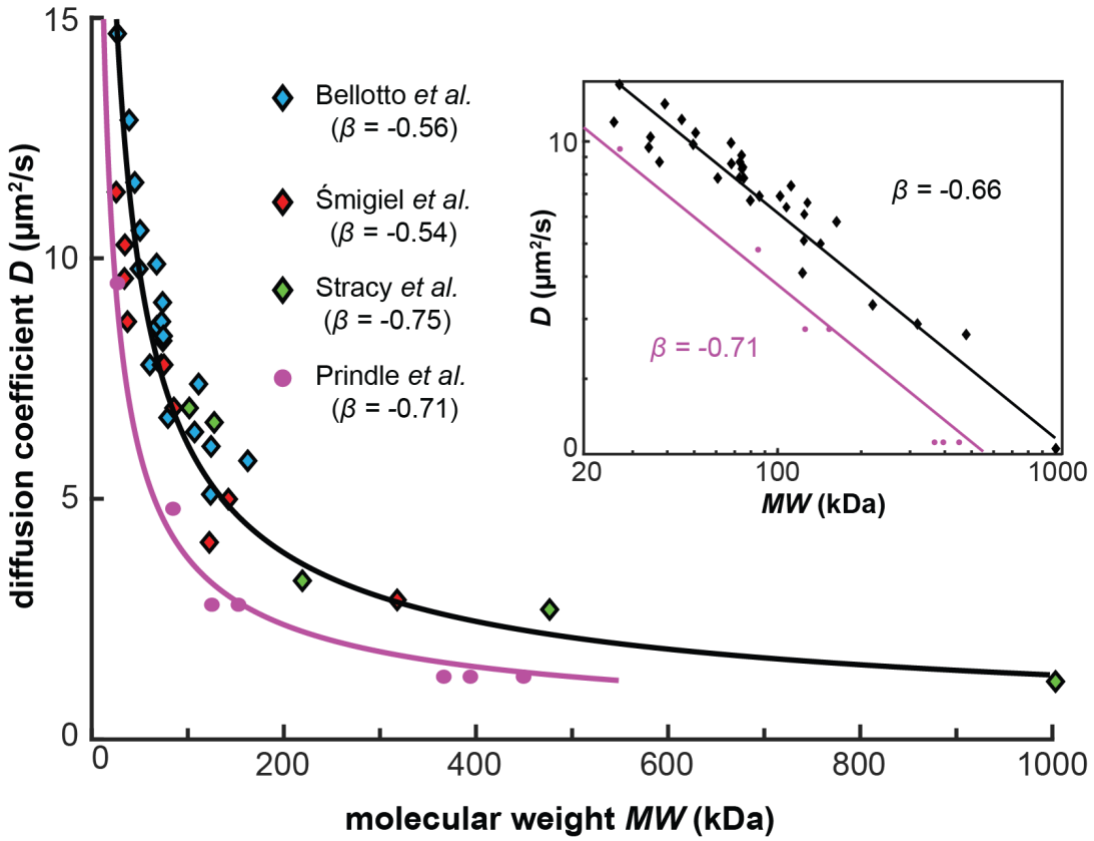

Figure 6.1. Power-law scaling of diffusion coefficients vs. molecular weight data of different proteins and protein complexes. Measurements obtained from three independent studies in *E. coli* (black diamonds){Bellotto, 2022 #6447;Śmigiel, 2022 #6412;Stracy, 2021 #6358} results in a scaling exponent β of -0.66 (black line). In each study, cells were grown in minimal growth media on the day of imaging. In *Y. enterocolitica* (pink circles), a scaling exponent β of -0.71 (pink line) was obtained for freely-diffusing proteins in osmotically-upshifted cells. Inset: log-log plot of same data.

6.2.2 Effector protein shuttling by *Ye*SctQ-containing complexes

The data presented in Chapter 5 supports an effector protein shuttling model by *Ye*SctQ-containing complexes, yet additional data is required to support this model. Interestingly, rates of effector protein secretion by injectisomes in *Y. enterocolitica* has not been previously reported. Such information is required to validate theoretical rates of protein secretion (**Fig 5.3**). Previously published data in *S.* Typhimurium revealed that the effector SipA was secreted at a rate of 7 – 60 effector proteins per second with an average of 10 – 100 injectisomes per cell (139–141). These rates were determined from time-lapse fluorescence microscopy experiments.

The authors of the study (140) took a two-step approach at visualizing effector protein delivery into eukaryotic host cells: 1) measure the rate of SipA accumulation in the eukaryotic host cell and 2) measure the rate of SipA depletion in the bacterial cell. Given that proteins must be unfolded into their primary structure for secretion, tagging of SipA with a fluorescent protein was not possible. However, the chaperone protein for SipA, InvB, is not secreted. The authors expressed GFP-InvB in *S.* Typhimurium and in Cos7 host cells, and modulations in GFP-InvB fluorescence intensity were visualized during infection. Specifically, GFP-InvB expressed from Cos7 host cells accumulated at the site of infection upon secretion of SipA from the bacterial cytosol. The accumulation of fluorescence intensity at the site of infection was then measured over time. From the slopes of this linear phase of GFP-InvB recruitment, the authors estimated the individual rates of SipA injection by single bacteria.

