Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Jiayue Sun 07/09/2020

Date

Structural and Functional Characterization of Extraintestinal Pathogenic *Escherichia coli* (ExPEC) PasTI complex

By
Jiayue Sun
Master of Science
Chemistry

Christine Dunham
Advisor

Christine Dunham
Committee Member

Stefan Lutz
Committee Member

Vincent Conticello
Committee Member

Accepted:
Lisa A. Tedesco, Ph.D. Dean of the James T. Laney School of Graduate Studies
07/09/2020

Structural and Functional Characterization of Extraintestinal Pathogenic *Escherichia coli* (ExPEC) PasTI complex

By Jiayue Sun

Advisor: Christine M. Dunham, Dr.

An abstract of
A thesis submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of Master of Science in Chemistry
2020

Abstract

Structural and Functional Characterization of Extraintestinal Pathogenic *Escherichia coli* (ExPEC) PasTI complex

By Jiayue Sun

PasTI appears to be a novel type II toxin-antitoxin (TA) module found in Extraintestinal Pathogenic *Escherichia coli* (ExPEC) which is critical for ExPEC cell survival in kidneys (Norton and Mulvey 2012). PasT is a putative toxin protein that causes ExPEC growth arrest, whereas PasI is PasT's cognate antitoxin which counteracts PasT's toxic activities under normal growth conditions. The *pasTI* gene causes and enhances ExPEC persister cell formation upon exposure to antibiotics, nutrient starvation, oxidative and nitrosative stress conditions (Norton and Mulvey 2012). Colonization of ExPEC in extraintestinal environments causes biofilm-like communities to form resulting in antibiotic tolerance and causing an array of chronic human diseases such as urinary tract infections and meningitis (Norton and Mulvey 2012, Wiles, Norton et al. 2013). *Escherichia coli* (*E. coli*) RatA (also known as YfjG toxin), is a homologue of PasT toxin and binds to the 50S ribosomal subunit to block its association with 30S subunit (Zhang and Inouye 2011). This blockage inhibits the 70S formation and protein synthesis (Zhang and Inouye 2011). Due to the high sequence identity between PasT and RatA, I hypothesize PasT functions in a similar manner to inhibit ribosomal subunit association. This characterization is important for determining how PasT toxin interferes with vital cellular processes, identifying the binding site of the PasT protein on the ribosome, and studying its function. I also propose to solve the structure of the PasT toxin, PasI antitoxin and PasTI complex using x-ray crystallography, and to understand the function of the PasTI system and decipher the interaction between PasT and its bound ribosomal subunit.

Structural and Functional Characterization of Extraintestinal Pathogenic *Escherichia coli* (ExPEC) PasTI complex

By Jiayue Sun

Advisor: Christine M. Dunham, Dr.

A thesis submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of Master of Science in Chemistry
2020

TABLE OF CONTENTS

Motivation

The discovery of antibiotics is a milestone in the early twentieth century, however, the effect of antibiotic treatments has decreased over time. Antibiotic exposure to bacteria will lead them to evolve into antibiotic-tolerant state and make antibiotics ineffective(Page and Peti 2016). Every year, over 2 million people in the United States are infected with antibiotic-resistant bacterial pathogens, and these infections cause nearly 23000 deaths every year (Centers for Disease Control and Prevention 2015). Persistence is a mechanism that upon exposure to antibiotics or other environmental stresses, causes a subpopulation of the bacteria to switch into a dormant, non-dividing state and can survive the stress (Lewis 2010). Thus, persisters can evade antibiotic treatment. Once relieved from the stress, the persisters can revert back to the actively growing state (Page and Peti 2016).

Extraintestinal pathogenic *Escherichia coli* (ExPEC) cells are facultative pathogens found in multiple human organs, and they live as commensals to the healthy *E. coli* populations (Kohler and Dobrindt 2011). ExPEC cells are the cause for the majority of urinary tract infections (UTIs), bacteremia, and neonatal meningitis. The PasTI chromosomal TA system was discovered in ExPEC strains in kidneys, and it was shown to enhance the survival and persistence of ExPEC in kidney infections (Norton and Mulvey 2012). Persister cells are commonly found in biofilms (Lewis 2010), and the established biofilm is not easily targeted by antibiotics. The PasTI system can be a new bacteria-specific target for the development of antibiotics to treat ExPEC infections caused by ExPEC biofilm formation. The goal of this project is to decipher the role of PasTI in gene regulation, persister cell and biofilm formation. This information will define the cellular functions of PasTI and provides insight to its biotechnological applications in clinical uses.

Background

Introduction of toxin-antitoxin systems

Bacterial toxin-antitoxin (TA) modules are abundant genetic elements composed of a stable toxin protein which inhibits cell growth by interfering with cellular processes, and a labile cognate antitoxin which counteracts the activity of toxin protein(Page and Peti 2016). Under normal growth conditions, antitoxins inhibit the toxin activity, while stress conditions lead to antitoxin degradation which frees toxin to inhibit essential cellular processes such as protein translation, cell division, and DNA replication (Page and Peti, 2016). TA systems are classified into five groups corresponding to the action and nature of cognate antitoxins (**Figure 1**). In type I and III, antitoxins are RNAs that regulates cellular levels of toxin protein, whereas in type II, IV, V, antitoxins are proteins that either sequester and counterbalance the activity of toxin protein, or directly inhibit the formation of toxin components (Goeders and Van Melderen 2014).

TA genes are abundant on bacterial chromosomes, but the composition and the number of TA loci vary between different bacterial species. For instance, *E. coli* K12 MG1655 strain encodes 19 type I TA loci, 13 type II TA loci, and 3 type IV TA loci while its relative *Salmonella enterica* Typhimurium encodes 6 type I TA loci and 21 type II TA loci (Harms, Brodersen et al. 2018). Although there is no clear correlation between the number of TA loci and environment conditions, it has been shown that high numbers of TA loci being activated upon exposure to unfavorable or hostile environments and promotes horizontal gene transfer (Leplae, Geeraerts et al. 2011).

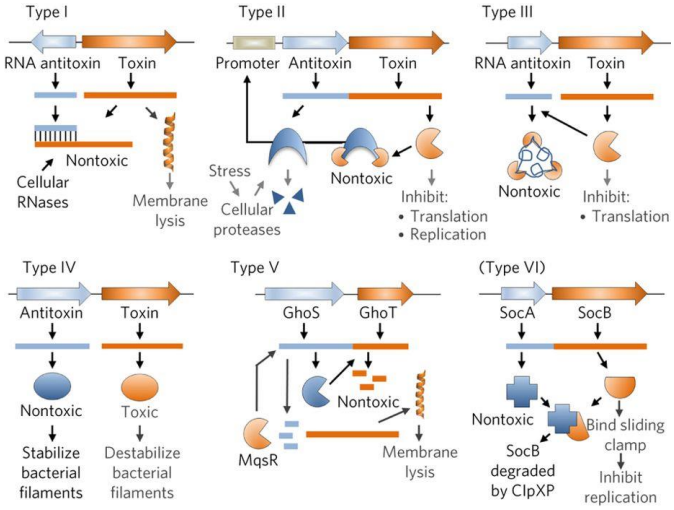


Figure 1. Types of TA systems. Toxins are shown in orange; antitoxin are shown in blue; nontoxic activities are shown in black font; toxic activities are shown in gray font. (Figure taken from Page and Peti, 2016)

TA systems have been studied for years, and they have shown biological importance in gene regulation, stress response, biofilm formation, and persister cell formation. There is no correlation that certain families of toxins would behave in similar biological functions (Harms, Brodersen et al. 2018). For example, the TacAT TA system, which is a type II TA module of *Salmonella enterica* Typhimurium, was shown to contribute to persister formation whereas its homologue model GmvAT in *Shigella* Sonnei, which has a high sequence similarity to TacAT, has been shown to be a potent post-segregational killing (PSK) module (Helaine and Kugelberg 2014, McVicker and Tang 2016). This

suggests that TA modules, even from the same family, may be adapted to divergent biological functions.

