



Western Michigan University
ScholarWorks at WMU

Masters Theses

Graduate College

8-2012

Catalytic Engineering of the Flagellin Protein

Alexandra M. Haase
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/masters_theses



Part of the Cell and Developmental Biology Commons

Recommended Citation

Haase, Alexandra M., "Catalytic Engineering of the Flagellin Protein" (2012). *Masters Theses*. 27.
https://scholarworks.wmich.edu/masters_theses/27

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Masters Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



CATALYTIC ENGINEERING OF THE FLAGELLIN PROTEIN

by
Alexandra M. Haase

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master of Science
Department of Biological Sciences
Advisor: Brian C. Tripp, Ph.D.

Western Michigan University
Kalamazoo, Michigan
August 2012

THE GRADUATE COLLEGE
WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MICHIGAN

Date July 2, 2012

WE HEREBY APPROVE THE THESIS SUBMITTED BY

Alexandra M. Haase

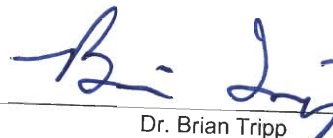
ENTITLED CATALYTIC ENGINEERING OF THE FLAGELLIN PROTEIN

AS PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF Master of Science

Biological Sciences

(Department)

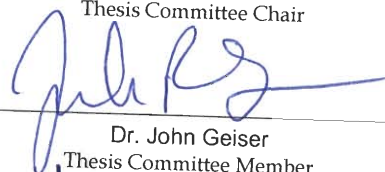


Dr. Brian Tripp

Thesis Committee Chair


Biological Sciences

(Program)



Dr. John Geiser

Thesis Committee Member



Dr. Todd Barkman

Thesis Committee Member

APPROVED



Dean of The Graduate College

Date August 2012

CATALYTIC ENGINEERING OF THE FLAGELLIN PROTEIN

Alexandra M. Haase, M.S.

Western Michigan University, 2012

Flagellin is the protein monomer that comprises the bacterial flagella for most bacteria, including *Salmonella typhimurium*. This protein has attracted attention for protein engineering because it is exported out of the cell, polymerized into stable fibers, is produced in large quantities and is relatively simple to purify. Using rational design with computer modeling, potential active sites in the flagellin protein structure were modeled after the human carbonic anhydrase II tetrahedral zinc binding site. In total, three locations were selected as potential active sites and the necessary mutations were successfully introduced. Flagella formation for the flagellin variants was demonstrated through TEM microscopy. When analyzed for metal ion content bound by the flagellin variants by ICP-ES, it appears that two of the variants have some level of zinc (II) association. One of these flagellin variants has diminished motility in motility agar that can be rescued upon the addition of the metal ion chelator EDTA and is less resistant to trypsin digestion upon the removal of metal ions with EDTA. Human carbonic anhydrase II can hydrolyze 4-NPA, however the flagellin variants have no detectable esterase activity. These preliminary results suggest that a metal binding site has successfully been introduced into bacterial flagellin while still retaining the ability of the protein to export, assemble and function as a motility organelle.

© 2012 Alexandra M. Haase

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Brian Tripp for all the suggestions and support over the years. I would also like to thank my committee members, Dr. Todd Barkman and Dr. John Geiser for their continued guidance. I would also like to thank the Biology Department, the Graduate College, and the STEM Workforce program for various funding.

I would like to thank Jacob Blanchard for all the help with my work in addition to all the encouragement. I would also like to thank Vanessa Revindran, Ph.D. for all the good advice. A special thanks to my family, especially my parents and Dana, for their continued support and for believing in me.

Alexandra M. Haase

TABLE OF CONTENTS

ACKNOWLEDGMENTS.....	ii
LIST OF TABLES.....	vi
LIST OF FIGURES.....	viii
CHAPTER	
1. REVIEW OF BACTERIAL FLAGELLIN AND CARBONIC ANHYDRASE PROTEIN STRUCTURE AND FUNCTION.....	1
1.1 Protein Engineering	1
1.2 Bacterial Flagellin Protein.....	4
1.3 Flagella and Bacterial Motility	6
1.4 Transition Metal Binding Sites in Proteins.	9
1.5 Flagellin Engineering	11
1.6 FliTrx Flagellin Peptide Display System	12
1.7 Flagellin Protein Engineering Difficulties	13
1.8 Carbonic Anhydrase	14
1.9 Carbonic Anhydrase Engineering	16
2. ENGINEERING A METAL BINDING SITE WITH THE POTENTIAL FOR CATALYTIC ACTIVITY INTO BACTERIAL FLAGELLIN.....	18
2.1 Introduction.....	18
2.1.1 Background and Rationale for Flagellin Engineering.....	18
2.1.2 Design of Catalytic Flagellin Mutations	21
2.2 Methods and Procedures	25
2.2.1 Media Preparation.....	25
2.2.2 Plasmid.....	26

Table of Contents-continued

CHAPTER

2.2.3 Site-Directed Mutagenesis of the <i>Salmonella</i> Flagellin Gene	27
2.2.4 <i>Salmonella</i> FliC Flagellin Expression and Purification	31
2.2.5 Carbonic Anhydrase II Expression and Purification	35
2.2.6 <i>In vivo</i> Motility Assay	36
2.2.7 Trypsin Digest of Flagellin Proteins.....	36
2.2.8 Circular Dichroism Spectroscopy	37
2.2.9 Transmission Electron Microscope Imaging	38
2.2.10 Inductively Coupled Plasma-Emission Spectrometry (ICP-ES) Analysis of Protein Metal Content	39
2.2.11 4-NPA Esterase Activity Assay.....	40
2.2.12 Sau Paulo Metallo β -Lactamase (SPM-1) Inhibitor Screening Assay	41
2.3 Results and Discussion	43
2.3.1 Flagellin Variant Protein Expression and Purification	43
2.3.2 <i>In vivo</i> Flagella Swarming Agar Motility Assay.....	45
2.3.3 Trypsin Digest Analysis of Flagellin Variant Stability	52
2.3.4 Circular Dichroism (CD) Spectroscopic Analysis of Flagellin Secondary Structure	56
2.3.5 Transmission Electron Microscope Images	58
2.3.6 Analysis of Flagellin Metal binding via ICP Metal Analysis .	60
2.3.7 4-NPA Esterase Assay	65

Table of Contents-continued

CHAPTER	
2.4 Bacterial Flagellin Engineering Main Conclusions	66
3. ENGINEERING THIOREDOXIN OF FLITRX TO PERFORM NUCLEOPHILIC CLEAVAGE OF 4-NPA ESTER SUBSTRATE	68
3.1 Background for Protein Engineering of FliTrx Catalytic Site	68
3.2 Methods and Procedures	70
3.2.1 Media Prep.....	70
3.2.2 Plasmid.....	70
3.2.3 Site-Directed Mutagenesis of Thioredoxin <i>trxA</i> Gene in pFliTrx Plasmid	71
3.3 FliTrx Engineering Conclusions	75
4. CATALYTIC ENGINEERING OF THE FLAGELLIN PROTEIN-FUTURE DIRECTIONS	77
4.1 FliTrx Engineering Future Directions	77
4.2 Flagellin Engineering Future Directions	78
APPENDIX	82
A. DNA Sequencing Results	82
B. Sequencing Primers.....	90
C. ICP-ES Protein-Metal Analysis Data.....	91
D. Recombinant DNA Approval	92
BIBLIOGRAPHY	93

LIST OF TABLES

2.1 Sequence Composition and Physical Properties of Wild-type <i>Salmonella typhimurium</i> Flagellin	19
2.2 Flagellin Amino Acid Sequence Comparison for Histidine and Cysteine in Various Species	20
2.3 CASTp Area and Volumes for Carbonic Anhydrase II, R-type and L-type Flagellins	22
2.4 <i>Salmonella</i> FliC Flagellin Site-Directed Mutagenesis Primers	28
2.5 PCR Reaction Volumes and Concentrations with Phusion™ Polymerase for Site-Directed Mutagenesis.....	28
2.6 PCR Thermocycler Settings for Site-Directed Mutagenesis on pTH890 Using Phusion™ Polymerase	29
2.7 Wild-type and Engineered Flagellin Variant Protein Extinction Coefficients and Theoretical Isoelectric Points (pI)	34
2.8 Predictions of Percent Alpha-Helix and Beta-Sheet of <i>Salmonella</i> FliC Protein Based on CD Spectroscopy Data Using K2D2 and K2D3 Prediction Software ..	58
2.9 Determining Metal:Protein Ratio With Zinc from ICP-ES Data.....	61
2.10 Determining Metal:Protein Ratio of Trace Metals from ICP-ES Data.....	64
2.11 4-NPA Hydrolysis Assay Percent Change of FliC Mutants Compared to CA2 and WT	65
3.1 FliTrx Site-Directed Mutagenesis Primers	72
3.2 PCR Reaction Volumes and Concentrations with Phusion™ Polymerase for Site-Directed Mutagenesis.....	73
3.3 PCR Thermocycler Settings for Site-Directed Mutagenesis on pFliTrx Using Phusion™ Polymerase	73
3.4 Colony PCR Reagent Volumes and Concentrations	74
3.5 Colony PCR Thermocycler Settings	74

LIST OF FIGURES

1.1 X-ray analysis of metal binding sites in a trypsin variant and human carbonic anhydrase II.....	3
1.2 Structure of <i>Salmonella typhimurium</i> flagellin protein and the flagellum fiber	5
1.3 L-type and R-type FliC ribbon structures superimposed	7
1.4 Schematic diagram of the bacterial flagellum apparatus for <i>Salmonella typhimurium</i>	8
1.5 Tetrahedral binding geometry of zinc	10
1.6 Representation of 4-nitrophenyl acetate hydrolysis	15
1.7 Computer modeling of human carbonic anhydrase II structure	16
2.1 Comparison of active-site-like solvent accessible pockets identified in carbonic anhydrase II, R-type straight flagellar filament and L-type straight supercoiled conformations of <i>Salmonella</i> FliC flagellin by the CASTp software	22
2.2 Computer modeled images of <i>Salmonella typhimurium</i> flagellin protein structure and the locations of rationally designed metal binding mutations..	24
2.3 Computer generated image of the pTH890 plasmid	27
2.4 Illustration of mechanical shearing of flagella fibers.....	33
2.5 Example of agarose gel electrophoresis of purified pTH890 plasmid DNA encoding several mutations in the <i>fliC</i> gene	43
2.6 SDS-PAGE gel demonstrating export of flagellin variants	44
2.7 Motility agar assay of <i>S. typhimurium</i> flagellin variant	46
2.8 Motility of full and partial pTH890 variants in SJW134	47
2.9 Motility of wild-type and flagellin variants in the presence of metals or EDTA	49
2.10 KNT flagellin variant motility as a function of pH, in the presence of EDTA and water	50

List of Figures - continued

2.11 Mechanical shearing of SJW134 cells used in partial mutant motility assay .	51
2.12 SDS-PAGE analysis of trypsin digests of polymeric WT and metal-site variant flagella fibers	52
2.13 SDS-PAGE analyses of trypsin digest of flagellin monomers with and without EDTA.....	54
2.14 CD spectra of the three flagellin variants compared with wild-type	56
2.15 TEM images of bacteria and flagella	59
2.16 Ribbon structure of the flagellin protein highlighting all the His, Asp and Glu residues	62
3.1 Computer generated image of the pFlITrx plasmid	71

CHAPTER 1

REVIEW OF BACTERIAL FLAGELLIN AND CARBONIC ANHYDRASE PROTEIN STRUCTURE AND FUNCTION

1.1 Protein Engineering

Protein engineering began in the 1980's by using site-directed mutagenesis to determine how changes in functional groups could alter enzymatic activity¹.

Mutagenesis of proteins evolved from a few amino acid changes to entire loops and protein domains. Nanobiocatalysis involves combining catalytic activity with a nanostructure, which can be useful for applications in industry and medicine^{2,3}. Bacterial flagella are a type of biologically produced, protein-based nanostructure that can be used for engineering new functions. Bacterial flagellins make an excellent protein engineering scaffold because they are exported out of the cell, can be produced in large quantities, self-assemble into long, homogenous fibers⁴, can withstand large deletions in the middle hypervariable region of the gene⁵, have a determined structure with known domain functions⁶ and are easy to purify.