A more reasonable approach for measuring the rates of secretion by individual injectisomes would be to fuse the N-terminal fragment of YopE (see Section 1.1.1) to a

protein to be destined for secretion, which has been shown to act as a signal for the export of any given protein in *Y. enterocolitica* (147,148). Utilizing the secretion signal to track any protein's export through the injectisome offers the possibility for more precise studies of effector protein secretion rates at the level of single injectisomes. In order for this to be possible, the designed system must be sensitive enough to detect single molecules via fluorescence, requiring us to visualize individual proteins before and after they are secreted. This secretion signal could also be more easily leveraged in diffraction-limited FRAP and fluorescence loss in photobleaching (FLIP) studies in an experimental setup similar to what has been used in the study of bacterial flagellar motors (158,159). For example, by employing a quantitative FLIP experiment, we could monitor the depletion of GFP within the bacterial cytoplasm, correlating the intracellular concentration decrease with the secretion events observed outside the cell. Experimentally-measured effector secretion rates could then be compared with the theoretical secretion rates determined in this work, which will further validate the effector shuttle model by *Ye*SctQ-containing complexes.

6.3 Conclusions

The work presented here provides the approximate stoichiometries and relative abundances of distinct protein complexes containing the injectisome sorting platform and ATPase in the cytosol of living *Y. enterocolitica*. By introducing the diffusion coefficient spectrum, robust analysis of *Ye*SctQ, *Ye*SctL, and *Ye*SctN tracking data in wild-type and deletion mutants was possible. One complex contains *Ye*SctQ, another contains *Ye*SctL and *Ye*SctN, and a third contains all three of these proteins. The average binding times of

*Ye*SctQ, *Ye*SctL, and *Ye*SctN were additionally measured. Collectively, these data support a model in which *Ye*SctQ-containing complexes shuttle effectors to the injectisome for secretion. Future work should be dedicated towards validating this model by 1) validating the Stokes-Einstein diffusion theory by carefully controlling protein size/molecular weight and 2) measuring the approximate rates of effector protein secretion in *Y. enterocolitica.*

References

1. Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, et al. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nat Rev Microbiol. 2015 Jun;13(6):343–59.

2. Green ER, Mecsas J. Bacterial Secretion Systems – An overview. Microbiol Spectr. 2016 Feb;4(1):10.1128/microbiolspec.VMBF-0012–2015.

3. Wagner S, Grin I, Malmsheimer S, Singh N, Torres-Vargas CE, Westerhausen S. Bacterial type III secretion systems: a complex device for the delivery of bacterial effector proteins into eukaryotic host cells. FEMS Microbiol Lett [Internet]. 2018 Aug 9 [cited 2019 Oct 21];365(19). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6140923/

4. Lara-Tejero M, Galán JE. The Injectisome, a Complex Nanomachine for Protein Injection into Mammalian Cells. EcoSal Plus [Internet]. 2019 Mar 29 [cited 2022 Apr 15];8(2). Available from: https://journals.asm.org/doi/full/10.1128/ecosalplus.ESP-0039-2018

5. Coburn B, Sekirov I, Finlay BB. Type III Secretion Systems and Disease. Clin Microbiol Rev. 2007 Oct;20(4):535–49.

6. Andor A, Trülzsch K, Essler M, Roggenkamp A, Wiedemann A, Heesemann J, et al. YopE of Yersinia, a GAP for Rho GTPases, selectively modulates Rac-dependent actin structures in endothelial cells. Cell Microbiol. 2001;3(5):301–10.

7. Larsson DGJ, Flach CF. Antibiotic resistance in the environment. Nat Rev Microbiol. 2022 May;20(5):257–69.

8. McShan AC, De Guzman RN. The Bacterial Type III Secretion System as a Target for Developing New Antibiotics. Chem Biol Drug Des. 2015 Jan;85(1):30–42.

9. Bai F, Li Z, Umezawa A, Terada N, Jin S. Bacterial type III secretion system as a protein delivery tool for a broad range of biomedical applications. Biotechnol Adv. 2018;36(2):482–93.

10. Miletic S, Goessweiner-Mohr N, Marlovits TC. The Structure of the Type III Secretion System Needle Complex. Curr Top Microbiol Immunol. 2020;427:67–90.

11. Dey S, Chakravarty A, Guha Biswas P, De Guzman RN. The type III secretion system needle, tip, and translocon. Protein Sci Publ Protein Soc. 2019 Sep;28(9):1582–93.

12. Deng W, Marshall NC, Rowland JL, McCoy JM, Worrall LJ, Santos AS, et al. Assembly, structure, function and regulation of type III secretion systems. Nat Rev Microbiol. 2017 Jun;15(6):323–37.

13. Bergeron JRC, Marlovits TC. Cryo-EM of the injectisome and type III secretion systems. Curr Opin Struct Biol. 2022 Aug 1;75:102403.

14. Jenkins J, Worrall LJ, Strynadka NCJ. Recent structural advances towards understanding of the bacterial type III secretion injectisome. Trends Biochem Sci. 2022 Sep 1;47(9):795–809.

15. Büttner D. Protein Export According to Schedule: Architecture, Assembly, and Regulation of Type III Secretion Systems from Plant- and Animal-Pathogenic Bacteria. Microbiol Mol Biol Rev. 2012 Jun;76(2):262–310.

16. Dietsche T, Mebrhatu MT, Brunner MJ, Abrusci P, Yan J, Franz-Wachtel M, et al. Structural and Functional Characterization of the Bacterial Type III Secretion Export Apparatus. PLOS Pathog. 2016 Dec 15;12(12):e1006071.

17. Kimbrough TG, Miller SI. Contribution of Salmonella typhimurium type III secretion components to needle complex formation. Proc Natl Acad Sci. 2000 Sep 26;97(20):11008–13.