Major Functions of TA systems

There are three major biological functions of TA modules: post-segregational killing (PSK), abortive infection and persister formation (**Figure 2**). PSK causes cell death of plasmid-free cells and their descendent which can no longer produce labile antitoxin to counteract the toxin activity (Harms, Brodersen et al. 2018). Abortive infection is a mechanism where the activation of TA systems in innate bacteria kills bacteriophage-infected cells through altruistic suicide prior to phage replication (Harms, Brodersen et al. 2018). For example, type II TA module MazEF obstructs phage T4 infections (Alawneh, Qi et al. 2016). Persisters are subpopulations of genetic-identical, metabolically slow-growing cells that exhibit tolerance to antibiotics and environmental stress conditions such as nutritional starvation due to phenotypic transitions into a dormant state where antibiotics or stress have no effect (Page and Peti 2016, Harms, Brodersen et al. 2018). Different TA modules tend to influence bacteria at different growth phases. For example, YafQ facilitates persister formation in biofilms, whereas HipBA and RelBE were implicated in stationary phase (Keren, Shah et al. 2004). It was also proposed that the persister cell formation mediated by type II TA modules of *E. coli* is related to the second messenger guanosine tetra- or penta- phosphate ((p)ppGpp) and consequent Lon activation, for example, persister formation via type II TA module activation in *Salmonella enterica* Typhimurium is induced by phagocytosis of macrophages under starvation-induced (p)ppGpp signaling (Helaine and Kugelberg 2014). The importance of type II PasTI complex in persister formation was shown using animal infection and uropathogenic *E. coli* models (Norton and Mulvey 2012).

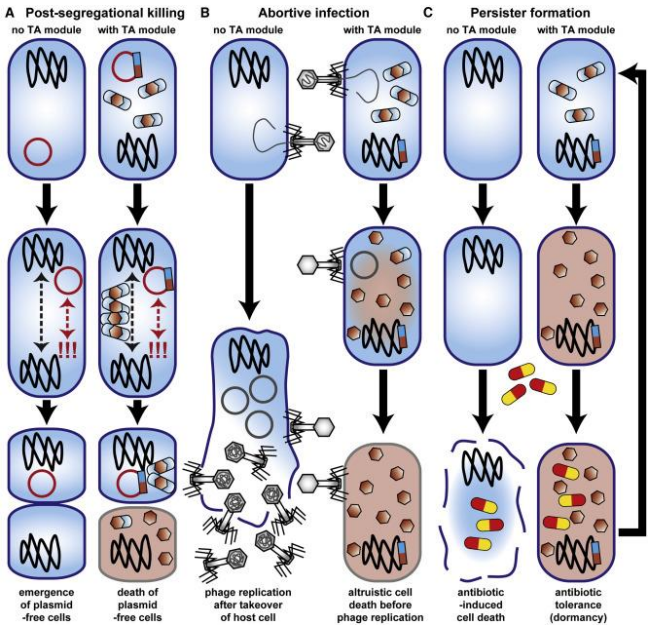


Figure 2. **Three major biological functions of TA modules.** Post-segregational killing (PSK) includes the activation of toxins to cause plasmid-free cell death, abortive infection includes activation of toxin to kill phage infected cells to prevent phage replication, and mechanism of persister cell formation with TA systems (Figure taken from Harms et al.2018)

Type II TA system

Type II antitoxins usually contains two domains: an N-terminal DNA-binding domain which is critical for transcriptional autoregulation, and a C-terminal domain which directly binds and inactivates the toxin (Chan, Espinosa et al. 2016). The proposed mechanisms of antitoxin inactivating toxins are either interfering with the toxins' catalytic domains or sterically hinder the toxins' binding to their cellular targets (Harms, Brodersen et al. 2018). In *E. coli*, most antitoxins are degraded by Lon proteases while some can be degraded by ClpP with the help of its adapter proteins ClpA or ClpX (Muthuramalingam, White et al. 2016). Antitoxins are typically labile because they lack functional protein secondary structures and their loop tails are the targets for proteases. The differential stability between toxin and its cognate antitoxin is critical to TA module's biological activities. For example, in RnlAB TA module, RnlA toxin has a half-life approximately ten times longer than that of its cognate antitoxin RnlB, so that during T4 phage infection, host gene expression has been shut off resulting in RnlB degradation and RnlA activation (Koga, Otsuka et al. 2011).

Prior research

The Mulvey lab discovered the PasTI system from ExPEC kidney infections in 2012 (Norton and Mulvey 2012). They found that low expression of PasT protects ExPEC from nutrient starvation, oxidative stress, and nitrosative stress (a nitric oxide-mediated nitrosylation of redox-sensitive thoils links to regulation of signal transduction, gene expression, cell growth and apoptosis (Na, Chung et al. 2008)), while overexpression of PasT causes bacterial death (Norton and Mulvey 2012). The toxicity of PasT is rescued by overexpression of PasI protein indicating PasTI is a type II TA system (Norton and Mulvey 2012). The Mulvey lab also showed that PasT toxicity and persister formation are linked to its N-terminal residues (Norton and Mulvey 2012). The PasT toxin protein was predicted to encode an oligoketide cyclase which catalyzes polyketide synthesis, while PasI was predicted to possess a ubiquitin-like β-grasp fold (Norton and Mulvey 2012) unlike any other type II TA modules. The *E. coli* RatA toxin binds to 50S ribosomal subunit and blocks the association with 30S subunit and leads to inhibition of 70S subunit formation and translation (Zhang and Inouye 2011). Initiation Factor 3 (IF3) binds to 30S ribosome and cause anti-association of 30S and 50S and will not cause the dissociation of 70S (Hirokawa, Nijman et al. 2005). Although RatA was shown to bind to 50S by western blot analysis of polysome profiles, RatA functions in a similar way to IF3 in anti-association of ribosomal subunits indicating there may be some functional overlaps between IF3 and RatA (Zhang and Inouye 2011). Due to the high sequence homology between PasT and RatA, we hypothesize PasT will act similarly on the ribosome.

Scientific/ Technical Approach

Specific Aim 1. Biochemical characterization of the PasTI system. It is hypothesized that PasT binds to the 50S ribosomal subunit of its highly homology to RatA. Biochemical assays including polysome profiling coupled with western blot assay and mass spectrometry will be used to determine what PasT interacts. Furthermore, whether overexpression of PasI antitoxin freeing PasT competes with normal translation will be tested, if so, how ribosomes react to PasT association and dissociation: recycled or still functions in protein synthesis. To better understand how PasTI functions within its host cells, we aim to solve the structures of PasT toxin protein, PasI antitoxin protein, and PasTI complex using x-ray crystallography.