Using an existing protein as a scaffold for enzyme engineering has many benefits over designing a protein *de novo*. Compared to *de novo* scaffolds, existing protein scaffolds tend to be much more stable, as the tertiary protein fold has already been arrived at by evolutionary selection pressures and is typically over-encoded by the primary sequence⁷. Nature has reused the same protein folds for many different functions, as indicated by the limited number of protein folds observed in globular

proteins and protein domains via structural studies that are deposited in the Protein Data Bank, and categorized into similar families in several online protein fold databases such as CATH⁸ (<http://www.cathdb.info>), SCOP⁹ (<http://scop.mrc-lmb.cam.ac.uk/scop/>), and FSSP (<http://www1.icsg.org/fss/help/document.html>). Existing protein fold scaffolds typically have the increased ability to withstand amino acid substitutions, insertions and deletions, are well characterized, easy to purify in high yield and generally more stable as compared to *de novo* proteins, while typically retaining some of the wild-type characteristics⁷.

There are two common ways to engineer enzymes to perform desired catalytic functions; directed evolution and rational design. Directed evolution is achieved by using error-prone PCR to produce genes encoding modified enzymes, followed by screening or selection for the desired catalytic activity of interest. This requires no need for structural data on the protein and produces many variants¹⁰. The other approach, rational design, requires detailed knowledge of the amino acid sequence and tertiary/quaternary structure of the protein of interest. Rational design involves selecting a few logical sites to introduce selected mutations using site-directed mutagenesis. This, in theory, should produce a fewer number of, and higher quality variants¹¹. Rational design methods for enzyme engineering have been met with some success in the past, which demonstrates the possibility of introducing a metal binding site into an existing scaffold. Some examples include a metal binding antibody^{12,13}, trypsin variants that bind copper (II), nickel (II) and zinc (II) ions^{14,15}, “histidine-patch”

thioredoxin with a nickel binding surface region¹⁶, zinc (II) ion binding sites in charbdoxin^{17,18}, and in mammalian serum retinal-binding protein¹⁹.

When modeling mutations for an active site into a protein based on another, most of the locations do not exactly match the model structure. However, flexibility within a protein to accommodate metal ions has been demonstrated through protein structure analysis²⁰. When the apo- *versus* holoenzyme X-ray diffraction structures of a metal-binding trypsin mutant were compared to one another, a considerable amount of local conformational change was observed as the metal-binding region side-chains of the protein moved into the correct orientation for metal binding (Figure 1.1A). When the holoenzyme structure was compared to the human carbonic anhydrase II (CA2)

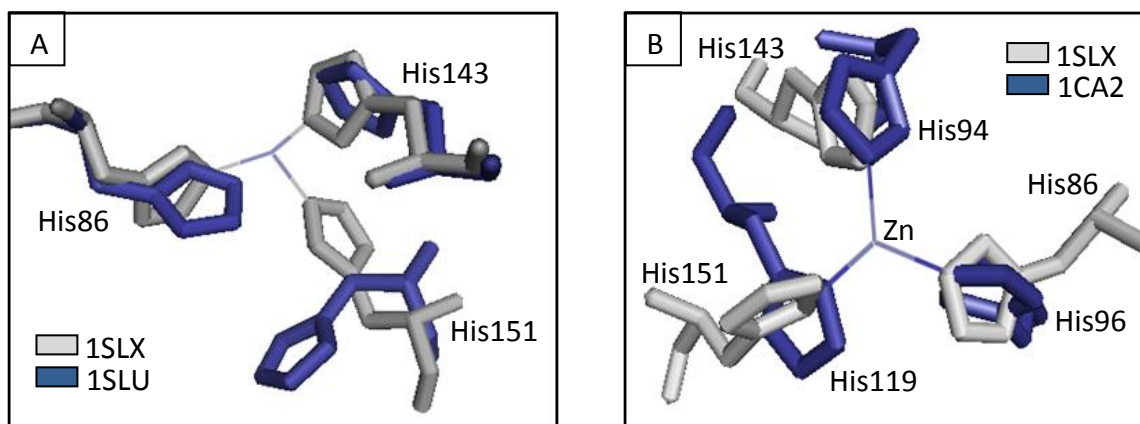


Figure 1.1

X-ray analysis of metal binding sites in a trypsin variant and human carbonic anhydrase II. (A) Overlay of trypsin metal-binding variant with and without zinc. The light colored image is the trypsin variant with zinc bound and the dark colored image is the trypsin variant without zinc bound. (B) Overlay of x-ray structures of zinc binding trypsin variant and CA2. The darker image is CA2 and the lighter image is the trypsin mutant variant. Trypsin variant crystal structures solved by Brinen *et al.*²⁰. Images were generated using PyMOL using PDB ID 1CA2 for carbonic anhydrase II, 1SLU for apo- and 1SLX for holo-trypsin with zinc.

crystal structure, which served as the model for the metal binding site, they observed some residue positioning differences, which demonstrates some plasticity in the sites selected for mutagenesis (Figure 1.1B)²⁰. Therefore, the goal of this project was to design and introduce a transition metal binding site into the wild-type *Salmonella* FliC flagellin that could potentially function as a catalytic center.

1.2 Bacterial Flagellin Protein

The bacterial flagellum fiber is composed upwards of 20,000 of flagellin protein (FliC, FljB) subunits on motile strains of *Salmonella typhimurium* (*S. typhimurium*) and *Escherichia coli* (*E. coli*)⁴. Each flagellum fiber is approximately 200 Å in diameter and can be up to 15 µm long²¹. The flagellum fiber grows at the distal tip, furthest from the cell body, as new flagellin proteins are added with the help of the chaperone cap protein, FliD²². The complete structures have been determined for *Salmonella* flagellin in two different conformations, the R-type (PDB ID 1IO1, 1UCU)^{6,23} and more recently, the L-type (PDB ID 3A5X)²⁴. These structures can be found in the Protein Data Bank (<http://www.rcsb.org/pdb/home/home.do>). The central core of the flagellum is hollow with a diameter of approximately 20 Å. FliC is transported through the central channel partially folded and only folds into the final conformation once it reaches the end of the fiber^{25,26}. FliC is a 494 amino acid protein composed of four domains; D0, D1, D2, and D3 (Figure 1.2B). When compared to flagellins from other bacterial species, the amino acid sequences of the D0 and D1 domains are highly conserved. This highly conserved region

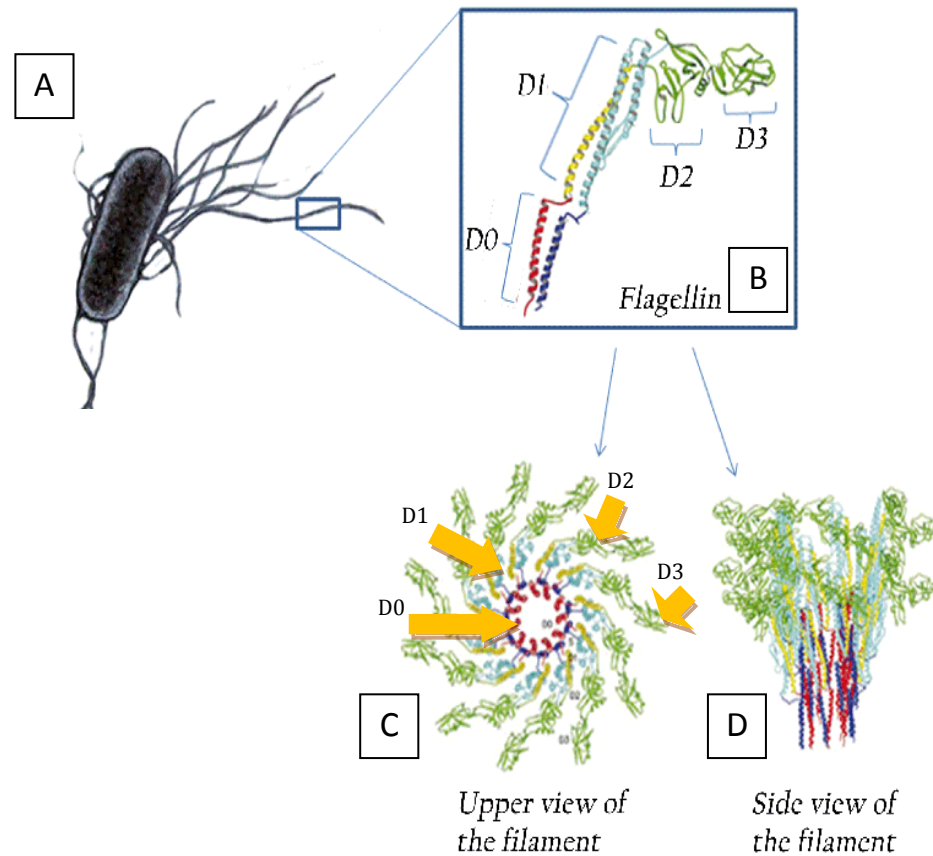


Figure 1.2

Structure of *Salmonella typhimurium* flagellin protein and the flagellum fiber. (A) Illustration of peritrichous bacterium. (B) The ribbon diagram depicts the flagellin monomer that composes the bacterial flagellum with all four domains labeled. (C) The ribbon diagram of the flagellin monomer polymerized as you would look down the length of the fiber. All four domains are labeled. (D) The ribbon diagram of the flagellin monomer upon polarization as you would look at it lengthwise. Colors of domains in picture B correspond to the same regions in C and D. Image adapted from <http://2008.igem.org/Team:Slovenia/Background/Flagellin>.

consists of approximately 140 amino acids at the N-terminus and approximately 40 amino acids at the C-terminus²⁷. The D0 domain, which is composed of the N- and C-termini of the protein, is essential for self-assembly into flagella fibers²⁸, and the C-terminus is responsible for protein export²⁹. The innate immune system recognizes the conserved N-terminal region via interactions with the Toll-like receptor 5, making the monomer more antigenic than the polymer³⁰. The polymerized FliC structure analysis indicates that the D0 and D1 domains stack on top of each other in the central channel of the fiber in an 11 protofilament interaction pattern (Figure 1.2C and D)^{6,23}. The curliness of the protofilament is determined by the ratio of the two distinct conformations, L- or R-type, with all 11 protofilaments of one conformation resulting in a straight fiber⁶. In contrast, the D3 domain is solvent accessible on the surface of the flagellin protein and the flagellar filament. The D3 domain amino acid sequence is highly variable across bacterial species²⁷, and can withstand large deletions^{5,28} or can be completely deleted while not eliminating the ability of flagellin to self-assemble into flagella fibers *in vivo* or *in vitro*. When the two FliC conformations are superimposed, D0 and part of D1 are shifted; however, the D2 and D3 domains remain relatively unchanged (Figure 1.3).