18. Wagner S, Königsmaier L, Lara-Tejero M, Lefebre M, Marlovits TC, Galán JE. Organization and coordinated assembly of the type III secretion export apparatus. Proc Natl Acad Sci. 2010 Oct 12;107(41):17745–50.

19. Diepold A, Amstutz M, Abel S, Sorg I, Jenal U, Cornelis GR. Deciphering the assembly of the Yersinia type III secretion injectisome. EMBO J. 2010 Jun 2;29(11):1928–40.

20. Kubori T, Sukhan A, Aizawa SI, Galán JE. Molecular characterization and assembly of the needle complex of the Salmonella typhimurium type III protein secretion system. Proc Natl Acad Sci. 2000 Aug 29;97(18):10225–30.

21. Cordes FS, Komoriya K, Larquet E, Yang S, Egelman EH, Blocker A, et al. Helical Structure of the Needle of the Type III Secretion System of Shigella flexneri*. J Biol Chem. 2003 May 9;278(19):17103–7.

22. Galkin VE, Schmied WH, Schraidt O, Marlovits TC, Egelman EH. The Structure of the *Salmonella typhimurium* Type III Secretion System Needle Shows Divergence from the Flagellar System. J Mol Biol. 2010 Mar 12;396(5):1392–7.

23. Loquet A, Sgourakis NG, Gupta R, Giller K, Riedel D, Goosmann C, et al. Atomic model of the type III secretion system needle. Nature. 2012 Jun;486(7402):276–9.

24. Erhardt M, Singer HM, Wee DH, Keener JP, Hughes KT. An infrequent molecular ruler controls flagellar hook length in Salmonella enterica. EMBO J. 2011 Jul 20;30(14):2948–61.

25. Journet L, Agrain C, Broz P, Cornelis GR. The Needle Length of Bacterial Injectisomes Is Determined by a Molecular Ruler. Science. 2003 Dec 5;302(5651):1757–60.

26. Veenendaal AKJ, Hodgkinson JL, Schwarzer L, Stabat D, Zenk SF, Blocker AJ. The type III secretion system needle tip complex mediates host cell sensing and translocon insertion. Mol Microbiol. 2007;63(6):1719–30.

27. Björnfot AC, Lavander M, Forsberg Å, Wolf-Watz H. Autoproteolysis of YscU of Yersinia pseudotuberculosis Is Important for Regulation of Expression and Secretion of Yop Proteins. J Bacteriol. 2009 Jul;191(13):4259–67.

28. Edqvist PJ, Olsson J, Lavander M, Sundberg L, Forsberg Å, Wolf-Watz H, et al. YscP and YscU Regulate Substrate Specificity of the Yersinia Type III Secretion System. J Bacteriol. 2003 Apr;185(7):2259–66.

29. Zarivach R, Deng W, Vuckovic M, Felise HB, Nguyen HV, Miller SI, et al. Structural analysis of the essential self-cleaving type III secretion proteins EscU and SpaS. Nature. 2008 May;453(7191):124–7.

30. Deane JE, Graham SC, Mitchell EP, Flot D, Johnson S, Lea SM. Crystal structure of Spa40, the specificity switch for the Shigella flexneri type III secretion system. Mol Microbiol. 2008;69(1):267–76.

31. Wiesand U, Sorg I, Amstutz M, Wagner S, van den Heuvel J, Lührs T, et al. Structure of the Type III Secretion Recognition Protein YscU from *Yersinia enterocolitica*. J Mol Biol. 2009 Jan 23;385(3):854–66.

32. Lountos GT, Austin BP, Nallamsetty S, Waugh DS. Atomic resolution structure of the cytoplasmic domain of Yersinia pestis YscU, a regulatory switch involved in type III secretion. Protein Sci. 2009;18(2):467–74.

33. Rudolph M, Carsten A, Kulnik S, Aepfelbacher M, Wolters M. Live imaging of Yersinia translocon formation and immune recognition in host cells. PLOS Pathog. 2022 May 23;18(5):e1010251.

34. Russo BC, Duncan JK, Wiscovitch AL, Hachey AC, Goldberg MB. Activation of Shigella flexneri type 3 secretion requires a host-induced conformational change to the translocon pore. PLOS Pathog. 2019 Nov 14;15(11):e1007928.

35. Young LN, Villa E. Bringing Structure to Cell Biology with Cryo-Electron Tomography. Annu Rev Biophys. 2023;52(1):573–95.

36. Keck C, Enninga J, Swistak L. Caught in the act: In situ visualization of bacterial secretion systems by cryo-electron tomography. Mol Microbiol [Internet]. [cited 2024 Feb 29];n/a(n/a). Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/mmi.15186

37. Hu B, Lara-Tejero M, Kong Q, Galán JE, Liu J. In Situ Molecular Architecture of the Salmonella Type III Secretion Machine. Cell. 2017 Mar 9;168(6):1065-1074.e10.

38. Hu B, Morado DR, Margolin W, Rohde JR, Arizmendi O, Picking WL, et al. Visualization of the type III secretion sorting platform of Shigella flexneri. Proc Natl Acad Sci. 2015 Jan 27;112(4):1047–52.

39. Tachiyama S, Chang Y, Muthuramalingam M, Hu B, Barta ML, Picking WL, et al. The cytoplasmic domain of MxiG interacts with MxiK and directs assembly of the sorting platform in the Shigella type III secretion system. J Biol Chem. 2019 Dec 13;294(50):19184–96.