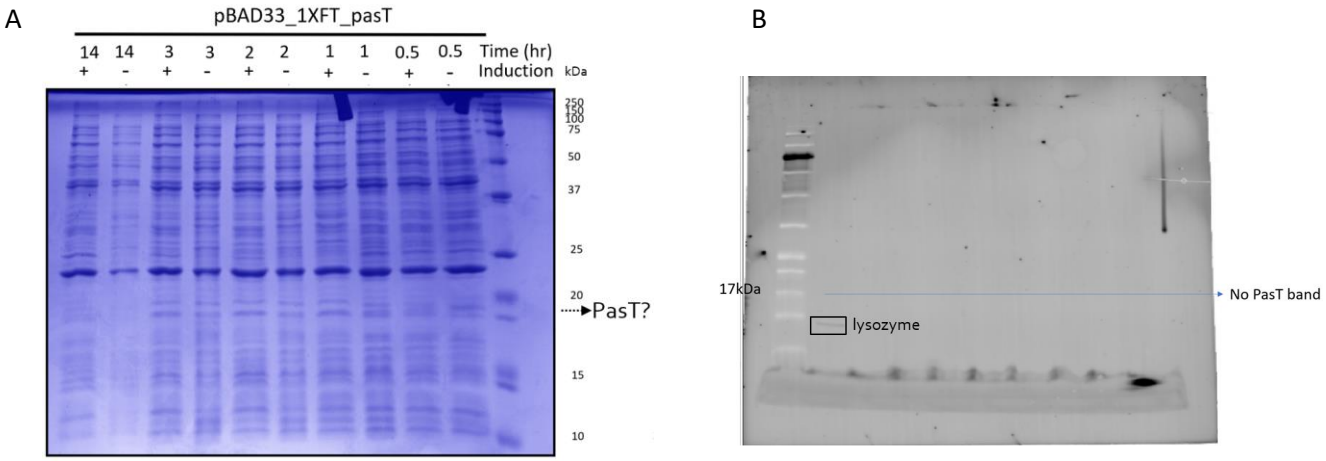
Specific Aim 2. **Defining the cellular functions of PasI**. PasI is predicted to be a ubiquitin-like protein (Norton and Mulvey 2012), and ubiquitylation is a common protein modification pathway that targets protein for proteolysis in eukaryotes (Burns and Darwin 2010). However, in prokaryotes, the prokaryotic

ubiquitin-like protein (Pup) protein in *Mycobacteria* functions analogously to ubiquitylation (Pearce, Mintseris et al. 2008) demonstrating there is potential protein modification mechanisms in Gram-negative bacteria system. Overexpression of PasTI-HA without protease treatment has shown ladder-like banding indicating PasTI-HA ot more likely PasI-HA may bind to different proteins. Although there is no evidence that shows PasI degrades proteins, its ubiquitin-like structure may implicate its function of tagging proteins and triggering proteolysis. PasI, a type II antitoxin, could be a putative target of the protease, thus if PasI binds to other proteins, proteolysis is more likely to happen. To test this, I will overexpress PasTI in *Δlon* and *ΔclpX, ΔclpP, ΔclpA E. coli* strains to determine if any specific proteases target PasI.

Results and Conclusions

PasT and PasTI cannot be expressed in bacterial system

Our lab has tried to overexpress PasT using pBAD18 vector, but there was no PasT protein seen. A possible reason for this is the overexpression of PasT causes cell death so that the live cells do not contain a detectable amount of PasT protein. Thus, I inserted the *pasT* gene into pBAD33 vector which is widely used for overexpressing various toxin and antitoxin proteins in our lab, with N-terminal FLAG tag which will prevent the toxicity of PasT to the *E. coli* cells (Norton and Mulvey 2012). Test samples were collected 1 hour and 3 hours after induction with 0.2% arabinose. The predicted 17.7 kDa PasT bands were not detected on the gel or on the western blot **(Figure 3A and 3B)**. The western blot setup and FLAG Tag antibody have been used widely and produced positive results on different TA systems in the lab, thus I excluded the system error of getting no bands of PasT. The remaining possible explanations for this issue are either no PasT expression using pBAD promoter or expressed PasT proteins were degraded in a short period of time by something unknown. Since the Mulvey lab has shown the overexpression of PasT causes bacteria growth arrest, I think it is more likely that PasT was degraded in the expression process, and it is not a stable protein. To test if PasTI can be overexpressed, I inserted PasTI into pET28a system. After induction with 0.5mM IPTG, test samples were collected at different time points. Neither PasT nor PasI were detected on the gel **(Figure 3C)**, indicating PasTI, as a complex, may not be stable as well.



C

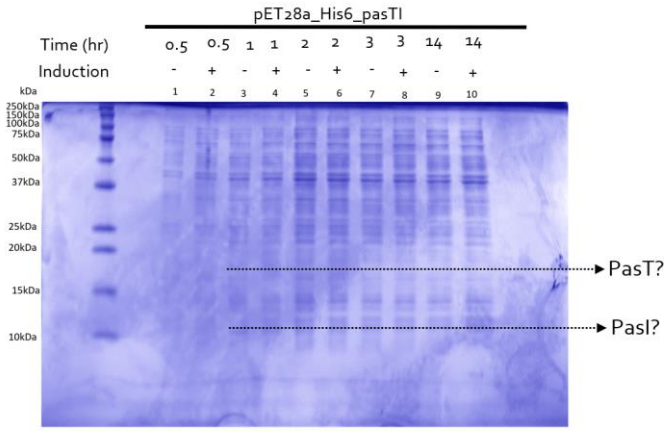


Figure 3. Induction trials of pBAD33_1XFT_PasT, pET28a_His6_PasTI. (A) DH5a competent cells were transformed with pBAD33_1XFT_PasT and incubated in M9+glycerol system, samples were collected at different time points for comparison. (B) western blot of induction trial of pBAD33_1XFT_PasT, lysozyme was used as a reference. (C) DH5a competent cells were transformed with pET28a_His6_PasTI and incubated in LB system, samples were collected at different time points for comparison.

Toward the functional characterization of the PasT protein

Since PasT could not be overexpressed and one possibility is that PasT was degraded, I inserted the *pasT* gene into pCOLD vector containing trigger factor (TF). This vector contains the cold shock protein A (*cspA*) promoter for expressing recombinant protein in *E. coli* with high yield and purity (Etchegaray and Inouye 1999) (**Supplementary Figure 1**). TF will be covalently bound to PasT to prevent PasT degradation after induction. TF is not only a protein tag that is hard to degrade, but also a protein folding chaperone that helps expressed PasT to fold into its native form (Haldar, Tapia-Rojo et al. 2017). TF tagged PasT was detected as a 75 kDa fusion protein on the gel (**Figure 4A and 4B**), and the fusion protein was confirmed using mass spectrometry. To separate PasT and TF, both thrombin and Factor Xa enzymes were tested to optimize the cleavage efficiency. Thrombin showed a higher cleavage efficiency than Factor Xa, and thus it was picked to cleave TF_PasT under 4°C for 72 hours. The cleaved PasT and TF mix was separated on HisTrap nickel column where the polyhistidine affinity-tagged trigger factor was expected to bind to the nickel resin and freed PasT will be directly eluted from the nickel column and saved in the direct-flowthrough. However, the collected direct-flowthrough portion contained both PasT and TF demonstrating that the interaction between PasT and TF has not been eliminated which outcompetes the polyhistidine affinity-tagged TF's binding affinity to HisTrap so that polyhistidine affinity-tagged trigger factor was not trapped on the nickel column . To disrupt this non-covalent interaction between PasT and TF, I tried three conditions: 1M NaCl, 1% 100X-Triton, and 6M guanidine-HCl, but none of these showed a well separation of PasT and TF (**Figure 5A and 5B**). The binding motif of TF to other proteins was identified as a stretch of basic, aromatic eight amino acids with a net positive charge (Petzalt et al. 2001). High salt will strengthen the hydrophobic interaction, and thus enhance both the TF and PasT interaction and polyhistidine affinity-tagged TF onto HisTrap. Nonionic detergents such as Triton or tween is often used in releasing intracellular materials such as protein-protein interaction without denaturing the protein (Johnson 2013). The reason for using X-100 Triton is to form micelles around PasT so that it can be freed from TF, however, X-100 Triton may also decrease the Histag binding affinity to the HisTrap. The use of 6M guanidine-HCl was intended to denature PasT and TF to break their intracellular protein-protein interaction. The detergent and

denaturing method eluted more PasT than the high salt condition did, thus I decided to incubate cleaved PasT and trigger factor mix with 1% X-100 Triton for 30min and 6M guanidine-HCl for 24 hours respectively before loading onto HisTrap. The pre-incubation with X-100 Triton did not elute more PasT than TF (**Figure 6C**), however, the pre-incubation with guanidine reduced the TF contamination in PasT elution (**Figure 6D**). Denatured PasT with little TF contamination was collected and refolded by dialysis and shock-refolding, but neither of the refolding methods successfully refolded PasT that PasT could be eluted at the 15 kDa~20 kDa range on size exclusion column.