1.3 Flagella and Bacterial Motility

Many species of bacteria use flagella to provide a means of propulsion in semi-liquid/liquid media by rotating the fibers^{31,32}. Some species, such as *S. typhimurium* and *E. coli* are peritrichious, which means that they have flagella located at random positions

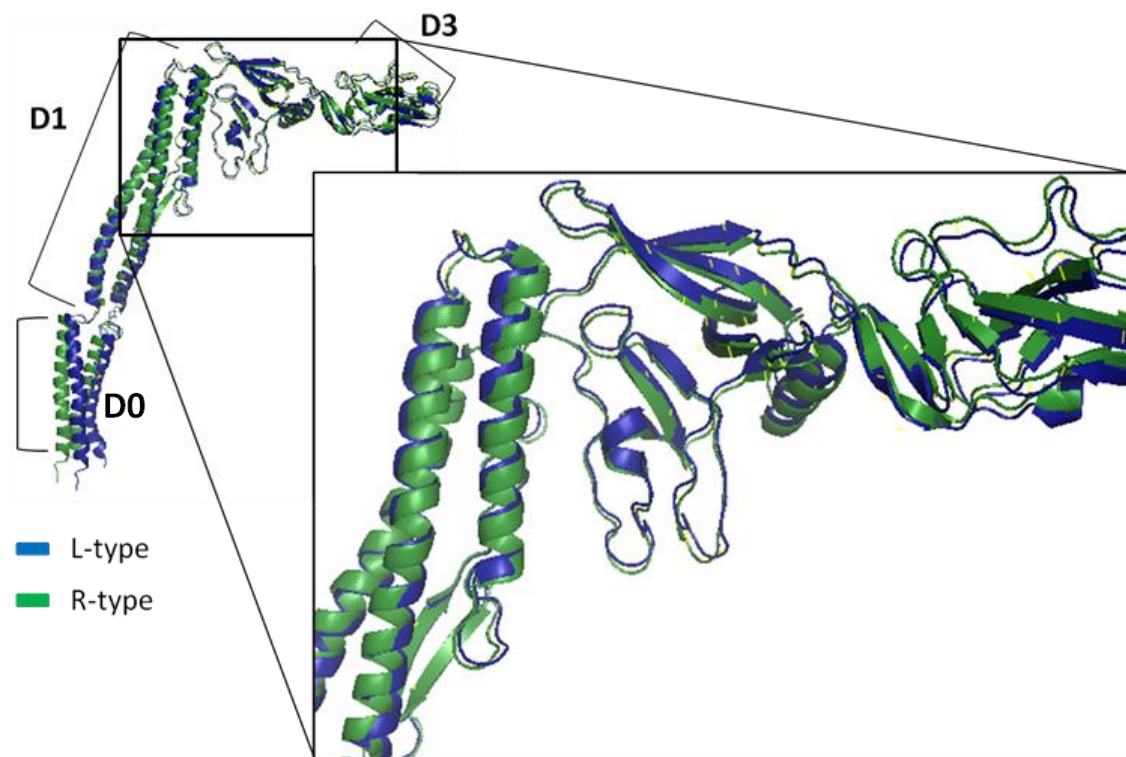


Figure 1.3

L-type and R-type FliC ribbon structures superimposed. Both the L- and R-type FliC are very similar with just a slight shift observed in the D0 and D1 domains. D2 and D3 show very little change in overall fold. Images generated using PyMOL software using PDB files 1UCU and 3A5X.

all over the surface of the cell body. The flagellum apparatus is comprised of a joined inner and outer membrane pore with a transmembrane proton potential motor that rotates the extracellular filament, which is comprised of the hook, flagellin and the cap protein⁴. Figure 1.4 is a representation of the flagellum apparatus showing in detail all the different parts and proteins that make up the entire system. The flagellum apparatus is a modified type III secretion system with which it shares many structural features⁴. For *Salmonella*, when the individual motors spin the flagella in a

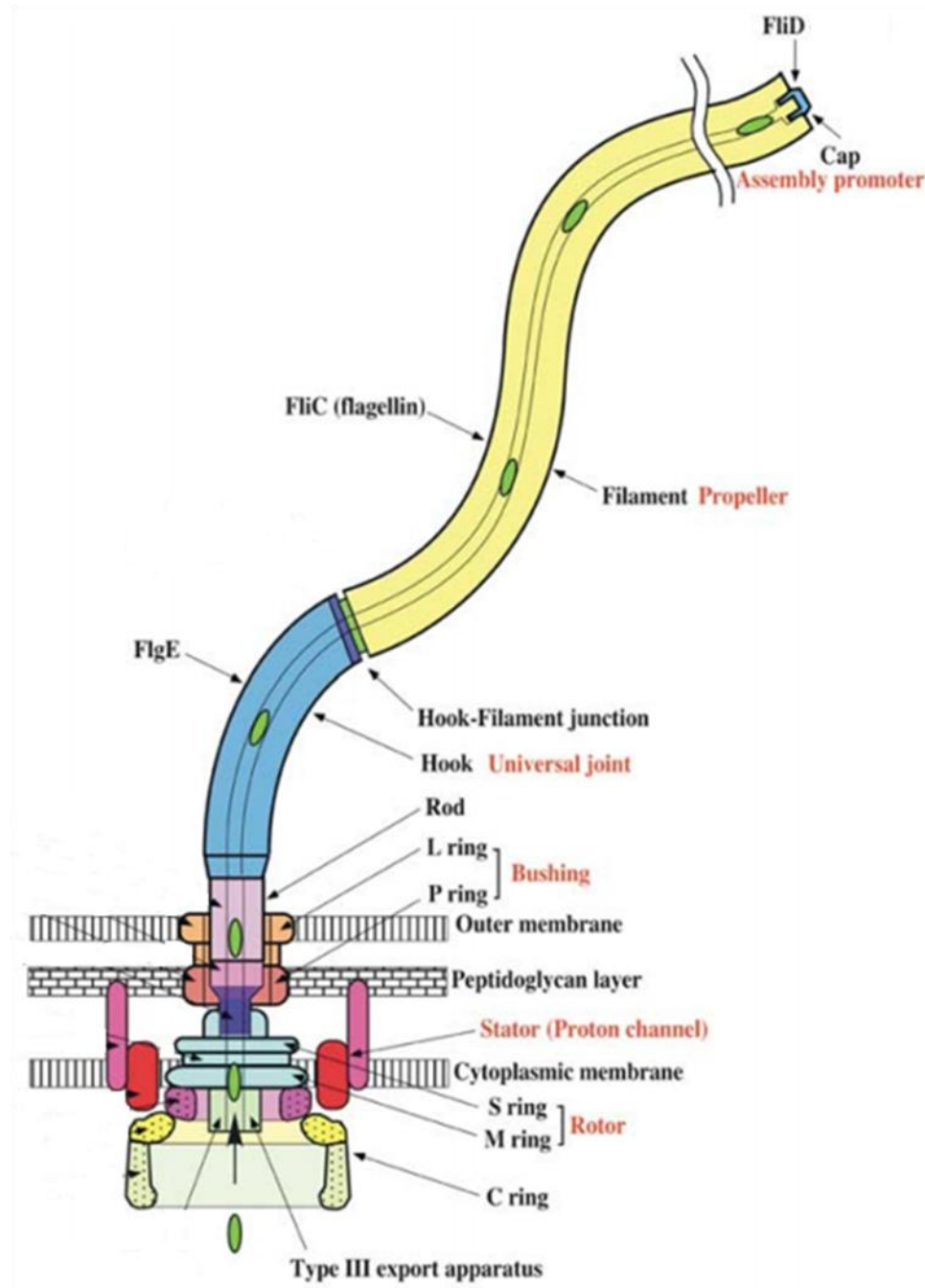


Figure 1.4

Schematic diagram of the bacterial flagellum apparatus for *Salmonella typhimurium*. The flagella apparatus consists of many parts as shown here. The main parts include the multiple membrane pores, the proton motor and the filament which is primarily composed of the flagellin protein, FliC. Figure adapted from Yonekura *et al.*³³.

counterclockwise motion, the flagella collapse to form a single bundle, and the bacteria is propelled forward, which is called running³⁴. When the motors rotate the flagella in a clockwise motion, the flagella unbundle and become disordered individual fibers, and the bacteria spin around at random, which is called tumbling^{35,36}. This random motion reorients the bacterium, which can then swim in a different direction, once the flagellar motor switches direction to counterclockwise rotation³⁷. At the molecular level, the abrupt change in rotation by the flagella motor shifts the interaction between the D0 and D1 domains of the FliC subunits²⁴, which stack up to form an 11 subunit protofilament⁶. When the bacteria are swimming or in motion, i.e., running, the flagella are in a left-handed helical shape. When the flagellar rotation is reversed, resulting in tumbling, the flagellin subunits switch their organization to a right-handed helical shape³⁴. This change in helical shape of the flagella can also be induced by changes in pH as well as ionic strength³⁸⁻⁴⁰.

1.4 Transition Metal Binding Sites in Proteins

Metalloproteins make up approximately one third of all structurally characterized proteins⁴¹. Metal ions are incorporated into proteins for a variety of reasons including stabilizing a protein's folded structure, making enzymes active or enhancing, diversifying or tuning the enzyme's function, depending on what metal it binds⁴². Metal binding sites in proteins can often accommodate many different types of metal ions. All four of the following transition metal ions, Zn^{2+} , Ni^{2+} , Co^{2+} and Cu^{2+} , can be observed in metal binding sites which are typically comprised of three or four

histidine (His) and/or cysteine (Cys) residues^{15,20,43,44} that are arranged in a slightly distorted tetrahedral fashion⁴⁵ (Figure 1.5). The zinc II ion (Zn^{2+}) has a filled d shell outer electron configuration $[\text{Ar}]3d^{10}$ and can accept four electron pairs to form four covalent, i.e. “metal coordinate” bonds with certain other elements⁴⁵. Two common zinc (II) ion ligands in proteins are either of the two nitrogen atoms in the imidazole ring of the His side chain or the thiol group sulfur atom in Cys side chains, with His being the most commonly observed⁴⁶. It has been observed that Co^{2+} can easily substitute for Zn^{2+} because they have identical charges and ionic radii⁴⁴. The substitution of Co^{2+} for Zn^{2+} causes few changes in the structure or activity of some enzymes, as observed with carbonic anhydrase⁴⁷. Furthermore, the Zn^{2+} ion, when permanently coordinated to only three other Cys / His ligands, can function as a Lewis acid, i.e., an electron pair acceptor, in some enzymes, such as carbonic anhydrase⁴⁸. Proteolysis, and the hydration of CO_2 , are two of the typical reactions for enzymes that use zinc ions for their catalytic activity⁴⁵.

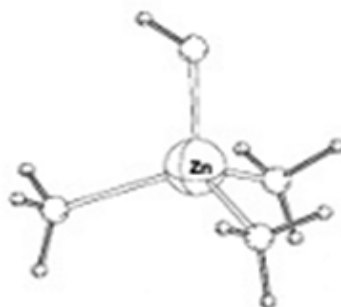


Figure 1.5

Tetrahedral binding geometry of zinc. For human carbonic anhydrase II (CA2), there would be three nitrogens from histidines and an oxygen from a water molecule in the solvent coordinating the Zn^{2+} ion. Figure is adapted from Diekmann *et al.*⁴⁹.

1.5 Flagellin Engineering

Due to the many favorable aspects of bacterial flagellin, it has been used as an engineering scaffold for diverse applications. Cheang *et al.*⁵⁰ added a biotin group to the flagellum fiber and used it as a bio-linker. This allowed the attachment of two magnetic beads that would spin under a rotating magnetic field. The flagellum fiber that connects the two beads added traction and induced the propulsion of the beads. Genetically engineered flagellins have also been used as a nanotube scaffold for peptide display and assembly of nanoparticles and biomineralized coatings because of its unique ability to self-assemble into hollow, long fibers with a high aspect ratio, i.e. the length-to-width ratio. For example, genetically modified flagellins have been used to form flagella nanotubes that were mineralized with TiO_2 ^{51,52}. By inserting large Cys loops, flagellin has been used to create palladium and gold nanotubes⁵³. Kumara also demonstrated this to some extent with various metal coatings and attachment of gold metal nanoparticles⁵⁴.

As noted above, the D3 domain in larger *E. coli* and *Salmonella* flagellins can be removed while still permitting flagella fiber formation⁵. Furthermore, the D3 domain can be replaced with other proteins and peptides of interest. Various heterologous peptides from other proteins such as egg white lysozyme were successfully inserted into the D3 domain of the flagellin protein²⁸. Adhesive peptide flagellin mutants can be isolated⁵⁵ and then mixed with other adhesive hybrid flagellins and allowed to polymerize, creating hybrid, multifunctional flagella⁵⁶. Some groups have used flagellin to produce vaccines⁵⁷⁻⁶¹ because it elicits an immune response⁶². The ability of multiple engineered flagellin proteins to polymerize together allows for a single, multiexpression system.

Internal fusions of complete proteins into the D3 domain of flagellin are also possible as demonstrated with thioredoxin⁶³ and more recently with green fluorescent protein (GFP)⁶⁴ and a xylanase enzyme⁶⁵. However, *in vivo* export and assembly of these internal fusion flagellin hybrids into functional flagella fibers is often problematic for larger hybrid flagellin proteins, requiring separate purification and *in vitro* polymerization steps for the GFP⁶⁴ and xylanase⁶⁵ flagellin proteins.