40. Berger C, Ravelli RBG, López-Iglesias C, Kudryashev M, Diepold A, Peters PJ. Structure of the Yersinia injectisome in intracellular host cell phagosomes revealed by cryo FIB electron tomography. J Struct Biol. 2021 Mar 1;213(1):107701.

41. Henderson R. Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. Proc Natl Acad Sci. 2013 Nov 5;110(45):18037–41.

42. van Heel M. Finding trimeric HIV-1 envelope glycoproteins in random noise. Proc Natl Acad Sci. 2013 Nov 5;110(45):E4175–7.

43. Soto JE, Galán JE, Lara-Tejero M. Assembly and architecture of the type III secretion sorting platform. Proc Natl Acad Sci. 2022 Dec 20;119(51):e2218010119.

44. Tachiyama S, Skaar R, Chang Y, Carroll BL, Muthuramalingam M, Whittier SK, et al. Composition and Biophysical Properties of the Sorting Platform Pods in the Shigella Type III Secretion System. Front Cell Infect Microbiol. 2021;11:488.

45. Diepold A, Kudryashev M, Delalez NJ, Berry RM, Armitage JP. Composition, Formation, and Regulation of the Cytosolic C-ring, a Dynamic Component of the Type III Secretion Injectisome. PLOS Biol. 2015 Jan 15;13(1):e1002039.

46. Diepold A, Sezgin E, Huseyin M, Mortimer T, Eggeling C, Armitage JP. A dynamic and adaptive network of cytosolic interactions governs protein export by the T3SS injectisome. Nat Commun. 2017 Jun 27;8(1):1–12.

47. Akeda Y, Galán JE. Chaperone release and unfolding of substrates in type III secretion. Nature. 2005 Oct;437(7060):911–5.

48. Minamino T, Morimoto YV, Kinoshita M, Aldridge PD, Namba K. The bacterial flagellar protein export apparatus processively transports flagellar proteins even with extremely infrequent ATP hydrolysis. Sci Rep. 2014 Dec 22;4(1):7579.

49. Case HB, Dickenson NE. MxiN Differentially Regulates Monomeric and Oligomeric Species of the Shigella Type Three Secretion System ATPase Spa47. Biochemistry. 2018 Apr 17;57(15):2266–77.

50. Blaylock B, Riordan KE, Missiakas DM, Schneewind O. Characterization of the Yersinia enterocolitica Type III Secretion ATPase YscN and Its Regulator, YscL. J Bacteriol. 2006 May;188(10):3525–34.

51. Lara-Tejero M, Kato J, Wagner S, Liu X, Galán JE. A sorting platform determines the order of protein secretion in bacterial type III systems. Science. 2011 Mar 4;331(6021):1188–91.

52. Bzymek KP, Hamaoka BY, Ghosh P. Two Translation Products of Yersinia yscQ Assemble To Form a Complex Essential to Type III Secretion. Biochemistry. 2012 Feb 28;51(8):1669–77.

53. McDowell MA, Marcoux J, McVicker G, Johnson S, Fong YH, Stevens R, et al. Characterisation of Shigella Spa33 and Thermotoga FliM/N reveals a new model for C-ring assembly in T3SS. Mol Microbiol. 2016;99(4):749–66.

54. Lara-Tejero M, Qin Z, Hu B, Butan C, Liu J, Galán JE. Role of SpaO in the assembly of the sorting platform of a Salmonella type III secretion system. PLoS Pathog [Internet]. 2019 Jan 22 [cited 2019 Oct 21];15(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6358110/

55. Wimmi S, Balinovic A, Jeckel H, Selinger L, Lampaki D, Eisemann E, et al. Dynamic relocalization of cytosolic type III secretion system components prevents premature protein secretion at low external pH. Nat Commun. 2021 Mar 12;12(1):1625.

56. Rocha J, Corbitt J, Yan T, Richardson C, Gahlmann A. Resolving Cytosolic Diffusive States in Bacteria by Single-Molecule Tracking. Biophys J. 2019 May 21;116(10):1970–83.

57. Betzig E, Patterson GH, Sougrat R, Lindwasser OW, Olenych S, Bonifacino JS, et al. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. Science. 2006 Sep 15;313(5793):1642–5.

58. Hess ST, Girirajan TPK, Mason MD. Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy. Biophys J. 2006 Dec 1;91(11):4258–72.

59. Rust MJ, Bates M, Zhuang X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat Methods. 2006 Oct;3(10):793–6.

60. Small A, Stahlheber S. Fluorophore localization algorithms for super-resolution microscopy. Nat Methods. 2014 Mar;11(3):267–79.

61. Huang F, Hartwich TMP, Rivera-Molina FE, Lin Y, Duim WC, Long JJ, et al. Video-rate nanoscopy using sCMOS camera–specific single-molecule localization algorithms. Nat Methods. 2013 Jul;10(7):653–8.

62. Reymond L, Huser T, Ruprecht V, Wieser S. Modulation-enhanced localization microscopy. J Phys Photonics. 2020 Jul;2(4):041001.

63. Balzarotti F, Eilers Y, Gwosch KC, Gynnå AH, Westphal V, Stefani FD, et al. Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. Science. 2017 Feb 10;355(6325):606–12.

64. Gwosch KC, Pape JK, Balzarotti F, Hoess P, Ellenberg J, Ries J, et al. MINFLUX nanoscopy delivers 3D multicolor nanometer resolution in cells. Nat Methods. 2020 Feb;17(2):217–24.