Toward the functional characterization of the PasI protein

To overexpress and purify the PasI protein, I inserted the *pasI* gene into the same pCOLD vector with TF to overexpress TF_PasI fusion protein. TF_PasI fusion protein appeared as two bands. The top band showed up at 70 kDa, and the bottom band showed up at 60 kDa (**Figure 4A and 4B**). The identity of these two bands were analyzed by mass spectrometry (**Supplementary data 2**). Both bands had trigger factor and PasI components detected by mass spectrometry, but the lower band has fewer typsin-cleaved peptide detected than the top band indicating TF_PasI may be degraded.

Toward the functional characterization of the PasTI complex

To overexpress and purify the PasTI complex, I inserted the *pasTI* gene into the same pCOLD vector with TF to overexpress TF_PasTI. The TF_PasTI complex appeared as a TF_PasT band at 75 kDa, but the PasI band cannot be visualized at 10 kDa (**Figure 4A and 4B**). However, after purifying the total cell lysate on HisTrap, there is band appearing at 10kDa that may be PasI (**Figure 6A**). Since type II antitoxins typically do not covalently bind to the toxin protein, the interaction to its toxin will be broken by SDS detergent. Thus, TF_PasT band was expected to be seen instead of an 86 kDa large protein band where PasT and PasI are still tightly linked on SDS-PAGE. The fusion protein was cleaved by thrombin under 4°C for 72 hours, but to disrupt the interaction between TF and PasTI, I incubated the mix with 6M guanidine-HCl for 24 hours followed by HisTrap purification. The eluted sample has weak PasI band showing up at around 10 kDa and PasT with some TF contamination (**Figure 6B**). The eluted sample was collected and undergone dialysis and shock-refolding respectively to refold PasTI, but size exclusion did not show positive result of refolded PasTI.

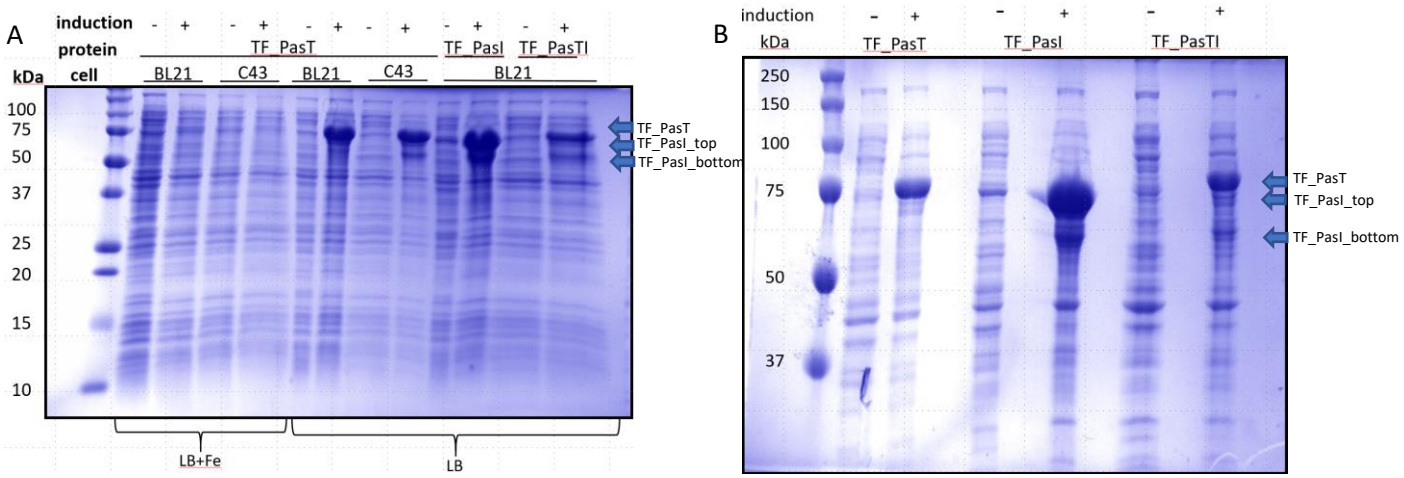


Figure 4. Induction trials of TF_PasT, TF_PasI and TF_PasTI. (A) left part shows the induction trials of TF_PasT supplemented with iron sources using *E. coli* BL21(DE3), BL21-C43(DE3) cells. Right part shows the induction trials of TF_PasT, TF_PasI, TF_PasTI in regular LB medium. The overexpression TF_PasI and TF_PasTI was using BL21(DE3) cells. The overexpression of TF_PasT was tested using BL21(DE3) and BL21-C43(DE3) cells. (B) To clarify two bands of TF_PasI, the samples were run on the SDS-PAGE for 100min.