1.6 FliTrx Flagellin Peptide Display System

A specially designed flagellin peptide display system called FliTrx is comprised of a small, 109 residue *E. coli* thioredoxin protein that was genetically inserted into the middle of the solvent accessible, outer D2/D3 domain region of *E. coli* flagellin, between residues Gly 243 and Ala 352⁶³. Various types of random and rationally designed peptide loops can be inserted into the thioredoxin domain as loop fusions via cassette mutagenesis with a suitable restriction site in the corresponding plasmid DNA. The corresponding expressed proteins are exported and assembled *in vivo* into functional flagella fibers and can be screened for interaction with other proteins or materials using suitable strains of live *E. coli* cells, e.g. *E. coli* strain GI826⁶³. The FliTrx flagellin peptide display system has been utilized in a variety of peptide expression and panning experiments as well as various metal binding applications. A metal-binding FliTrx random mutant was selected via panning against an IL-8 monoclonal antibody that would precipitate out of solution upon binding of Zn²⁺ ions⁶⁶. Dong, *et al.*⁶⁷ screened a FliTrx random 12-mer peptide library for nickel binding functionality; consequently, one

FliTrx mutant was identified that would bind nickel at high enough affinity to sequester free nickel ions from a solution. Using a 10X His-tag loop, Woods, *et al.*⁶⁸ were able to successfully bind nickel and cobalt with a FliTrx protein. The FliTrx protein was also engineered to bind metals using a rationally-designed sequential histidine loop peptide on thioredoxin to form metallic and semi-conductor nanowires⁵⁴. Large bundles of disulfide-bond cross-linked flagella were also engineered by Kumara, *et al.*⁶⁹ using the FliTrx system, demonstrating the covalent-bond polymerization of flagella into larger fibers. These examples indicate that the FliTrx system has a number of possible applications as an engineered flagellin expression system that allows for further genetic and chemical modifications.

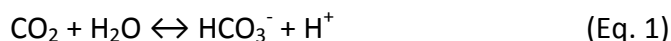
1.7 Flagellin Protein Engineering Difficulties

One difficulty in inserting large, stable proteins into the D3 domain of FliC flagellin is the potential for altering the ability of FliC to export and polymerize. The D0 domain in the monomeric form is disordered and therefore very susceptible to protease digestion^{70,71}, but forms an alpha-helical coiled-coil structure when polymerized into flagella fibers. Consequently, degradation of this region by bacterial proteases during purification of flagellin monomers will result in flagella that are no longer capable of polymerization *in vitro*²⁹. Therefore, extra precautions should be taken to prevent protease digestion if purifying the monomeric form of flagellin. The wild-type flagellin monomer is thought to be at least partially unfolded during export through the ~20 Å diameter central pore of the flagella fiber, prior to assembly at the distal tip of the

growing flagella fiber^{25,26}. Thus, any hybrid flagellins containing inserted proteins in the D2/D3 domain region will also need to be unfolded to some extent, and, insertion of a large and highly stable protein such as GFP can prevent the hybrid flagellin protein from exporting with high efficacy^{64,72}. In fact, GFP and xylanase internal flagellin fusions were not exported *in vivo* from live bacterial cells and only polymerized *in vitro* under high salt conditions^{64,65}. However, after careful protease-free *in vitro* purification, both fusions could be assembled to form either fluorescent or enzymatic bioengineered flagella.

1.8 Carbonic Anhydrase

Carbonic anhydrase (CA) enzymes are found in most eukaryotic and prokaryotic organisms⁷³. There are a total of 5 classes of CAs with the zinc-dependent α class being the most clinically relevant⁷⁴. In mammalian organisms, these CAs are responsible for pH control, bicarbonate metabolism and the regulation of intracellular osmotic pressure⁷⁵. *In vivo*, CAs catalyze the reversible hydration of carbon dioxide, CO₂, into a bicarbonate anion, HCO₃⁻, and a proton (Eq. 1).



For some CAs, e.g. mammalian CA2, this enzymatic reaction has one of the fastest rates known in nature, with k_{cat} (turnover number) values in excess of $1 \times 10^6 \text{ s}^{-1}$, which is near diffusion rate-limited kinetics⁷⁶. Some CAs, e.g., mammalian CA2, can also cleave some

ester bonds⁷⁷, allowing relatively simple monitoring of their catalytic activity with colorimetric ester bond substrates such as 4-nitrophenyl acetate (4-NPA or *p*NPA) (Figure 1.6)^{78,79}. The active site of CA2 contains a zinc center coordinated by three

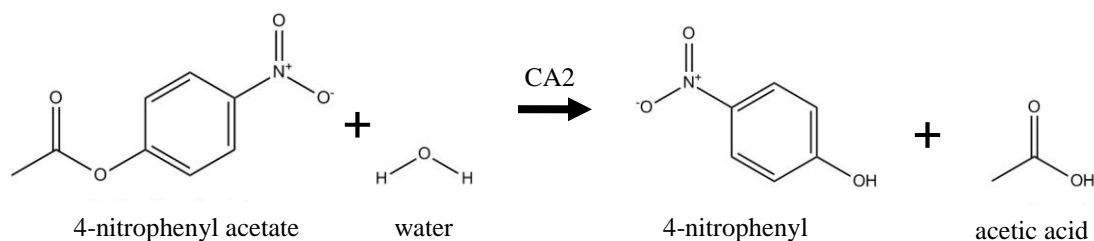


Figure 1.6

Representation of 4-nitrophenyl acetate hydrolysis. With 4-nitrophenyl acetate (pNPA or 4-NPA) as the starting product, CA2 can cleave the ester bond through hydrolysis to create 4-nitrophenol and acetic acid. The 4-nitrophenol product gives off a bright yellow color that can easily be detected through UV-Vis spectroscopy. Images generated using ChemBioDraw Ultra (Version 12).

histidines (His 94, His 96 and His 119) from the protein, and a single water molecule from the solvent, creating a tetrahedral shape (Figures 1.5 and 1.7)⁸⁰. There is another histidine (His 64) that acts as a proton shuttle upon the conversion of zinc-bound water to zinc-bound hydroxide⁸¹. When this His 64 residue is mutated to a residue that cannot shuttle protons (e.g., His 64 --> Ala), the enzymatic activity is reduced by up to 10-fold⁸². Other important active site residues include Thr 199, which accepts a hydrogen bond from zinc-bound hydroxide, and Trp 209, Val 121, Val 143 and Leu 198, which make the active site pocket more hydrophobic⁸³. This demonstrates a relatively simple design for metal binding but also the intricacy of enzymatic site residue functions.

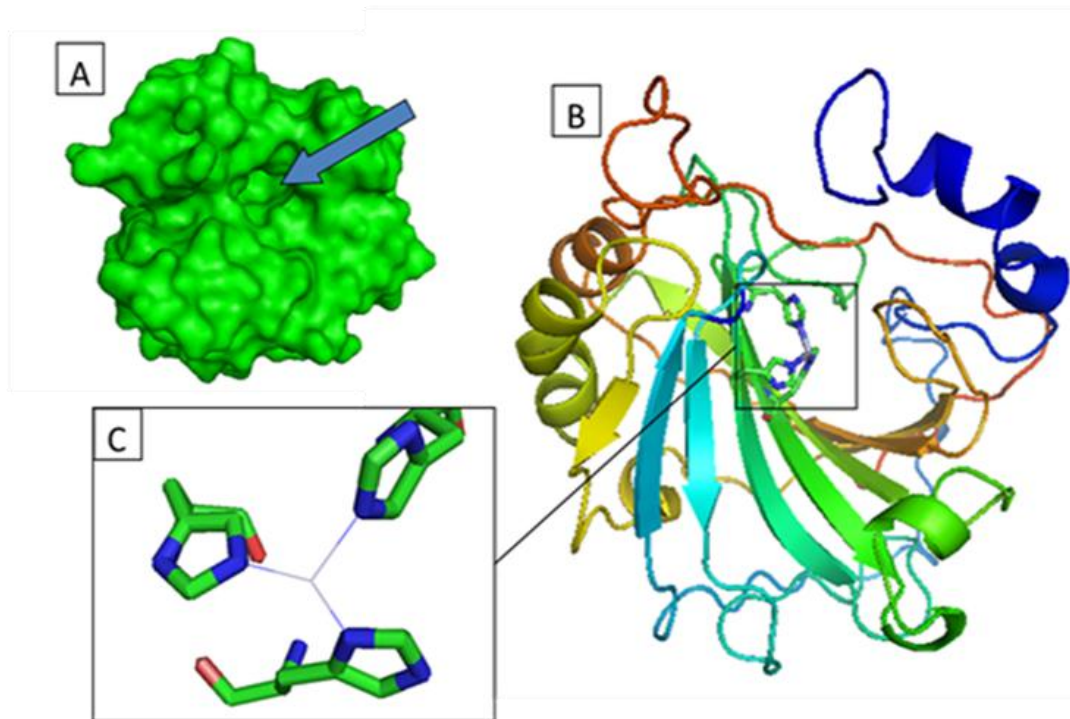


Figure 1.7

Computer modeling of human carbonic anhydrase II structure. Images were generated using PDB ID 1CA2 using PyMOL software. (A) Space-filling representation of the globular CA2 structure. The arrow points to the pocket in which the active site is located. (B) Ribbon diagram of CA2 with the three active site histidine residues that bind the zinc ion emphasized. (C) Close-up image of the three histidines that bind the zinc ion in CA2. The light blue lines show the metal coordinate covalent bonds between imidazole nitrogen atoms and the zinc ion, located at the center of the three lines (not shown).

1.9 Carbonic Anhydrase Engineering

Human CA2 has been the subject of enzyme manipulation through site-directed mutagenesis and chemical modifications to alter catalytic activity and substrate specificity. For example, some mutations have made the active site pocket larger or more hydrophobic to allow for bulkier or more hydrophobic substrates⁸⁴. The active site zinc ion can also be replaced with other transition metals. CA2 has been found to perform the reaction in Eq. 1 when Co^{2+} is substituted for Zn^{2+} ⁴⁷. However, if the zinc ion is replaced with manganese, the enzymatic activity is altered and instead shows

peroxidase activity⁸⁵. With just a few modifications, the CA2 catalytic activity can be altered to act on an even broader range of substrates or with cofactors not typical of the wild-type enzyme. Knowing what types of changes can help create new and modified enzyme function may give insight for additional mutation types to be introduced into flagellin, once metal binding and/or enzyme activity is established.

Several studies have utilized the three-histidine zinc-binding site of human CA2 as a model to design metal binding sites into other proteins. Several sites were identified on the light chain of the fluorescein-binding antibody and were mutated to allow for zinc (II) binding^{12,13}. Other examples include a metal binding site being introduced into charbdotoxin^{17,86}, mammalian serum retinal-binding protein¹⁹, and the serine protease, trypsin¹⁵. Using the CA2 three-histidine zinc-binding center as a model for engineering metal binding sites into other proteins has been proven successful in past experiments. In this study, we explored the possibility of engineering three novel zinc-binding sites into the *Salmonella* FlhC flagellin protein and present preliminary data for metal binding in the resulting flagellin variants.