65. Schmidt R, Weihs T, Wurm CA, Jansen I, Rehman J, Sahl SJ, et al. MINFLUX nanometer-scale 3D imaging and microsecond-range tracking on a common fluorescence microscope. Nat Commun. 2021 Mar 5;12(1):1478.

66. Eilers Y, Ta H, Gwosch KC, Balzarotti F, Hell SW. MINFLUX monitors rapid molecular jumps with superior spatiotemporal resolution. Proc Natl Acad Sci. 2018 Jun 12;115(24):6117–22.

67. Masullo LA, Steiner F, Zähringer J, Lopez LF, Bohlen J, Richter L, et al. Pulsed Interleaved MINFLUX. Nano Lett. 2021 Jan 13;21(1):840–6.

68. Marklund E, van Oosten B, Mao G, Amselem E, Kipper K, Sabantsev A, et al. DNA surface exploration and operator bypassing during target search. Nature. 2020 Jul;583(7818):858–61.

69. Enderlein J. Tracking of fluorescent molecules diffusing within membranes. Appl Phys B. 2000 Nov 1;71(5):773–7.

70. Levi V, Ruan Q, Kis-Petikova K, Gratton E. Scanning FCS, a novel method for three-dimensional particle tracking. Biochem Soc Trans. 2003 Oct 1;31(5):997–1000.

71. Berglund AJ, Mabuchi H. Tracking-FCS: Fluorescence correlation spectroscopy of individual particles. Opt Express. 2005 Oct 3;13(20):8069–82.

72. Hou S, Lang X, Welsher K. Robust real-time 3D single-particle tracking using a dynamically moving laser spot. Opt Lett. 2017 Jun 15;42(12):2390–3.

73. Hou S, Welsher K. An Adaptive Real-Time 3D Single Particle Tracking Method for Monitoring Viral First Contacts. Small. 2019;15(44):1903039.

74. Hou S, Exell J, Welsher K. Real-time 3D single molecule tracking. Nat Commun. 2020 Jul 17;11(1):3607.

75. Perillo EP, Liu YL, Huynh K, Liu C, Chou CK, Hung MC, et al. Deep and high-resolution three-dimensional tracking of single particles using nonlinear and multiplexed illumination. Nat Commun. 2015 Jul 29;6(1):7874.

76. Shechtman Y. Recent advances in point spread function engineering and related computational microscopy approaches: from one viewpoint. Biophys Rev. 2020 Nov 18;12(6):1303–9.

77. Huang B, Wang W, Bates M, Zhuang X. Three-dimensional Super-resolution Imaging by Stochastic Optical Reconstruction Microscopy. Science. 2008 Feb 8;319(5864):810–3.

78. Pavani SRP, Thompson MA, Biteen JS, Lord SJ, Liu N, Twieg RJ, et al. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. Proc Natl Acad Sci. 2009 Mar 3;106(9):2995–9.

79. Shechtman Y, Weiss LE, Backer AS, Sahl SJ, Moerner WE. Precise Three-Dimensional Scan-Free Multiple-Particle Tracking over Large Axial Ranges with Tetrapod Point Spread Functions. Nano Lett. 2015 Jun 10;15(6):4194–9.

80. Lelek M, Gyparaki MT, Beliu G, Schueder F, Griffié J, Manley S, et al. Single-molecule localization microscopy. Nat Rev Methods Primer. 2021 Jun 3;1(1):1–27.

81. Leake MC, Quinn SD. A guide to small fluorescent probes for single-molecule biophysics. Chem Phys Rev. 2023 Jan 13;4(1):011302.

82. Lew M, Lew* MD, von Diezmann* ARS, Moerner WE. Easy-DHPSF open-source software for three-dimensional localization of single molecules with precision beyond the optical diffraction limit. Protoc Exch [Internet]. 2013 Feb 25 [cited 2022 Apr 24]; Available from: http://www.nature.com/protocolexchange/protocols/2622

83. Rocha JM, Richardson CJ, Zhang M, Darch CM, Cai E, Diepold A, et al. Single-molecule tracking in live Yersinia enterocolitica reveals distinct cytosolic complexes of injectisome subunits. Integr Biol. 2018 Sep 17;10(9):502–15.

84. Paintdakhi A, Parry B, Campos M, Irnov I, Elf J, Surovtsev I, et al. Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. Mol Microbiol. 2016 Feb;99(4):767–77.

85. Michalet X, Berglund AJ. Optimal diffusion coefficient estimation in single-particle tracking. Phys Rev E. 2012 Jun 21;85(6):061916.

86. Persson F, Lindén M, Unoson C, Elf J. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. Nat Methods. 2013 Mar;10(3):265–9.

87. Heckert A, Dahal L, Tjian R, Darzacq X. Recovering mixtures of fast-diffusing states from short single-particle trajectories. Smal I, Akhmanova A, Paul MW, editors. eLife. 2022 Sep 6;11:e70169.

88. Karslake JD, Donarski ED, Shelby SA, Demey LM, DiRita VJ, Veatch SL, et al. SMAUG: Analyzing single-molecule tracks with nonparametric Bayesian statistics. Methods. 2021 Sep 1;193:16–26.

89. Kinz-Thompson CD, Ray KK, Gonzalez RL. Bayesian Inference: The Comprehensive Approach to Analyzing Single-Molecule Experiments. Annu Rev Biophys. 2021;50(1):191–208.

90. Stracy M, Lesterlin C, Garza de Leon F, Uphoff S, Zawadzki P, Kapanidis AN. Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid. Proc Natl Acad Sci. 2015 Aug 11;112(32):E4390–9.