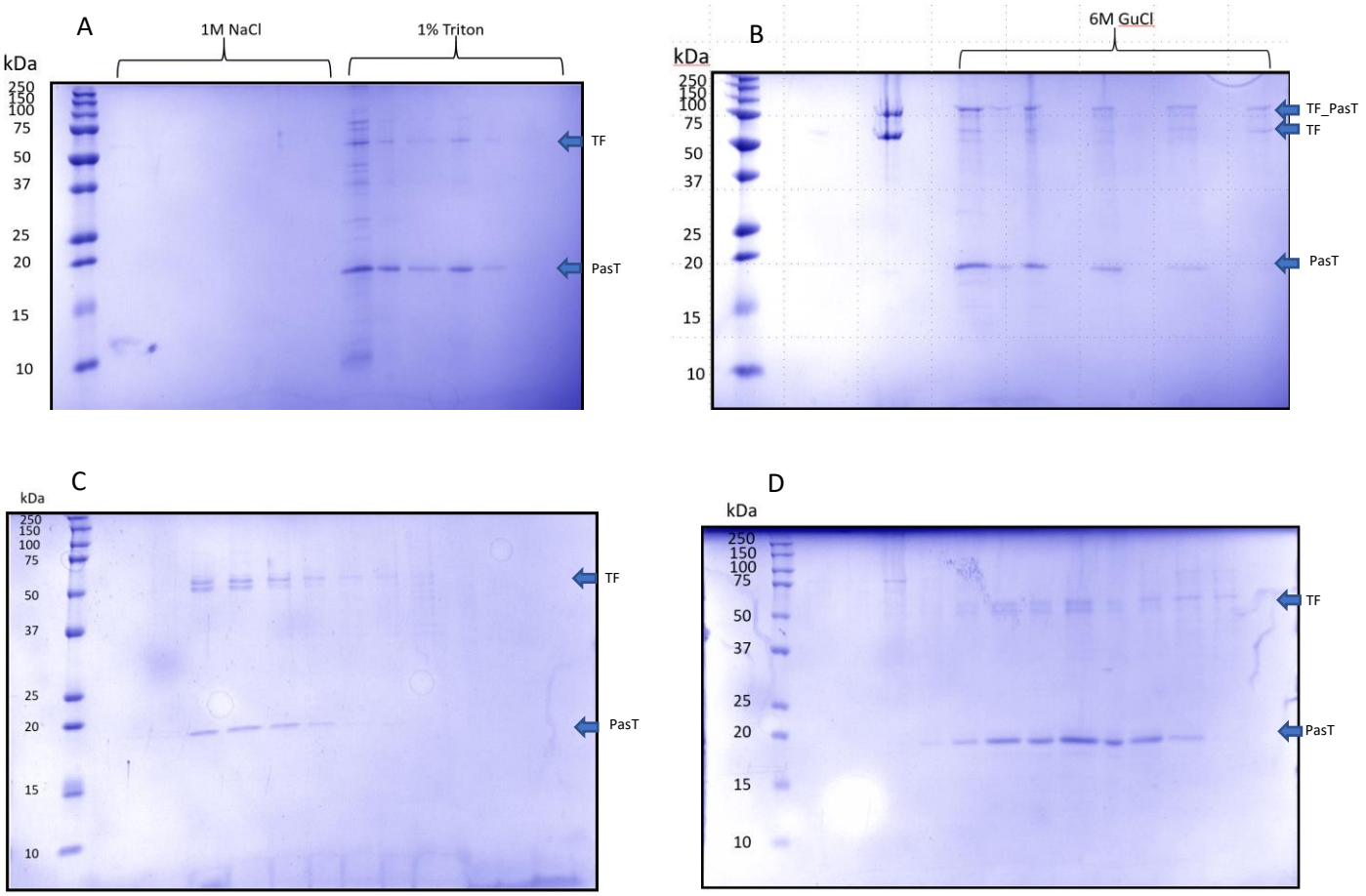


Figure 5. Purification of PasT. (A)Separation of TF and PasT using binding buffer supplemented with 1M NaCl and 1% X-100 Triton respectively. (B) Separation of TF and PasT using binding buffer supplemented with 6M Guanidine-HCl. (C) separation of TF and PasT after preincubation with 1% X-100 Triton. (D) Separation of TF and PasT after preincubation with 6M Guanidine-HCl for 24hrs. Akta data is shown in supplementary data (**Supplementary data 3**).

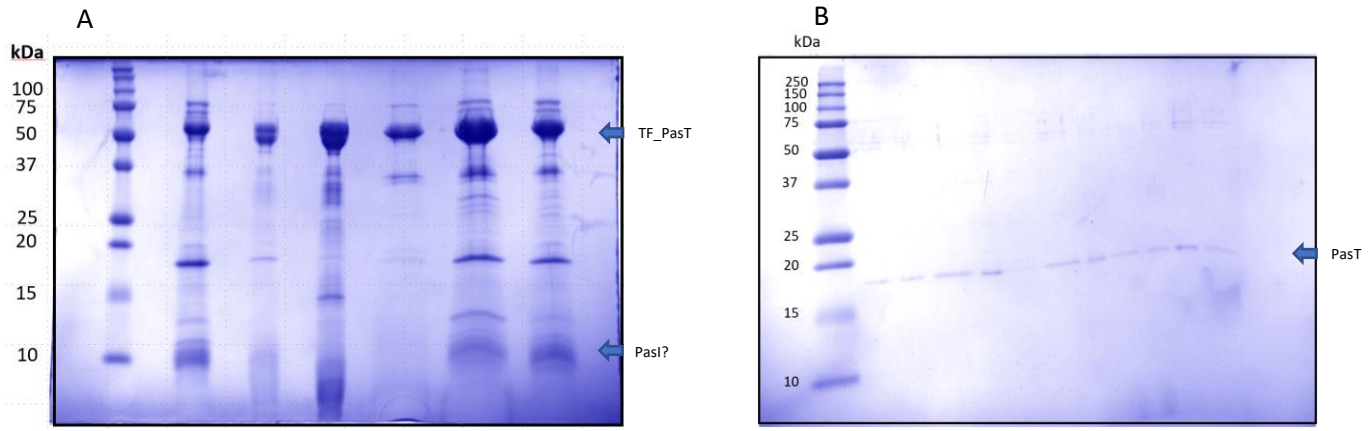


Figure 6. Purification of PasTI. (A)HisTrap purified cell lysate samples detect 10 kDa bands which may refer to PasI. (B) Separation of TF and PasTI after incubation with 6M Guanidine-HCl for 24hrs. Akta data is shown in supplementary data (**Supplementary data 3**).

Future Directions

Overexpression and purification of PasT, PasI, PasTI

Other than TF, ubiquitin-like SUMO protein tag is also widely used for recombinant protein purification (Lee, Sun et al. 2008). The Conn Lab from Biochemistry Department at Emory University kindly provided the commercially available pE_His6_SUMO vector which contains the pET system with T7 promoter. PasT or PasI or PasTI can be separated from polyhistidine affinity-tagged SUMO using commercially available SUMO protease and protease buffer. Thus, I plan to insert the *pasT, pasI, pasTI* genes into pE_His6_SUMO vector. Polyhistidine affinity-tagged SUMO can be trapped on nickel column to free PasT, PasI or PasTI. In comparison to trigger factor, SUMO is a smaller protein tag but also helps stabilize and increase the solubility of target protein (Panavas, Sanders et al. 2009). Due to the smaller size, SUMO tag is predicted to have less interaction with PasT or PasI or PasTI than TF does, thus it may result in an easier purification scheme. Other backup plans for PasT, PasI, PasTI purification is to use vectors containing other protein tags such as GST and MBP.

Protease recognition of PasI antitoxin

Type II antitoxins in general are degraded by proteases to free toxins, thus PasI is likely a target of certain protease(s). The Mulvey lab at the University of Utah has performed induction trials of PasTI with P*tac* promoter and C-terminal Human influenza hemagglutinin (HA) tag, and they found with the addition of protease inhibitors, PasI-HA appeared as a ladder of bands, however, with protease treatment there was only a single band. PasI is predicted to have a ubiquitin-like β-grasp fold and PasT is predicted to have oligoketide cyclase features (Norton and Mulvey 2012), PasT may function like a chaperon to promote the PasI's attachment to other proteins. Ubiquitin is a well-studied protein that targets proteins for degradation in eukaryotes (Burns and Darwin 2010), however, in prokaryotes, a protein modifier named Pup was found in *Mycobacterium tuberculosis* (Pearce, Mintseris et al. 2008). Pupylation works similarly to ubiquitylation in that Pup protein targets other proteins on lysine residues through its C-terminal region post-translationally and triggers proteolysis (Burns and Darwin 2010). PasI could be a potential "Pup" tagging proteins for their proteolysis in *E. coli*. I plan to transform *E. coli* BW25113 cells (*Δlon*, *ΔclpP, ΔclpX, ΔclpA*) with pET21c_PasTI_His6 construct to determine the target proteins that PasI recognizes. Further goal will be to analyze the correlation between PasT and PasI in pupylation or ubiquitylation.

PasTI native promoter(s) identification

The native promoter of PasTI system has not been clarified. I have tried online promoter-search tools, Promoterhunter and Bacpp to predict the potential PasTI promoter, and the results suggested PasTI does not have a general *E. coli* sigma factor. PasTI is a noncanonical type II TA system that the *pasT* toxin gene precedes the *pasT* antitoxin gene in the operon region. MqsRA, which is a type II TA system has the same reverse configuration in the operon region, was recently discovered two new MqsA promoters in the toxin *mqsR* gene region (Fraikin, Rousseau et al. 2019). Thus, PasI may also have more than 1 promoter. Better understanding the promoters of PasT and PasI will provide more insights into how transcriptional activity is regulated in this system. Most type II antitoxins include an N-terminal DNA binding domain and a C-terminal toxin binding domain (Page and Peti 2016), and antitoxins are likely to bind to its promoter(s) to process autoregulation of toxins. Thus, the PasI antitoxin protein may target more than one promoter to autoregulate transcriptional activity of the PasTI system.