CHAPTER 2

ENGINEERING A METAL BINDING SITE WITH THE POTENTIAL FOR CATALYTIC ACTIVITY INTO BACTERIAL FLAGELLIN

2.1 Introduction

2.1.1 Background and Rationale for Flagellin Engineering

Flagellin, encoded by the *fliC* gene (FliC protein), is one of the two types of protein monomer that makes up the flagella fiber in *Salmonella typhimurium* (the *Salmonella* genome also contains an alternate *fliB* flagellin gene), which is a model system for bacterial flagellum⁴ (Figure 1.2). This 494 amino acid protein contains a total of one His and zero Cys residues (Table 2.1). Beatson *et al.*²⁷ compared the amino acid sequences of the conserved D0 and D1 domains of 20 different diverse bacterial species to determine their sequence homology. The amino acid sequences of these 20 different flagellin proteins were further analyzed with respect to their His and Cys residue composition (Table 2.2). His and Cys are the two types of residues in proteins most likely to bind transition metal ions such as zinc (Zn^{2+}). This analysis indicated that most of the representative flagellins were similar to *Salmonella* FliC with regards to their composition of Cys and His residues. Specifically, all 20 of the flagellin sequences lacked any Cys residues and nearly all contained four or less His residues. *Vibrio parahaemolyticus* was the one exception to this trend; it contained ten His residues, which was the largest number of any of the compared flagellin sequences. Given that there is only one His residue and zero Cys residues in the entire 494 amino acid

Table 2.1
Sequence Composition and Physical Properties of Wild-type
***Salmonella typhimurium* Flagellin**

Amino acid	Number of Residues	Percentage composition
Ala (A)	61	12%
Arg (R)	14	2.8%
Asn (N)	42	8.5%
Asp (D)	37	7.5%
Cys (C)	0	0.0%
Gln (Q)	32	6.5%
Glu (E)	17	3.4%
Gly (G)	43	8.7%
His (H)	1	0.2%
Ile (I)	25	5.1%
Leu (L)	42	8.5%
Lys (K)	28	5.7%
Met (M)	2	0.40%
Phe (F)	6	1.2%
Pro (P)	5	1.0%
Ser (S)	38	7.7%
Thr (T)	57	11.5%
Trp (W)	0	0.0%
Tyr (Y)	12	2.4%
Val (V)	32	6.5%
Number of amino acids:	494	
Molecular weight:	51611.7	
Theoretical pI:	4.79	
Protein parameters were determined at ExPASy website using the Swiss-Prot accession number P06179: http://web.expasy.org/cgi-bin/protparam/protparam1?P06179@noft@		

sequence of the *Salmonella* FliC protein, it almost appears that transition metal binding was selected against in this flagellin. However, the overall charge of the protein is negative at neutral pH (pI = 4.8, Table 2.1), which could lead to some electrostatic

Table 2.2
Flagellin Amino Acid Sequence Comparison for Histidine and Cysteine in Various Species

Species and accession number	His	Cys
<i>Salmonella typhimurium</i> P06179	1	0
<i>Caulobacter crescentus</i> P18914	1	0
<i>Treponema pallidum</i> P21990	3	0
<i>Aquifex pyrophilus</i> P46210	2	0
<i>Helicobacter felis</i> Q9XB38	2	0
<i>Vibrio parahaemolyticus</i> Q9ZBA2	10	0
<i>Rhodobacter sphaeroides</i> O33578	3	0
<i>Azospirillum brasilense</i> Q43896	0	0
<i>Escherichia coli</i> Q5DY03	0	0
<i>Listeria monocytogenes</i> Q5Y833	1	0
<i>Yersinia pestis</i> Q8CZT1	4	0
<i>Aeromonas punctata</i> Q93TL9	2	0
<i>Rhizobium loti</i> Q98HD0	0	0
<i>Ralstonia solanacearum</i> Q9KGT9	0	0
<i>Bartonella clarridgeiae</i> Q9REF9	3	0
<i>Burkholderia cepacia</i> O68144	0	0
<i>Leifsonia xyli subsp. xyli</i> Q6AGB4	0	0
<i>Desulfovibrio vulgaris</i> Q729A8	4	0
<i>Bacillus anthracis</i> Q81SF2	0	0
<i>Bradyrhizobium japonicum</i> Q89NY8	0	0

Species and sequence names from Beatson *et al.*²⁷. Amino acid sequences were analyzed using ExPASy sequence analysis tools using the Swiss-Prot accession numbers listed after the specie's name (<http://web.expasy.org/protparam/>).

interactions with positively charged metal cations such as Zn^{2+} . Electrostatic interactions between the negatively charged exterior of the cell, which includes flagella, and positively charged metal surfaces have been previously observed⁸⁷. Flagella are known to interact with silver, because it is commonly used in flagella stains⁸⁸. Flagella can also be coated with TiO_2 ^{51,52}, although neither silver nor titanium are used as enzyme cofactors. Heavy metal ions in solution, including zinc and nickel, have been reported to increase the transcription of *fliC* in *E. coli*, although it is postulated to be an

escape mechanism from a toxic environment⁸⁹. The heavy metal up-regulation of *fliC* expression may explain why it could have been evolutionarily selected against for metal ion binding.

2.1.2 Design of Catalytic Flagellin Mutations

The *Salmonella typhimurium* wild-type flagellin protein (FliC) crystal structure was analyzed for solvent-accessible surface pockets using the CASTp program⁹⁰ (<http://sts.bioengr.uic.edu/castp/>). The program determines locations that a 1.4 Å solvent molecule, i.e. water, can fit on the surface of the molecule and determines the volume and surface area of the pockets. Predicted pocket sizes were compared to the human CA2 (PDB ID 1CA2) active site, which has a solvent-accessible surface area of 259 Å² and a volume of 282 Å³. Both the R-type straight supercoiled (PDB ID 1UCU) and the L-type straight supercoiled (PDB ID 3A5X) were analyzed for pockets similar to CA2 (Table 2.3) (Figure 2.1). There are some differences in the pocket sizes and number of pockets predicted between the two FliC conformations. However, the largest pocket still remains substantially larger than the CA2 active site pocket.

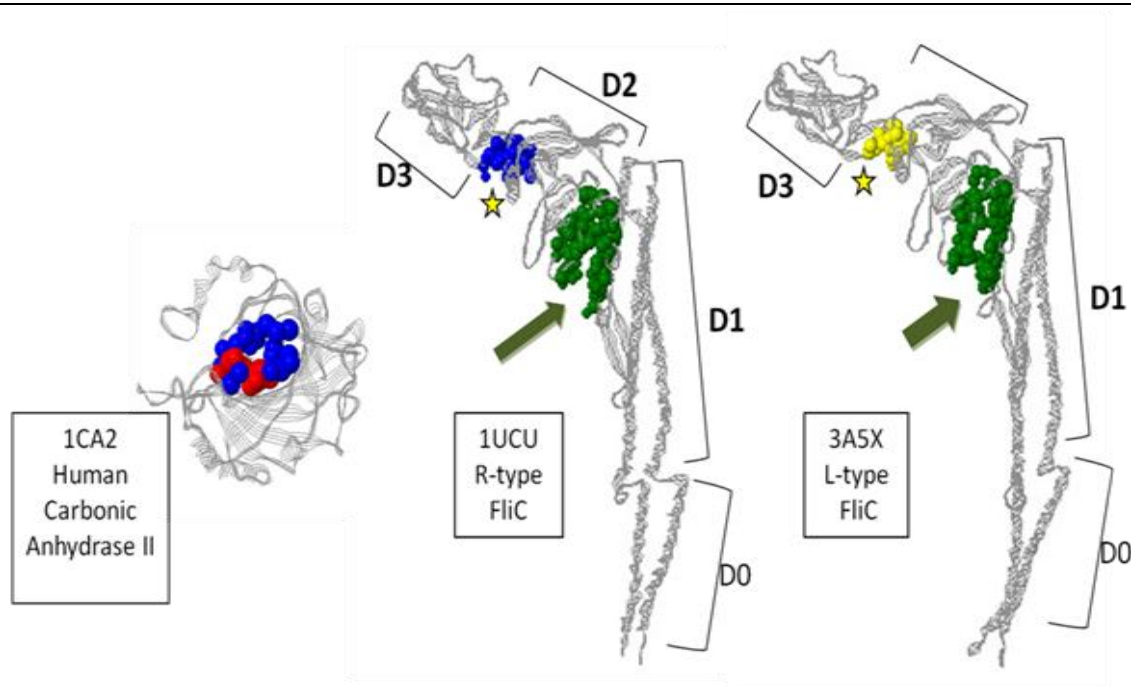
FliC flagellin has four globular protein domains (Figure 1.2), two of which, D1 and D2, are separated by a solvent accessible pocket. It has been observed that active sites in multidomain enzyme structures are typically located at interdomain regions⁹¹. These observations suggest that this pocket and other smaller pockets on the surface of FliC flagellin could potentially be engineered via molecular biology methods such as site-directed mutagenesis to function as enzyme-like catalytic sites and/or ligand/substrate

Table 2.3**CASTp Area and Volumes for Carbonic Anhydrase II, R-type and L-type Flagellins**

<u>1CA2 Carbonic Anhydrase II</u>			<u>1UCU R-type FliC</u>			<u>3A5X L-type FliC</u>		
<u>Name</u>	<u>Area</u> (\AA^2)	<u>Volume</u> (\AA^3)	<u>Name</u>	<u>Area</u> (\AA^2)	<u>Volume</u> (\AA^3)	<u>Name</u>	<u>Area</u> (\AA^2)	<u>Volume</u> (\AA^3)
32	202	346	48	946	1780	63	676	1230
31	259	282	47	292	344	62	154	129
30	125	152	46	162	229	61	129	145
29	93.6	85.0	45	151	163	60	148	173
28	84.5	69.1	44	117	98.1	59	166	179
27	108	78.8	43	133	168	58	139	185
26	42.7	33.9	42	135	144	57	163	152
25	50.8	31.0	41	144	125	56	121	212
24	45.6	30.8	40	137	176	55	195	138
23	81.3	59.5	39	131	119	54	134	128

Area and volume given for the top ten largest pockets predicted by CASTp.

Boxes around the different FliC pockets correspond to Figure 2.1.

**Figure 2.1**

Comparison of active-site-like solvent accessible pockets identified in carbonic anhydrase II, R-type straight flagellar filament and L-type straight supercoiled conformations of *Salmonella* FliC flagellin by the CASTp software. The active site pocket for CA2 and the largest pockets for FliC are denoted by the space-filling representations of residues lining each pocket. For FliC, the pocket located between domains D1 and D2 labeled with the arrow and the pocket labeled with a star were chosen for the locations to introduce the mutations. Pocket areas and volumes are given in Table 2.3.

binding sites. Therefore, the goal of this project was to design and introduce a transition metal binding site into the wild-type *Salmonella* FliC flagellin that could potentially function as a catalytic center. This structural protein has no previously characterized transition metal binding or known catalytic activity. Using rational design with computer modeling programs such as CASTp⁹⁰ and PyMOL Molecular Graphics Systems (Version 1.5.0.4) (Schrödinger, LLC), three sites were proposed by modeling them on the human carbonic anhydrase II (CA2) active site. CA2 has a zinc ion coordinated by three His residues and a water molecule in a tetrahedral fashion⁸⁰ (Figure 1.5). The residues to be mutated in FliC were modeled using PyMOL software by substituting other existing residues with His residues and manipulating the side chain bond angles of the imidazole side chain rotamers until they resembled the active site His residues of CA2. The sites chosen in flagellin for the mutations could allow the His imidazole side chains to be in close enough proximity to create the same binding geometry. Two of the three proposed metal-binding sites are located in a solvent-accessible pocket located between the flagellin D1 and D2 domains. The first of these two sites, termed KNT (Figure 2.2B), is comprised of the wild-type flagellin residues, Lys 348, Asn 120, and Thr 116. The second proposed site, termed QL (Figure 2.2C), is comprised of wild-type residues Gln 393, Leu 396, and His 388, and includes the only native His residue present in the wild-type flagellin; this His residue resides in the D2 domain. The third proposed metal-binding site, termed FQY (Figure 2.2D), is comprised of residues Phe 222, Gln 282, and Tyr 190. Introducing three His residues into these sites via replacement by site-directed mutagenesis could potentially create novel binding sites in the flagellin protein for