91. Mohapatra S, Choi H, Ge X, Sanyal S, Weisshaar JC. Spatial Distribution and Ribosome-Binding Dynamics of EF-P in Live Escherichia coli. mBio. 2017 Jun 6;8(3):10.1128/mbio.00300-17.

92. Xiang L, Yan R, Chen K, Li W, Xu K. Single-molecule displacement mapping unveils sign-asymmetric protein charge effects on intraorganellar diffusion. Nano Lett. 2023 Mar 8;23(5):1711–6.

93. Lasker K, von Diezmann L, Zhou X, Ahrens DG, Mann TH, Moerner WE, et al. Selective sequestration of signalling proteins in a membraneless organelle reinforces the spatial regulation of asymmetry in Caulobacter crescentus. Nat Microbiol. 2020 Mar;5(3):418–29.

94. Saurabh S, Chong TN, Bayas C, Dahlberg PD, Cartwright HN, Moerner WE, et al. ATP-responsive biomolecular condensates tune bacterial kinase signaling. Sci Adv. 2022 Feb 16;8(7):eabm6570.

95. Prindle JR, Wang Y, Rocha JM, Diepold A, Gahlmann A. Distinct Cytosolic Complexes Containing the Type III Secretion System ATPase Resolved by Three-Dimensional Single-Molecule Tracking in Live Yersinia enterocolitica. Microbiol Spectr. 2022 Nov 10;10(6):e01744-22.

96. Yuan Y xiang. A Review of Trust Region Algorithms for Optimization. ICM99 Proc Fourth Int Congr Ind Appl Math. 1999 Sep 12;

97. Gad AG. Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review. Arch Comput Methods Eng. 2022 Aug 1;29(5):2531–61.

98. Stone M. Cross-Validatory Choice and Assessment of Statistical Predictions. J R Stat Soc Ser B Methodol. 1974;36(2):111–33.

99. Diepold A. Assembly and Post-assembly Turnover and Dynamics in the Type III Secretion System. In: Wagner S, Galan JE, editors. Bacterial Type III Protein Secretion Systems [Internet]. Cham: Springer International Publishing; 2020 [cited 2022 Apr 19]. p. 35–66. (Current Topics in Microbiology and Immunology). Available from: https://doi.org/10.1007/82_2019_164

100. Diepold A, Armitage JP. Type III secretion systems: the bacterial flagellum and the injectisome. Philos Trans R Soc B Biol Sci. 2015 Oct 5;370(1679):20150020.

101. Burkinshaw BJ, Strynadka NCJ. Assembly and structure of the T3SS. Biochim Biophys Acta BBA - Mol Cell Res. 2014 Aug 1;1843(8):1649–63.

102. Galán JE, Wolf-Watz H. Protein delivery into eukaryotic cells by type III secretion machines. Nature. 2006 Nov;444(7119):567–73.

103. Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants | Microbiology and Molecular Biology Reviews [Internet]. [cited 2022 Feb 15]. Available from: https://journals.asm.org/doi/full/10.1128/MMBR.62.2.379-433.1998

104. Hu B, Lara-Tejero M, Kong Q, Galán JE, Liu J. In Situ Molecular Architecture of the Salmonella Type III Secretion Machine. Cell. 2017 09;168(6):1065-1074.e10.

105. Kawamoto A, Morimoto YV, Miyata T, Minamino T, Hughes KT, Kato T, et al. Common and distinct structural features of Salmonella injectisome and flagellar basal body. Sci Rep. 2013 Nov 28;3(1):3369.

106. Nans A, Kudryashev M, Saibil HR, Hayward RD. Structure of a bacterial type III secretion system in contact with a host membrane in situ. Nat Commun. 2015 Dec 11;6(1):10114.

107. Butan C, Lara-Tejero M, Li W, Liu J, Galán JE. High-resolution view of the type III secretion export apparatus in situ reveals membrane remodeling and a secretion pathway. Proc Natl Acad Sci. 2019 Dec 3;116(49):24786–95.

108. Oikonomou CM, Jensen GJ. Electron cryo-tomography of bacterial secretion systems. Microbiol Spectr. 2019 Mar;7(2):10.1128/microbiolspec.PSIB-0019–2018.

109. Bernal I, Börnicke J, Heidemann J, Svergun D, Horstmann JA, Erhardt M, et al. Molecular Organization of Soluble Type III Secretion System Sorting Platform Complexes. J Mol Biol. 2019 Sep 6;431(19):3787–803.

110. Brubaker RR, Surgalla MJ. The Effect of Ca^{++} and Mg^{++} on Lysis, Growth, and Production of Virulence Antigens by Pasteurella pestis. J Infect Dis. 1964;114(1):13–25.

111. Straley SC, Bowmer WS. Virulence genes regulated at the transcriptional level by Ca2+ in Yersinia pestis include structural genes for outer membrane proteins. Infect Immun. 1986 Feb;51(2):445–54.

112. Michiels T, Wattiau P, Brasseur R, Ruysschaert JM, Cornelis G. Secretion of Yop proteins by Yersiniae. Infect Immun. 1990 Sep;58(9):2840–9.

113. Straley SC, Plano GV, Skrzypek E, Haddix PL, Fields KA. Regulation by Ca2+ in the Yersinia low-Ca2+ response. Mol Microbiol. 1993;8(6):1005–10.

114. Pettersson J, Nordfelth R, Dubinina E, Bergman T, Gustafsson M, Magnusson KE, et al. Modulation of Virulence Factor Expression by Pathogen Target Cell Contact. Science. 1996 Aug 30;273(5279):1231–3.