Correlation between *E. coli* persister formation, biofilm formation and PasTI system

PasTI system has been shown to be critical to ExPEC survival in kidneys that, ExPEC increases its survival rate in mouse urinary tract infection models (Norton and Mulvey 2012). PasT enhances the ExPEC persister cell formation (Norton and Mulvey 2012). PasTI system may play an important role in regulating persister cell formation and biofilm formation. Previous researches have shown that *csgD*

gene is critical to curli-major component of biofilm expression (Soo and Wood 2013), and major curli subunit is generated by *csgA* gene (Fraikin, Rousseau et al. 2019). I hypothesize that PasT or PasI may recognize these genes or their promoters to either promotes or inhibit their transcriptional activity. I plan to overexpress PasT, PasI and PasTI complex to study their impact on E. coli biofilms using Crystal Violet assay and Congo Red stain assay.

Experimental Methods

PasT constructs

The *pasT* gene was amplified by PCR using primers oJS001_F and oJS002_R **(Supplementary Data 4)**. The PCR product and pBAD33_1XFT backbone were digested with PstI-HF and HindIII-HF (NEB) and agarose gel purified using Qiagen gel extraction kit followed by ligation using T4 ligase (NEB) overnight under room temperature. The ligated product was transformed into DH5α competent cells and plated on LB agar with 25 μg/mL chloramphenicol and grown overnight at 37◦C. *E. coli* DH5a cells were transformed with ligated product, and the transformants were picked for plasmid extraction using plasmid miniprep kit (Qiagen). Plasmid sequences were confirmed by Genewiz DNA plasmid Sanger sequencing.

Western blot analysis of PasT

E. coli BW25113 cells were transformed with pBAD33_1XFT_pasT and grown in LB medium with 34 μg/mL chloramphenicol. A 1:100 dilution was used to inoculate 25 mL of fresh LB medium containing 34 μg/mL chloramphenicol. Cells were grown to OD_{600nm} reached 0.2 and induced with 0.2% arabinose to express PasT. 10 mL culture was pelleted 1 hour and 3 hours post induction respectively using Eppendorf 5801R centrifuge for 10 min at 10000 rpm. The pellets were then resuspended with 1 mL lysis buffer (20 mM Tris-HCl 7.5, 250 mM KCl, 5 mM 2-mercaptoethanol (βMe), 0.1% X-100 Triton) followed by sonication using Misonix sonicator 3000 (1 min total time, 1s on, 1s off, output level 1) and then hard spin to get rid of cell debris. Supernatant was collected as soluble protein samples, pellet was saved. 1mg and 2 mg of samples were loaded onto 4-20% SDS-PAGE at 150V for 60 min. The gel was removed and washed three times in western transfer buffer (25 mM Tris-HCl, 192 mM Glycine, 10% methanol, pH 7.6) and equilibrated in western transfer buffer for 15 min. Then the gel was placed into blotting cassette and transferred to Immobilon-FL (Millipore) membranes at 100V for 60 min. Membranes were removed and washed with TBST (20mM Tris-HCl, 150mM NaCl, 0.2% Tween-20, pH7.6) for 5 mins and blocked with BLOTTO (5% blocking buffer, BioRad) for 2 hours under room temperature followed by anti-FLAG polyclonal antibody (Sigma F7425) incubation overnight at 4°C. The next day, the membranes were rinsed with TBST and incubated with DyLight 550 (ThermoFisher) for 2 hours under room temperature, then analyzed by Typhoon Trio (GE) using the settings of A_{max}=562nm, E_{max}=576nm, high sensitivity for green light; and A_{max}=633nm, E_{max}=670nm, normal sensitivity for red light, with PMT 750V. Images were obtained using ImageQuant (GE).

Construction of pCOLD_TF_pasT, pCOLD_TF_pasI, pCOLD_TF_pasTI

The *pasT* gene, *pasI* gene, *pasTI* gene were amplified by PCR using primers oJS003_F, oJS003_R oJS004_pasT_F, oJS004_pasI_F and oJS004_pasI_R respectively (Supplementary Data 3). The PCR product and pCOLD_TF vector were digested with NdeI and HindIII-HF (NEB) and agarose gel purified using Qiagen gel extraction kit followed by ligation using Quick Ligase (NEB) for 5 min under room temperature. *E. coli* DH5α competent cells were transformed with the ligated products and plated on LB agar with 100 μg/mL ampicillin and grown overnight at 37◦C. Transformants were picked for plasmid

extraction using Qiagen plasmid miniprep kit. Plasmid sequences were confirmed by Genewiz DNA plasmid Sanger sequencing.

Workflow of PasT, PasI, PasTI purification

The three purifications share the same workflow. *E. coli* BL21 competent cells were transformed with Pcold plasmids and plated on LB agar with 100 μg/mL ampicillin under 37◦C overnight. The transformant was picked to grow the overnight. The next day, 10mL overnight was added to fresh 990mL LB with 100ug/mL ampicillin and let it grow under 37C at 200rpm. After OD_{600nm} of the culture reached 0.6, the culture was cold-shocked in ice for 30min, then 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added. The culture was grown under 15◦C at 200 rpm for 24 hours. The culture was then spin down using Beckman JA20 rotor at 3.5k rpm, 4◦C for 30 min. Cell pellet was resuspended in 25 mL lysis buffer (20 mM Tris-HCl 7.0, 250 mM KCl, 5 mM βMe, 0.1% X-100 Triton) and lysed using Avestin Emulsiflex C5. Cell debris was removed by hard spin using Eppendorf 5801R centrifuge for 15min at 10000 rpm, 4◦C. Supernatant was retrieved and filtered by steriflip filter (Millipore). Filtered sample was loaded on 5 mL Ni HisTrap column on the AKTA purifier system. Desired fusion proteins were eluted with 500 mM imidazole. The fusion proteins were then cleaved with thrombin for 72 hours at 4◦C (1 unit of thrombin per mg of protein). The cleaved products were saved in -80◦C for further characterization.

References

Alawneh, A. M., D. Qi, T. Yonesaki and Y. Otsuka (2016). "An ADP-ribosyltransferase Alt of bacteriophage T4 negatively regulates the Escherichia coli MazF toxin of a toxin-antitoxin module." Mol Microbiol **99**(1): 188-198.

Burns, K. E. and K. H. Darwin (2010). "Pupylation versus ubiquitylation: tagging for proteasome-dependent degradation." Cell Microbiol **12**(4): 424-431.

Burns, K. E. and K. H. Darwin (2010). "Pupylation : A Signal for Proteasomal Degradation in Mycobacterium tuberculosis." Subcell Biochem **54**: 149-157.

Centers for Disease Control and Prevention. in *Antibiotic/Antimicrobial Resistance Vol.* (2015)

Chan, W. T., M. Espinosa and C. C. Yeo (2016). "Keeping the Wolves at Bay: Antitoxins of Prokaryotic Type II Toxin-Antitoxin Systems." Front Mol Biosci **3**: 9.

Etchegaray, J. P. and M. Inouye (1999). "CspA, CspB, and CspG, major cold shock proteins of Escherichia coli, are induced at low temperature under conditions that completely block protein synthesis." Journal of Bacteriology **181**(6): 1827-1830.

Fraikin, N., C. J. Rousseau, N. Goeders and L. Van Melderen (2019). "Reassessing the Role of the Type II MqsRA Toxin-Antitoxin System in Stress Response and Biofilm Formation: mqsA Is Transcriptionally Uncoupled from mqsR." mBio **10**(6).

Goeders, N. and L. Van Melderen (2014). "Toxin-antitoxin systems as multilevel interaction systems." Toxins (Basel) **6**(1): 304-324.

Haldar, S., R. Tapia-Rojo, E. C. Eckels, J. Valle-Orero and J. M. Fernandez (2017). "Trigger factor chaperone acts as a mechanical foldase." Nat Commun **8**(1): 668.