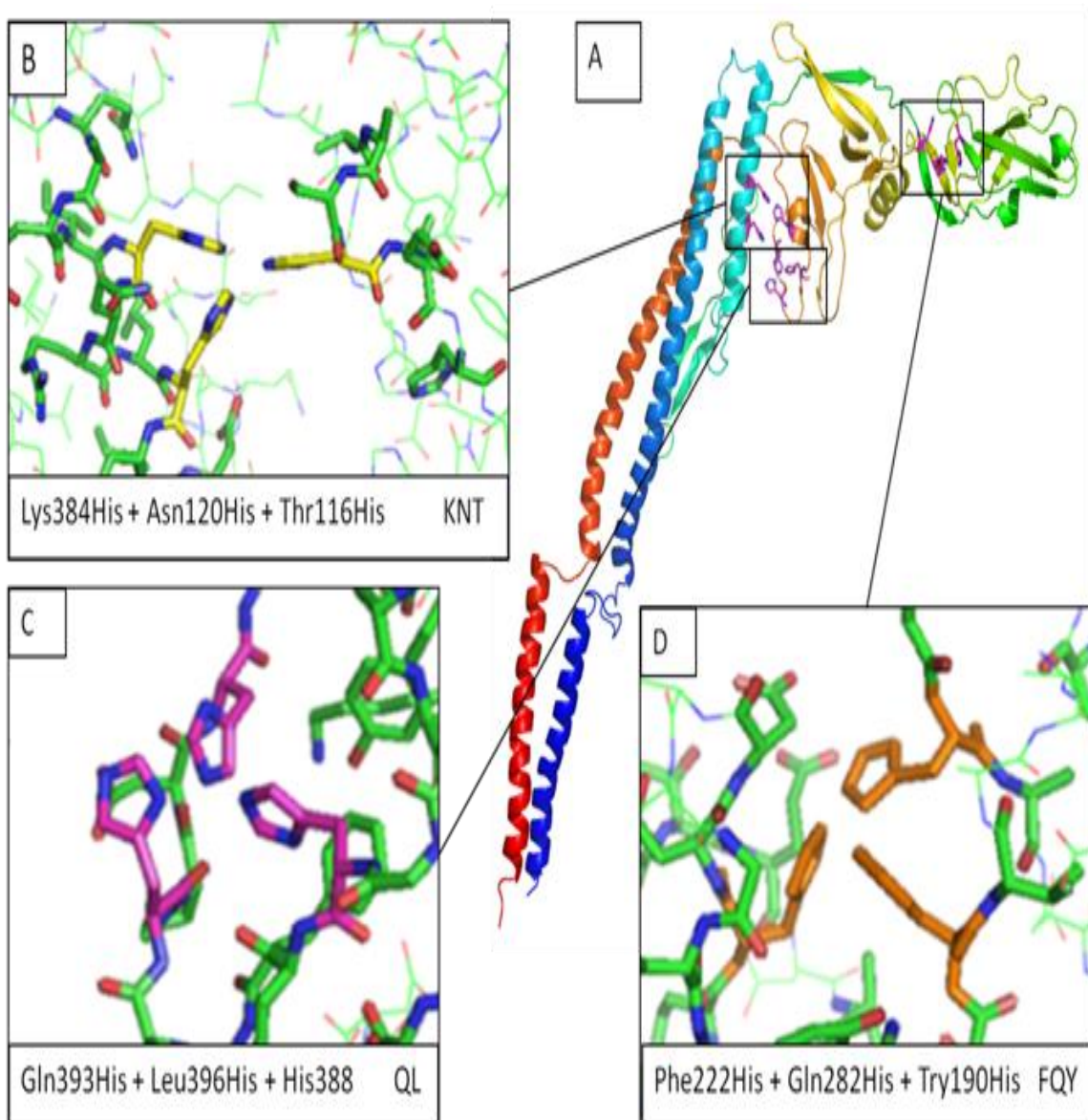


Figure 2.2

Computer modeled images of *Salmonella typhimurium* flagellin protein structure and the locations of rationally designed metal binding mutations. Images were prepared using PyMOL software. (A) Ribbon diagram of wild-type flagellin protein (PDB file 1UCU) showing the three solvent-accessible interdomain locations chosen for the histidine mutations. (B)(C)(D) Close up illustrations of the three rationally designed histidine residue clusters, KNT, QL and FQY, generated by site-directed mutagenesis, with lines connecting each zoomed in region to their location in the flagellin protein structure.

transition metal cations such as zinc II (Zn^{2+}), cobalt II (Co^{2+}), copper II (Cu^{2+}), and nickel II (Ni^{2+}). Furthermore, any successfully engineered metal-binding sites could potentially have catalytic activities such as hydrolysis and esterase activity, similar to that of CA2. By rationally designing an enzyme active site into flagellin, a greater understanding of metal binding proteins and enzymes will be gained. If introducing a new function into flagellin is successful, either metal binding and possibly enzymatic activity, it can then be further modified and optimized by rational design and/or random selection mutagenesis techniques to perform desired reactions while still allowing the flagella fiber to self-assemble under mild conditions.

2.2 Methods and Procedures

2.2.1 Media Preparation

Luria-Bertani (LB) Broth was prepared by dissolving 10 g tryptone (Fisher Scientific, Pittsburg, PA), 5 g yeast extract (EM Science, Gibbstown, NJ), and 10 g NaCl (Sigma-Aldrich, St. Louis, MO) in 900 ml deionized water. The pH was then adjusted to 7.0 and the volume brought up to 1 L. The solution was then autoclaved for 20 minutes on a liquid cycle. Once the media cooled, it was stored at 4 °C.

Luria-Bertani (LB) Agar was prepared by dissolving 10 g tryptone, 5 g yeast extract, 10 g NaCl, and 15 g agar (CalBioChem, San Diego, CA) in 900 ml deionized water. The pH was then adjusted to 7.0 and the volume was brought up to 1 L. The solution was then autoclaved for 20 minutes on a liquid cycle. Once it cooled to approximately

55 °C, the antibiotic was added at the desired concentration, mixed well and then the plates were poured. Plates were allowed to cool then stored at 4 °C in the dark.

Super Optimal broth with Catabolite repression (SOC) was prepared by dissolving 20 g of tryptone, 5 g of yeast extract, 0.5 g of NaCl, and 10 ml of 250 mM KCl (EM Science, Gibbstown, NJ) in 950 ml deionized water. The pH was then adjusted to 7.0 with NaOH and the volume was adjusted to 1 L. The solution was then autoclaved for 20 minutes on a liquid cycle. When the solution had cooled to approximately 55 °C, 10 ml of 1 M MgCl₂ and 7.2 ml of 50% glucose were aseptically added. SOC media was then stored at 4 °C.

2.2.2 Plasmid

The wild-type *Salmonella typhimurium* phase 1 flagellin *fliC* gene is encoded on the pTH890 expression plasmid, which was derived from the pTrc99A plasmid, with the *fliC* gene cloned into the XbaI/HindIII restriction sites. The pTrc plasmids are under control of a trp/lac hybrid, IPTG or lactose inducible promoter that is not fully repressed in the uninduced state. Thus, the *fliC* gene is constitutively transcribed at a moderate level, even when no induction agent is added, e.g. IPTG or lactose. The plasmid has a β -lactamase ampicillin (Amp) resistance gene as a selective marker (Figure 2.3). The pTH890 plasmid was a kind gift from the late Dr. Robert Macnab (Yale University).

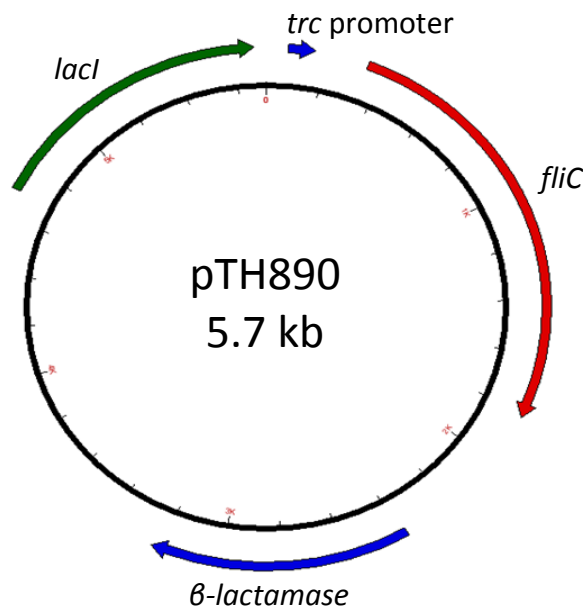


Figure 2.3

Computer generated image of the pTH890 plasmid. Genes are labeled accordingly. Image generated using Discovery Studio Gene (Version 1.5).

2.2.3 Site-Directed Mutagenesis of the *Salmonella* Flagellin Gene

Site-directed mutagenesis on the pTH890 plasmid containing the *Salmonella typhimurium fliC* gene was performed using synthetic DNA oligonucleotide primers encoding the desired mutations obtained from Integrated DNA Technologies (Coralville, IA) listed in Table 2.4. A PCR-type reaction was performed using Phusion™ High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA), with a final reaction volume of 25 μ l in a 200 μ l PCR tubes (Midwest Scientific, Valley Park, MO). Primers were designed using Agilent Technology's Quikchange Primer Design (<https://www.genomics.agilent.com/collectionsubpage.aspx?pagetype=tool&subpagetype=toolqcpd&pageid=15>). See Table 2.5 for reaction volume details and Table 2.6 for

Table 2.4
***Salmonella* FliC Flagellin Site-Directed Mutagenesis Primers**

Mutation Name	Primer Sequence	Nearest N. Temp.
Q393H	5' AGGTCACAACCTTTAAAGCA <u>CAC</u> CCTGAT <u>CAT</u> GCGGAAGCGGC 3' 5' GCCGCTTCCGC <u>ATG</u> ATCAGG <u>GTG</u> TGCTTTAAAGTTGTGACCT 3'	72
L396H	5' CTTTAAAGCACAGCCTGAT <u>CAT</u> GCGGAAGCGGCTGCTACAAC 3' 5' GTTGTAGCAGCCGCTTCCGC <u>ATG</u> ATCAGGCTGTGCTTTAAAG 3'	73
Y190H	5' GCTGCAACTGTTACAGGA <u>CAT</u> GCCGATACTACGATTG 3' 5' CAATCGTAGTATCGGC <u>ATG</u> TCTGTAAACAGTTGCAGC 3'	68
F222H	5' GATGGCGATTAAAA <u>CAT</u> GATGATACGACTGG 3' 5' CCAGTCGTATCATC <u>ATG</u> TTTTAAATCGCCATC 3'	61
Q282H	5' GAGGATGTGAAAAATGT <u>CAC</u> GTTGCAAATGCTG 3' 5' CAGCATTTGCAAC <u>GTG</u> TACATTTTTCACATCCTC 3'	63
N120H/ T116H	5' CCATCCAGGCTGAAATC <u>CAC</u> CAGCGCCTG <u>CAT</u> GAAATCGACCGTGTATC 3' 5' GATACACGGTCGATTC <u>ATG</u> CAGGCGCTG <u>GTG</u> GATTCAGCCTGGATGG3'	75
K384H	5' GGTAACCTTACGCTGCAAGT <u>CAC</u> GCCGAAGGTCACAAC 3' 5' GTTGTGACCTTCGGC <u>GTG</u> ACTTGCAGCGTAAGTTTTACC 3'	69

^aPrimers used for site-directed mutagenesis are listed with the name of the mutation and the nearest neighbor temperatures used to calculate the annealing temperature.

^bThe bolded and underlined base pairs indicate which codon was being mutated. The bolded text pairs are codons that have already been changed and have to be accounted for in the next primer. Primers are grouped by which site they are located.

Table 2.5
PCR Reaction Volumes and Concentrations with Phusion™ Polymerase for Site-Directed Mutagenesis

Item	Volume (μL)	Concentration
Phusion™ 5X HF Buffer	5	5X
DNA template	1	5-100 ng
Forward Primer	0.5	125 ng
Reverse Primer	0.5	125 ng
dNTP	0.5	10 mM
Sterile water	17.5	
NEB Phusion™ Polymerase	0.25	2 U/μL
Total	25.25	

Table 2.6
PCR Thermocycler Settings for Site-Directed
Mutagenesis on pTH890 Using Phusion™
Polymerase

Cycle step	Temp.	Time	Cycles
Initial denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	25
Annealing	NN-3	30 s	
Extension	°C	2m	
	72 °C	45s	
Final extension	72 °C	5 min	1
	4 °C	Hold	

thermocycler settings, which were based on a suggested New England Biolabs PCR protocol (M0530) found at <http://www.neb.com/nebecomm/products/protocol631.asp>.