115. Elf J, Barkefors I. Single-Molecule Kinetics in Living Cells. Annu Rev Biochem. 2019;88(1):635–59.

116. Rocha J, Corbitt J, Yan T, Richardson C, Gahlmann A. Resolving Cytosolic Diffusive States in Bacteria by Single-Molecule Tracking. Biophys J. 2019 May 21;116(10):1970–83.

117. Mika JT, Van Den Bogaart G, Veenhoff L, Krasnikov V, Poolman B. Molecular sieving properties of the cytoplasm of Escherichia coli and consequences of osmotic stress. Mol Microbiol. 2010;77(1):200–7.

118. Mika JT, Schavemaker PE, Krasnikov V, Poolman B. Impact of osmotic stress on protein diffusion in Lactococcus lactis. Mol Microbiol. 2014;94(4):857–70.

119. Minamino T, Macnab RM. FliH, a soluble component of the type III flagellar export apparatus of Salmonella, forms a complex with FliI and inhibits its ATPase activity. Mol Microbiol. 2000;37(6):1494–503.

120. González-Pedrajo B, Fraser GM, Minamino T, Macnab RM. Molecular dissection of Salmonella FliH, a regulator of the ATPase FliI and the type III flagellar protein export pathway. Mol Microbiol. 2002;45(4):967–82.

121. Bernal I, Römermann J, Flacht L, Lunelli M, Uetrecht C, Kolbe M. Structural analysis of ligand-bound states of the Salmonella type III secretion system ATPase InvC. Protein Sci Publ Protein Soc. 2019 Oct;28(10):1888–901.

122. Okabe M, Minamino T, Imada K, Namba K, Kihara M. Role of the N-terminal domain of FliI ATPase in bacterial flagellar protein export. FEBS Lett. 2009 Feb 18;583(4):743–8.

123. Bai F, Morimoto YV, Yoshimura SDJ, Hara N, Kami-ike N, Namba K, et al. Assembly dynamics and the roles of FliI ATPase of the bacterial flagellar export apparatus. Sci Rep. 2014 Oct 6;4(1):6528.

124. Imada K, Minamino T, Uchida Y, Kinoshita M, Namba K. Insight into the flagella type III export revealed by the complex structure of the type III ATPase and its regulator. Proc Natl Acad Sci U S A. 2016 Mar 29;113(13):3633–8.

125. Muthuramalingam M, Whittier SK, Lovell S, Battaile KP, Tachiyama S, Johnson DK, et al. The Structures of SctK and SctD from Pseudomonas aeruginosa Reveal the Interface of the Type III Secretion System Basal Body and Sorting Platform. J Mol Biol. 2020 Dec 4;432(24):166693.

126. Kinoshita M, Namba K, Minamino T. A positive charge region of Salmonella FliI is required for ATPase formation and efficient flagellar protein export. Commun Biol. 2021 Apr 12;4(1):1–14.

127. Chen S, Beeby M, Murphy GE, Leadbetter JR, Hendrixson DR, Briegel A, et al. Structural diversity of bacterial flagellar motors. EMBO J. 2011 Jun 14;30(14):2972–81.

128. Wimmi S, Balinovic A, Jeckel H, Selinger L, Lampaki D, Eisemann E, et al. Dynamic relocalization of cytosolic type III secretion system components prevents premature protein secretion at low external pH. Nat Commun. 2021 Mar 12;12(1):1625.

129. Mika JT, Poolman B. Macromolecule diffusion and confinement in prokaryotic cells. Curr Opin Biotechnol. 2011 Feb;22(1):117–26.

130. Stracy M, Schweizer J, Sherratt DJ, Kapanidis AN, Uphoff S, Lesterlin C. Transient non-specific DNA binding dominates the target search of bacterial DNA-binding proteins. Mol Cell. 2021 Apr;81(7):1499-1514.e6.

131. Bellotto N, Agudo-Canalejo J, Colin R, Golestanian R, Malengo G, Sourjik V. Dependence of diffusion in Escherichia coli cytoplasm on protein size, environmental conditions and cell growth [Internet]. bioRxiv; 2022 [cited 2022 Mar 4]. p. 2022.02.17.480843. Available from: https://www.biorxiv.org/content/10.1101/2022.02.17.480843v1

132. Śmigiel WM, Mantovanelli L, Linnik DS, Punter M, Silberberg J, Xiang L, et al. Protein diffusion in Escherichia coli cytoplasm scales with the mass of the complexes and is location dependent. Sci Adv. 2022 Aug 12;8(32):eabo5387.

133. Johnson S, Blocker A. Characterization of soluble complexes of the Shigella flexneri type III secretion system ATPase. FEMS Microbiol Lett. 2008 Sep 1;286(2):274–8.

134. Kaniga K, Delor I, Cornelis GR. A wide-host-range suicide vector for improving reverse genetics in Gram-negative bacteria: inactivation of the blaA gene of Yersinia enterocolitica. Gene. 1991 Dec 20;109(1):137–41.

135. Biteen JS, Thompson MA, Tselentis NK, Bowman GR, Shapiro L, Moerner WE. Super-resolution imaging in live Caulobacter crescentus cells using photoswitchable EYFP. Nat Methods. 2008 Nov;5(11):947–9.

136. Thompson MA, Lew MD, Badieirostami M, Moerner WE. Localizing and Tracking Single Nanoscale Emitters in Three Dimensions with High Spatiotemporal Resolution Using a Double-Helix Point Spread Function. Nano Lett. 2010 Jan 13;10(1):211–8.