Harms, A., D. E. Brodersen, N. Mitarai and K. Gerdes (2018). "Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology." Mol Cell **70**(5): 768-784.

Helaine, S. and E. Kugelberg (2014). "Bacterial persisters: formation, eradication, and experimental systems." Trends Microbiol **22**(7): 417-424.

Hirokawa, G., R. M. Nijman, V. S. Raj, H. Kaji, K. Igarashi and A. Kaji (2005). "The role of ribosome recycling factor in dissociation of 70S ribosomes into subunits." RNA **11**(8): 1317-1328.

Johnson M (2013). "Detergents: Triton X-100, Tween-20, and More". *MATER METHODS*. **3**:163.

Keren, I., D. Shah, A. Spoering, N. Kaldalu and K. Lewis (2004). "Specialized persister cells and the mechanism of multidrug tolerance in Escherichia coli." J Bacteriol **186**(24): 8172-8180.

Koga, M., Y. Otsuka, S. Lemire and T. Yonesaki (2011). "Escherichia coli rnlA and rnlB compose a novel toxin-antitoxin system." Genetics **187**(1): 123-130.

Kohler, C. D. and U. Dobrindt (2011). "What defines extraintestinal pathogenic Escherichia coli?" Int J Med Microbiol **301**(8): 642-647.

Lee, C. D., H. C. Sun, S. M. Hu, C. F. Chiu, A. Homhuan, S. M. Liang, C. H. Leng and T. F. Wang (2008). "An improved SUMO fusion protein system for effective production of native proteins." Protein Sci **17**(7): 1241-1248.

Leplae, R., D. Geeraerts, R. Hallez, J. Guglielmini, P. Dreze and L. Van Melderen (2011). "Diversity of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis of novel families." Nucleic Acids Res **39**(13): 5513-5525.

Lewis, K. (2010). "Persister cells." Annu Rev Microbiol **64**: 357-372.

McVicker, G. and C. M. Tang (2016). "Deletion of toxin-antitoxin systems in the evolution of Shigella sonnei as a host-adapted pathogen." Nat Microbiol **2**: 16204.

Muthuramalingam, M., J. C. White and C. R. Bourne (2016). "Toxin-Antitoxin Modules Are Pliable Switches Activated by Multiple Protease Pathways." Toxins (Basel) **8**(7).

Na, H. J., H. T. Chung, K. S. Ha, H. Lee, Y. G. Kwon, T. R. Billiar and Y. M. Kim (2008). "Detection and measurement for the modification and inactivation of caspase by nitrosative stress in vitro and in vivo." Methods Enzymol **441**: 317-327.

Norton, J. P. and M. A. Mulvey (2012). "Toxin-antitoxin systems are important for niche-specific colonization and stress resistance of uropathogenic Escherichia coli." PLoS Pathog **8**(10): e1002954.

Page, R. and W. Peti (2016). "Toxin-antitoxin systems in bacterial growth arrest and persistence." Nat Chem Biol **12**(4): 208-214.

Panavas, T., C. Sanders and T. R. Butt (2009). "SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems." Methods Mol Biol **497**: 303-317.

Pearce, M. J., J. Mintseris, J. Ferreyra, S. P. Gygi and K. H. Darwin (2008). "Ubiquitin-like protein involved in the proteasome pathway of Mycobacterium tuberculosis." Science **322**(5904): 1104-1107.

Soo, V. W. and T. K. Wood (2013). "Antitoxin MqsA represses curli formation through the master biofilm regulator CsgD." Sci Rep **3**: 3186.

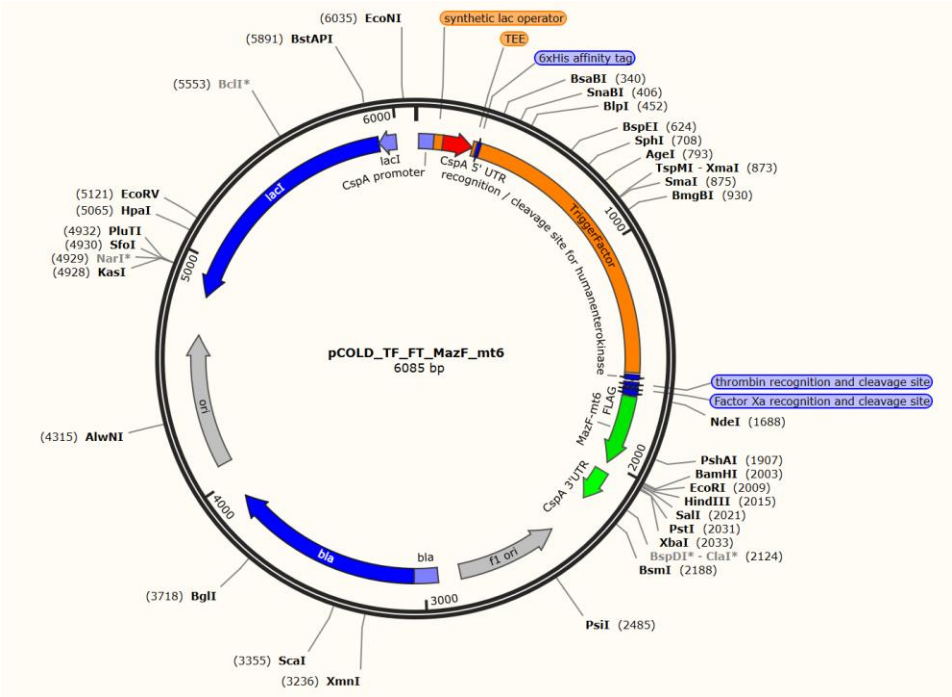
Wiles, T. J., J. P. Norton, S. N. Smith, A. J. Lewis, H. L. Mobley, S. R. Casjens and M. A. Mulvey (2013). "A phyletically rare gene promotes the niche-specific fitness of an E. coli pathogen during bacteremia." PLoS Pathog **9**(2): e1003175.

Patzelt H., Rudiger S., Brehmer D., Kramer G., Vorderwuibecke S., Schaffitzel E., Waitz A., Hesterkamp T., Dong L., Schneider-Mergener J., Bukau b., Deuerling E. (2001). Binding specificity of Escherichia coli trigger factor. PNAS. **98**(25):14244-14249.

Zhang, Y. and M. Inouye (2011). "RatA (YfjG), an Escherichia coli toxin, inhibits 70S ribosome association to block translation initiation." Mol Microbiol **79**(6): 1418-1429.