Annealing temperatures were calculated via a “nearest neighbor minus three” method, where the nearest neighbor temperatures were derived using the calculator at

<http://www.basic.northwestern.edu/biotools/oligocalc.html>. A Techne Touchgene

Gradient Thermocycler (Bibby Scientific US, Burlington, NJ) was used for the PCR

reaction. After completion, the PCR DNA-template DNA mixture was incubated with

0.75 µL DpnI restriction enzyme (New England Biolabs, Ipswich, MA) at 37 °C for one

hour to digest the methylated plasmid template DNA. Using 1 µl of the PCR product and

25 µl of chemically competent XL-1 Blue Supercompetent *E. coli* cells (Stratagene, Cedar

Creek, TX), the plasmid DNA was transformed by incubating the plasmid DNA and cells

on ice for 30 minutes, heating at 42 °C for 45 seconds using a heat block, then

incubating on ice for two minutes. After the two minutes, 200 µl of room temperature

SOC was added to the transformed cells and the cells were incubated with shaking at 30 °C for 1 hour. Transformed cells were then plated on Luria-Bertani (LB) plates with 100 µg/ml Amp and incubated overnight at 37 °C. Isolated colonies were then picked with sterile toothpicks using sterile technique and used to inoculate 7 ml of LB-Amp broth and grown at 30 °C overnight in 34 ml PYREX® glass culture tubes (Corning, Tewksbury, MA) in a Lab-Line CEL-GRO® tissue culture rotary incubator (ThermoFisher Scientific, Waltham, MA) rotating at a speed setting of “7” (approximately 100 rpm). Plasmid DNA containing the desired mutations was purified from the overnight cultures using a QIAprep® Spin Miniprep kit (Qiagen, Valencia, CA). Samples of the purified plasmid DNA were analyzed by gel electrophoresis with 1% agarose gel with 0.5 µg/ml of ethidium bromide in a 1X Tris-acetate-EDTA (TAE) electrophoresis buffer (242 g Tris base (EM Science, Gibbstown, NJ), 57.1 ml of glacial acetic acid, 100 ml of 0.5 M EDTA (EM Science, Gibbstown, NJ), pH 8.0, made up to a final volume of 1 L) at 80 volts for 30 minutes. Gels were stained with ethidium bromide and were visualized using a White/2UV transilluminator (Ultraviolet Products, Upland, CA) at UV₃₀₂ with a Kodak imaging station with Kodak 1D™ LE 3.6 Software to confirm DNA plasmid size and concentration. Plasmid DNA for each mutation was then submitted for DNA sequencing at the University of Michigan Sequencing Core. Results were aligned using ClustalW2 at EBI through ExPASy tools (<http://www.expasy.ch/tools/>) and are listed in the Appendix.

2.2.4 *Salmonella* FliC Flagellin Expression and Purification

Purified pTH890 plasmid DNA isolated from *E. coli* XL-1 Blue, encoding the desired mutations as confirmed by DNA sequencing, was then transformed into chemically competent *Salmonella typhimurium* strain JR501. The two *S. typhimurium* strains used for wild-type motility and FliC variant expression, SJW1103 and SJW134, are restriction-proficient modification-proficient (R^+M^+). Consequently, transformation of either of these strains with a plasmid originating from other strains or species such as *E. coli* undergo severe restriction. *Salmonella typhimurium* JR501 is restriction-deficient modification-proficient (R^-M^+) and can accept plasmid DNA from other species, which can then be easily transformed into the flagellin-deficient *S. typhimurium* strain, SJW134⁹². Transformed cells were plated on LB-Amp plates and incubated overnight at 37 °C. An isolated colony was picked using aseptic technique and used to inoculate 7 ml of LB-Amp broth in a 34 ml Pyrex® glass culture tube that was incubated at 37 °C in a Cel-Gro® tissue culture rotary incubator rotating at speed setting “7” (approx. 100 rpm). The methylated plasmid DNA was then purified from the culture with a Qiagen QIAprep® Spin Miniprep Kit (Valencia, CA). The purified, methylated plasmid DNA was then transformed into the *Salmonella typhimurium* SJW134 expression strain via chemical heat shock. This SJW134 strain has both phase 1 (*fliC*) and phase 2 (*fljB*) flagellin genes deleted, making it non-motile, unless a functional flagellin gene is introduced via transformation with a suitable expression plasmid, e.g. pTH890. Transformed SJW134 strains were grown on LB-Amp plates overnight at 37 °C. Isolated colonies were used to inoculate 7 ml of LB-Amp broth using aseptic technique. These

cultures were then mixed with sterile glycerol to a final concentration of 25% glycerol (EM Science, Gibbstown, NJ) and frozen in liquid nitrogen and stored at -80 °C for future use.

Crude purifications of extracellular flagella were performed to demonstrate flagella export and assembly by centrifuging 1 ml of a 7 ml overnight culture at 13,000 rpm in a 1.7 ml graduated microtube (BioExpress, Kaysville, UT), using an Eppendorf Centrifuge 5430 (Hauppauge, NY) for 30 minutes at 4 °C to pellet the cells. The pellet was resuspended in 100 µL of 25 mM Tris, 150 mM NaCl, pH 7.3 buffer, vortexed on high for 30 seconds to mechanically shear off the extracellular flagella fibers (Figure 2.4), and then pelleted via centrifugation at 13,000 rpm using the same centrifuge as above at 4 °C for 20 minutes to pellet the cells. The resulting supernatant contained a suspension of sheared flagella fibers. The supernatant was then mixed with an equal volume of 2x sodium dodecyl sulfate (SDS) loading dye (1.52 g Tris, 20 ml glycerol, 2 g SDS, 2 ml 2-mercaptoethanol, 1 mg bromophenyl blue, adjusted to a final volume of 100 ml with DI water and pH adjusted to 6.8 with 1N HCl) and heated at 90 °C for 5 minutes in a heat block. Samples were then analyzed by electrophoresis on a 15 % acrylamide sodium dodecyl sulfate- polyacrylamide gel electrophoresis (SDS-PAGE) gel at 160 volts for 90 minutes and stained with Coomassie Brilliant Blue protein dye to estimate protein size, concentration and purity.

For full scale purifications, 1 L cultures of SJW134 strains of *Salmonella typhimurium* transformed with the pTH890 plasmid of interest were grown at 37 °C to an OD₆₀₀ of 0.8-1.2 and then chilled on ice for 20 minutes. Cultures were then

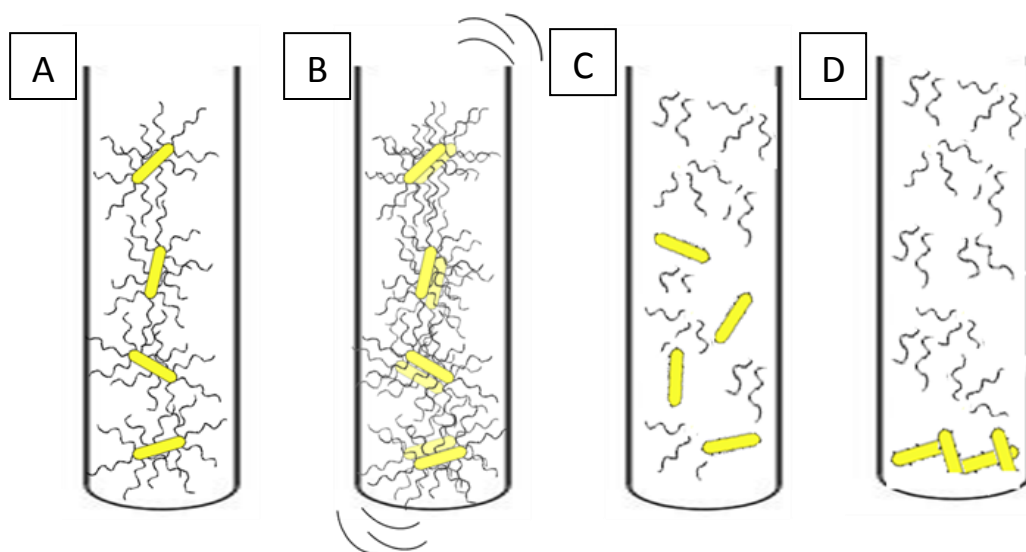


Figure 2.4

Illustration of mechanical shearing of flagella fibers. (A) Culture of flagellated bacteria. (B) Bacteria are then subjected to intense shaking either by vortexing or blending for two minutes. (C) Force of shaking beaks the flagella fibers off the bacterial cell. (D) Deflagellated cells are then pelleted using centrifugation leaving the flagella fibers in solution.

transferred to sterile 1 L Nalgene® centrifuge bottles (ThermoFisher Scientific, Waltham, MA) and spun at 7500 x g for 25 minutes at 4 °C using a Sorvall RC-3B refrigerated centrifuge with a Du Pont Sorvall H4000 swinging bucket rotor to pellet the cells. The supernatant was discarded and cells were resuspended in 30 ml of 150 mM NaCl, 25 mM Tris, pH 7.3 buffer by gentle swirling. Resuspended cell solutions were transferred to a Hamilton Beach 10 Blend Master blender and subjected to mechanical shearing on the high blend setting for 2 minutes to remove flagella from the bacterial cells. The solution was then transferred to sterile centrifuge tubes and spun at 5500 x g for 15 minutes at 4 °C using a Beckman Coulter Allegra X-22R benchtop centrifuge and a Beckman F0685 fixed angle rotor. The supernatant was transferred to sterile ultra centrifuge tubes in which the sample volumes were adjusted to yield the same weight

for each tube. Using a Beckman Optima XL-100K Ultracentrifuge, the protein solutions were spun at 105,000 x *g* for 1 hour at 4 °C using a Beckman SW 28 swinging bucket rotor. The pelleted flagellin protein was resuspended in 100 mM NaCl, 15 mM MOPS (CalBioChem, San Diego, CA), 50 mM EDTA pH 7.5. The pellet was then washed and spun down two more additional times. The resulting flagellin protein was mixed and heated with loading dye and then analyzed via SDS-PAGE with a 15% acrylamide gel and stained with Coomassie Brilliant Blue protein dye to determine size and purity of the protein. The protein was frozen in liquid nitrogen and stored at -80 °C. Using a Shimadzu UV-1650PC spectrophotometer (Shimadzu, Koyoto, Japan), UV₂₈₀ absorbance measurements were used to determine the concentration of the protein monomer using the extinction coefficients and molecular weights given in Table 2.7.

Table 2.7
Wild-type and Engineered Flagellin Variant Protein Extinction Coefficients and Theoretical Isoelectric Points (pI)

Flagellin Variant Name	Ext. Coefficient ^a (M ⁻¹ cm ⁻¹)	Isoelectric Point (pI) ^b
WT	17880	4.79
FQY	16390	4.94
KNT	17880	4.89
QL	17880	4.89

^{a,b} Predicted molar extinction coefficient values at a wavelength of 280 nm and pI values were calculated from the mature 494 residue amino acid sequences of wild-type and engineered flagellin variant using the ExPASy ProtParam algorithm (<http://web.expasy.org/protparam/>).

2.2.5 Carbonic Anhydrase II Expression and Purification

Recombinant human CA2 was expressed and purified using previously described methods⁹³. The pACAI plasmid was transformed into BL21 Star™ DE3 chemically competent cells (Invitrogen, Grand Island, NY). Cells were incubated in SOC media at 30 °C for 45 minutes and were then plated on an LB-Amp plate and incubated overnight at 30 °C. An isolated colony was picked using aseptic technique with a sterile toothpick and used to inoculate 25 ml of LB-Amp media shaken at 225 rpm at 37 °C overnight. Using 10 ml from the overnight culture, 1 L LB-Amp and 150 µM ZnSO₄ was inoculated and grown in a 2 L Pyrex® Erlenmeyer flask (Corning, Tewksbury, MA) at 225 rpm at 37 °C. Once the culture attained an OD₆₀₀ of 0.8-1.0, expression of the CA2 enzyme was induced by addition of 0.1 mM IPTG; 350 µM ZnSO₄ was also added to the media at this point, and the culture was grown in shaking at 225 rpm for 3 hours. After three hours, the culture was spun down at 7,500 x g for 20 minutes at 4 °C using a Sorvall RC-3B refrigerated centrifuge with a Du Pont Sorvall H4000 swinging bucket rotor to pellet the cells. The cells were then resuspended and lysed in 20 mM, pH 7.5 Tris buffer containing 1.0 mg/ml lysozyme, 5 µg/ml deoxyribonuclease I, and 5 mM MgSO₄ and magnetically stirred on ice for 30 minutes. The bacterial cell lysate was then centrifuged at 8,000 x g for 15 minutes at 4 °C using a Beckman Coulter Allegra X-22 centrifuge. The supernatant was then filtered with a 0.2 µm cellulose acetate membrane filter (VWR International, Radnor, PA) and loaded onto an Amersham Biosciences SP Sepharose Fast Flow cation exchange column using an Amersham Pharmacia Biotech ÄKTA FPLC system. The protein solution was captured on the SP column and washed with 15 mM MOPS, pH 7.5

buffer and the bound protein was then eluted by application of a linear gradient of 0-1 M NaCl, 15 mM MOPS pH 7.5. Protein-containing fractions were identified by UV₂₈₀ absorbance readings by the FPLC system. Fractions with the highest apparent concentrations of protein were frozen in liquid nitrogen and stored at -80 °C for future use.