137. Wimmi S, Balinovic A, Brianceau C, Pintor K, Vielhauer J, Turkowyd B, et al. Cytosolic sorting platform complexes shuttle type III secretion system effectors to the injectisome in Yersinia enterocolitica. Nat Microbiol. 2024 Jan;9(1):185–99.

138. Diepold A, Sezgin E, Huseyin M, Mortimer T, Eggeling C, Armitage JP. A dynamic and adaptive network of cytosolic interactions governs protein export by the T3SS injectisome. Nat Commun. 2017 Jun 27;8(1):15940.

139. Zhang Y, Lara-Tejero M, Bewersdorf J, Galán JE. Visualization and characterization of individual type III protein secretion machines in live bacteria. Proc Natl Acad Sci. 2017 Jun 6;114(23):6098–103.

140. Schlumberger MC, Müller AJ, Ehrbar K, Winnen B, Duss I, Stecher B, et al. Real-time imaging of type III secretion: Salmonella SipA injection into host cells. Proc Natl Acad Sci U S A. 2005 Aug 30;102(35):12548–53.

141. Kubori T, Matsushima Y, Nakamura D, Uralil J, Lara-Tejero M, Sukhan A, et al. Supramolecular structure of the Salmonella typhimurium type III protein secretion system. Science. 1998 Apr 24;280(5363):602–5.

142. Makino F, Shen D, Kajimura N, Kawamoto A, Pissaridou P, Oswin H, et al. The Architecture of the Cytoplasmic Region of Type III Secretion Systems. Sci Rep. 2016 Sep 30;6(1):33341.

143. Gillespie DT. Stochastic simulation of chemical kinetics. Annu Rev Phys Chem. 2007;58:35–55.

144. Lecca P. Stochastic chemical kinetics. Biophys Rev. 2013 Jul 30;5(4):323–45.

145. Enninga J, Mounier J, Sansonetti P, Nhieu GTV. Secretion of type III effectors into host cells in real time. Nat Methods. 2005 Dec;2(12):959–65.

146. Mills E, Baruch K, Charpentier X, Kobi S, Rosenshine I. Real-time analysis of effector translocation by the type III secretion system of enteropathogenic Escherichia coli. Cell Host Microbe. 2008 Feb 14;3(2):104–13.

147. Ittig SJ, Schmutz C, Kasper CA, Amstutz M, Schmidt A, Sauteur L, et al. A bacterial type III secretion-based protein delivery tool for broad applications in cell biology. J Cell Biol. 2015 Nov 23;211(4):913–31.

148. Wölke S, Ackermann N, Heesemann J. The Yersinia enterocolitica type 3 secretion system (T3SS) as toolbox for studying the cell biological effects of bacterial Rho GTPase modulating T3SS effector proteins. Cell Microbiol. 2011;13(9):1339–57.

149. Ovesný M, Křížek P, Borkovec J, Švindrych Z, Hagen GM. ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. Bioinformatics. 2014 Aug 15;30(16):2389–90.

150. Kwapiszewska K, Szczepański K, Kalwarczyk T, Michalska B, Patalas-Krawczyk P, Szymański J, et al. Nanoscale Viscosity of Cytoplasm Is Conserved in Human Cell Lines. J Phys Chem Lett. 2020 Aug 20;11(16):6914–20.

151. Kalwarczyk T, Zięba cz N, Bielejewska A, Zaboklicka E, Koynov K, Szymański J, et al. Comparative Analysis of Viscosity of Complex Liquids and Cytoplasm of Mammalian Cells at the Nanoscale. Nano Lett. 2011 May 11;11(5):2157–63.

152. Elowitz MB, Surette MG, Wolf PE, Stock JB, Leibler S. Protein Mobility in the Cytoplasm ofEscherichia coli. J Bacteriol. 1999 Jan;181(1):197–203.

153. Mika JT, van den Bogaart G, Veenhoff L, Krasnikov V, Poolman B. Molecular sieving properties of the cytoplasm of Escherichia coli and consequences of osmotic stress. Mol Microbiol. 2010 Jul 1;77(1):200–7.

154. Mika JT, Poolman B. Macromolecule diffusion and confinement in prokaryotic cells. Curr Opin Biotechnol. 2011 Feb 1;22(1):117–26.

155. Xiang L, Chen K, Yan R, Li W, Xu K. Single-molecule displacement mapping unveils nanoscale heterogeneities in intracellular diffusivity. Nat Methods. 2020 May;17(5):524–30.

156. Watson JL, Juergens D, Bennett NR, Trippe BL, Yim J, Eisenach HE, et al. De novo design of protein structure and function with RFdiffusion. Nature. 2023 Aug;620(7976):1089–100.

157. Parry BR, Surovtsev IV, Cabeen MT, O'Hern CS, Dufresne ER, Jacobs-Wagner C. The Bacterial Cytoplasm Has Glass-like Properties and Is Fluidized by Metabolic Activity. Cell. 2014 Jan 16;156(1):183–94.

158. Branch RW, Sayegh MN, Shen C, Nathan VSJ, Berg HC. Adaptive remodelling by FliN in the bacterial rotary motor. J Mol Biol. 2014 Sep 23;426(19):3314–24.

159. Delalez NJ, Wadhams GH, Rosser G, Xue Q, Brown MT, Dobbie IM, et al. Signal-dependent turnover of the bacterial flagellar switch protein FliM. Proc Natl Acad Sci. 2010 Jun 22;107(25):11347–51.