Supplementary Data

Supplementary Figure 1. pCOLD vector sequence map

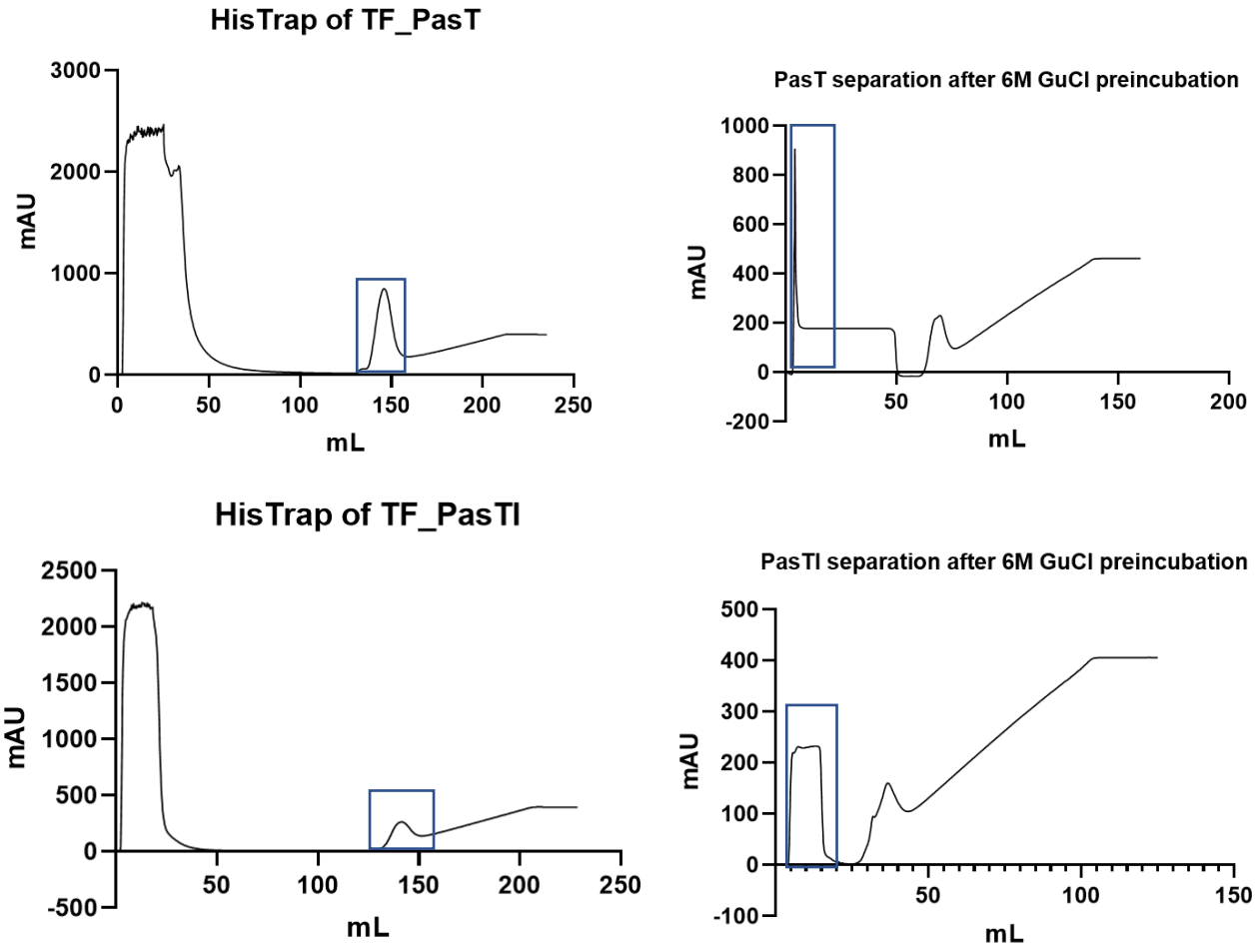


Supplementary Data 2. Mass Spectrometry result of TF_PasI

Checked	Protein FDI	Master	Accession	Description	Exp. q-valu	Sum PEP Sc	Coverage [%	# Peptides	# PSMs	# Unique P	# AAs	MW [kDa]	calc. pI	Score Sequ	# Peptides	Biological F
FALSE	High	Master Pro	P52119	UPF0125 p	0	58.362	75	10	202	10	96	10.8	9	383.05	10	
	Checked	Confidence	Annotated Sequence	Modificatic	Qvality PEI	Qvality q-v	# Protein G	# Proteins	# PSMs	Master Pro	Positions in	Modificatic	# Missed C	Theo. MH+	# Razor Qu	Abundance
	FALSE	High	[R].VTLQEGATVEEAIR.[A]		1.11E-07	0.001872	1	1	81	P52119	P52119 [2		0	1515.801	0	7.36E+08
	FALSE	High	[K].IAVEVAYALPEK.[Q]		1.54E-06	0.001872	1	1	44	P52119	P52119 [5-		0	1302.73	0	7.88E+08
	FALSE	High	[R].ASGLLELR.[T]		0.000527	0.001872	1	1	33	P52119	P52119 [3		0	858.5043	0	6.29E+08
	FALSE	High	[R].TDIDLTK.[N]		0.003661	0.001872	1	1	32	P52119	P52119 [4		0	805.4302	0	6.57E+08
	FALSE	High	[R].VEIYRPLIADPK.[E]		4.05E-06	0.001872	1	1	6	P52119	P52119 [7		0	1413.81	0	7.83E+08
	FALSE	High	[K].IAVEVAYALPEKQYLQR.[V]		1.77E-07	0.001872	1	1	2	P52119	P52119 [5-		1	1991.096	0	4766933
	FALSE	High	[K].VGIYSRPAK.[L]		0.107494	0.004547	1	1	1	P52119	P52119 [5		0	990.5731	0	571354.9
	FALSE	High	[R].VEIYRPLIADPKELR.[R]		0.008472	0.001872	1	1	1	P52119	P52119 [7		1	1812.038	0	146704.7
	FALSE	High	[R].TDIDLTKNK.[V]		0.013554	0.001872	1	1	1	P52119	P52119 [4		1	1047.568	0	851581.6
	FALSE	High	[R].ASGLLELRTDIDLTK.[N]		0.006064	0.001872	1	1	1	P52119	P52119 [3		1	1644.917	0	176173.2

Checked	Protein FDI	Master	Accession	Description	Exp. q-valu	Sum PEP Sc	Coverage [%	# Peptides	# PSMs	# Unique P	# AAs	MW [kDa]	calc. pI	Score Sequ	# Peptides	Biological F
FALSE	High	Master Pro	P52119	UPF0125 p	0	24.768	55	5	54	5	96	10.8	9	56.67	5	
	Checked	Confidence	Annotated Sequence	Modificatic	Qvality PEP	Qvality q-v	# Protein G	# Proteins	# PSMs	Master Pro	Positions in	Modificatic	# Missed Cl	Theo. MH+	# Razor Qu	Abundance
	FALSE	High	[R].VTLQEGATVEEAIR.[A]		3.19E-07	0.001875	1	1	41	P52119	P52119 [22		0	1515.801	0	25707227
	FALSE	High	[K].IAVEVAYALPEK.[Q]		1.67E-06	0.001875	1	1	7	P52119	P52119 [5-		0	1302.73	0	41204291
	FALSE	High	[R].ASGLLELR.[T]		0.001863	0.002371	1	1	2	P52119	P52119 [36		0	858.5043	0	13087308
	FALSE	High	[R].VEIYRPLIADPK.[E]		0.001065	0.001875	1	1	2	P52119	P52119 [72		0	1413.81	0	9635077
	FALSE	High	[R].TDIDLTK.[N]		0.033478	0.004121	1	1	2	P52119	P52119 [44		0	805.4302	0	14798992

Supplementary Data 3. Akta data: 5mL nickel column data



Supplementary Data 4. Primer list

Jiayue's primers	Date	Concentration	%GC content	Tm °	Sequence 5'~3'
oJS001_F	2019/6/23	100uM	40%	65.4°	CCCGGGGAACTGCAGCAATGATATTATTTGTTGGATTTTTGTTG
oJS001_R	2019/6/23	100uM	53.30%	63.5°	CCCAACTTTTACCTGGCACTGTAGACCTC
oJS003_F	2019/6/23	100uM	30.95%	51.99°	GGGAATTCCATATGATATTATTTGTTGGATTTTTGTTGATGG
oJS003_R	2019/6/23	100uM	53.57%	52.58°	CCCAAGCTTTTACCTGGCACTGTAGACC
oJS004_PasT_F	2019/7/22	100uM	30.90%	60.65°	GGGAATTCCATATGATATTATTTGTTGGATTTTTGTTGATGGAAATTGTTATGCC
oJS004_PasI_F	2019/7/22	100uM	50%	61.47°	GGGAATTCCATATGGTGCAAAAGAGGTCTACAGTGCCAGG
oJS004_PasI_R	2019/7/22	100uM	41.46%	60.6°	CCCAAGCTTTTATTTATTCGCTGATTTTTCTGCTCGTTGCC