2.2.6 *In vivo* Motility Assay

Salmonella typhimurium SJW134 strains containing the pTH890 plasmid of interest were grown overnight in 7 ml of LB-Amp broth in 34 ml Pyrex® glass culture tubes (Corning, Tewksbury, MA) in a CEL-GRO® tissue culture rotary incubator rotating at a speed setting “7” (approx. 100 rpm). These overnight cultures were used to inoculate 0.3% agar LB plates by injecting 2 µL into the agar, while attempting to avoid air bubbles and overflow on the surface. Inoculated plates were incubated at 30 °C for 6-8 hours in a humidified incubator with the agar side facing down. Plates were then imaged using a Kodak imaging station with Kodak 1D™ LE 3.6 software. Motility assays in the presence of metals ions and EDTA, were performed by injecting the desired test solution into the agar plate approximately 10 mm from the culture inoculation site.

2.2.7 Trypsin Digest of Flagellin Proteins

Overnight cultures of *Salmonella typhimurium* SJW134 previously transformed with mutated pTH890 plasmids of interest were centrifuged at 5500 x g for 20 minutes at 4 °C to pellet the bacterial cells using a Beckman Coulter Allegra X-22 centrifuge. The

cell pellet was resuspended in 500 μ L of 15 mM MOPS, 150 mM NaCl pH 7.5 buffer and vortexed on high for 90 seconds. The solution was then centrifuged to pellet the deflagellated cells at 5500 x g, 10 minutes at 4 °C. The supernatant typically contained a colloidal suspension of the sheared, polymerized flagella. Using a Shimadzu UV-1650PC spectrophotometer, UV₂₈₀ measurements were used to determine the flagellin protein concentration of each sample. For trypsin digests with monomeric flagellin, the solution was contained in a 1.7 ml graduated microtube (BioExpress, Kaysville, UT) and was heated at 65 °C for 10 minutes in a heat block. EDTA was included in the metal-free experimental samples at a final concentration of 1 mM. Trypsin was added to give a final molar trypsin to protein ratio of 1:300 and allowed to digest at room temperature. Samples of digested protein were taken at times 0, 1, 5, 10, 30, 60, 120, 180 minutes and mixed with an equal volume of SDS loading buffer and immediately heated for 5 minutes at 90 °C in a heat block. Samples were then analyzed via a 15% acrylamide SDS-PAGE gel run at 160 V for 90 minutes and stained with Coomassie Brilliant Blue dye.

2.2.8 Circular Dichroism Spectroscopy

Monomerized flagellin proteins were dialyzed into 10-20 mM sodium phosphate buffer pH 7.3 using 8,000 MWCO Spectra/Por Biotech membrane dialysis tubing (Spectrum Laboratories, Rancho Dominguez, CA). The buffer was changed a total of three times with a minimum of 8 hours dialysis between buffer changes, with gentle magnetic stirring at 4 °C. The FliC proteins were diluted to a concentration of 3 μ M in a 10 mm cuvette or 10 μ M in a 1 mm cuvette, prior to acquisition of CD spectra. Using a

Jasco J-815 Circular Dichroism Spectrophotometer (Easton, MD), each sample was scanned three times and averaged from 250 to 200 nm with a data pitch and data interval of 0.1 nm, a bandwidth of 1 nm at a scanning speed of 200 nm/min. The protein signal was then subtracted from the buffer signal using Spectra Manager™ Software.

2.2.9 Transmission Electron Microscope Imaging

All electron microscopy imaging was performed at the WMU Imaging Center. For each of the FliC mutants and WT, 10 μ L of an overnight culture was placed on a FCF400-Cu Formvar Carbon grid (Electron Microscopy Sciences, Hatfield, PA) and then allowed to adsorb on the grid for 10 seconds. The majority of the liquid culture was then removed from the grid from one edge via capillary action with Qualitative-Grade 615 Filter Paper (American Scientific Products, McGraw Park, IL) and allowed to air dry completely, which took approximately two minutes. Then 10 μ L of 2% (w/v) phosphotungstic acid, pH 5.2 staining solution was placed on the TEM grid and allowed to incubate with the previously deposited sample for 10 seconds. The solution was then removed via capillary action with Qualitative-Grade 615 Filter Paper to one side. The grid was then again allowed to air dry completely for approximately two minutes before inserting into a JEOL JEM-1230 Transmission Electron Microscope (Tokyo, Japan) operating at an accelerating voltage of 80 kV.

2.2.10 Inductively Coupled Plasma-Emission Spectrometry (ICP-ES) Analysis of Protein Metal Content

Samples of WT FliC, CA2 and the three mutant FliC proteins, QL, KNT and FQY, were prepared for ICP-ES analysis using the ultracentrifugation purification method described in Section 2.2.4, FliC Expression and Purification. Monomeric samples were heated at 65 °C for 10 minutes in 15 ml centrifuge tubes (VWR International, Radnor, PA) using a heat block. Protein samples were then dialyzed with 250 μ M ZnSO₄ in 25 mM MOPS pH 7.4 buffer, adjusted with KOH, for 12 hours at 4 °C. This was then dialyzed against three changes of 25 mM MOPS pH 7.4 buffer without zinc for a minimum of 6 hours, stirring gently at 4 °C each using 8,000 MWCO Spectra/Por biotech membrane dialysis tubing (Spectrum Laboratories). Proteins were at concentration of approximately 5 μ M, as determined by UV₂₈₀ absorbance measurements using a Shimadzu UV-1650PC Spectrophotometer. Protein solutions were then sterile filtered using a 0.2 μ m cellulose acetate membrane syringe filter (VWR International, Radnor, PA). Final UV₂₈₀ measurements were also collected after filtration to obtain the final concentration of each protein. Samples were then frozen in liquid nitrogen and shipped to the Chemical Analysis Laboratory at the Center for Applied Isotope Studies, University of Georgia (Athens, GA, <http://www.cais.uga.edu/index.htm>) where they performed a 20 metal ICP-ES analysis of the samples. To confirm either polymeric or monomeric state of the flagellin proteins, samples were analyzed for resistance to proteolysis by incubation with 0.5 μ l 1 mg/ml trypsin added to 50 μ l of 5 μ M protein for three minutes at room temperature and then mixed with 50 μ l of 2x SDS load dye and heated 90 °C for 5 minutes. Protein was then analyzed for presence of full-length versus digested protein

products using SDS-PAGE with a 15% acrylamide gel run at 160 V for 90 minutes and stained with Coomassie Brilliant Blue protein dye for visualization.

2.2.11 4-NPA Esterase Activity Assay

Using a clear Costar[®] 96 well flat-bottom, polystyrene microplate (Corning Costar, Corning, NY), *Salmonella* WT FliC flagellin and the three FliC variants were screened for hydrolysis activity using the commercially available ester compound 4-nitrophenyl acetate (4-NPA), obtained from Sigma-Aldrich (St. Louis, MO). FliC proteins used for this assay were prepared using methods described in Section 2.2.10, Inductively Coupled Plasma-Emission Spectrometry (ICP-ES) Analysis of Protein Metal Content. Each 96-well microplate was prepared with triplicate samples of each protein at approximately 5 μ M concentration in 25 mM MOPS pH 7.5 buffer, to give an initial sample volume of 90 μ L. Plates were mixed for 5 minutes at room temperature using a RotoMix Type 48200 microplate shaker (Barnstead/Thermolyne, Dubuque, IA) before the substrate was added. The 4-NPA substrate was dissolved in 10% vol/vol DMSO and then diluted to final volume with 25 mM MOPS, pH 7.5 buffer. A Multidrop-384 8-channel peristaltic pump dispenser (MTX Laboratory Systems, Vienna, VA) was used to rapidly dispense \sim 10 μ L of 5 mM 4-NPA substrate solution into each well of the microplate and Abs₃₄₈ absorbance readings were immediately measured by a Molecular Devices SpectraMax Plus 384 (Molecular Devices, Sunnyvale, CA) at a wavelength of 348 nm at 15 second intervals for 15 minutes, using the SoftMax Pro (version 4.7.1) software. The absorbance data was then converted to a Microsoft Excel file for analysis.

Using the start point and the end point of each data set, the amount of change in A_{348} observed for each of the three FliC metal-site variants (KNT, FQY and QL) were compared with CA2 esterase rates as the positive control and WT flagellin as the negative control. The average of the WTp (polymer) rates was used for the background correction for the other polymeric proteins and the average of the WTm (monomer) was used for the background correction of the monomeric proteins. The CA2 Abs₃₄₈ data was corrected by subtracting the average background change in the wells without protein. All changes were then divided by the corrected average of CA2 and multiplied by 100 to get a percent activity compared to CA2.

2.2.12 Sao Paulo Metallo β -Lactamase (SPM-1) Inhibitor Screening Assay

Previously, 19 potential inhibitors of a beta lactamase enzyme, type SPM-1, were identified by a colorimetric absorbance screen of the Genesis Plus 960 compound library, a collection of bioactive compounds (MicroSource Discovery Systems Inc., Gaylordsville, CT). These high throughput assays were preformed in 96-well microplate format, using CENTA, a chromogenic cephalosporin substrate (CalBioChem, San Diego, CA) as the substrate. All 19 compounds (MicroSouce Discovery Systems Inc., Gaylordsville, CT) were diluted to a concentration of 10 mM and frozen at -80 °C for future use. The inhibitors were screened with SPM-1 enzyme and CENTA substrate to re-verify activity. CENTA was diluted to 1 mM in 50 mM HEPES (Fisher Scientific, Pittsburg, PA) pH 7.0 buffer and the freezer stocks of each inhibitor were diluted to 2 mM using DMSO. Using a clear Costar® 96 well flat-bottom, polystyrene microplate

(Corning Costar, Corning, NY), 1 μ l of the 2 mM inhibitor was added to 89 μ l of 50 mM HEPES pH 7.0 buffer with 111 μ M ZnSO₄ and 60 nM SPM-1 enzyme. Negative control wells lacked the SPM-1 enzyme. Plates were incubated for 15 minutes using a RotoMix Type 48200 microplate shaker (Barnstead/Thermolyne, Dubuque, IA). Using a Multidrop-384 8-channel peristaltic pump dispenser (MTX Laboratory Systems, Vienna, VA), ~10 μ l of 1 mM CENTA substrate was added to the plate, followed by absorbance readings at 405 nm (Abs₄₀₅) using a Molecular Devices SpectraMax Plus 384 plate reader (Molecular Devices, Sunnyvale, CA). Absorbance readings for each microplate were collected at 15 second intervals for a total of 10 minutes using SoftMax Pro (Version 4.7.1) software.

The inhibitors were then screened for IC₅₀ against SPM-1 using CENTA as the substrate. CENTA was prepared using the same methods as above. Using the same type of 96 well plates, samples were prepared with and without Triton X-100 (LabChem Inc., Pittsburg, PA). Triton X was added at a final concentration of 0.01%. Higher concentrations of inhibitor were used at 300, 200, and 150 μ M. Dilution wells then started at 100 μ M and diluted 2-fold for 12 wells. Microplates were then incubated by shaking at room temperature for 15 minutes, followed by rapid addition of ~10 μ l 1mM CENTA substrate, and Abs₄₀₅ measurements at 15 second intervals for a total of 10 minutes, as described above.

2.3 Results and Discussion

2.3.1 Flagellin Variant Protein Expression and Purification

All of the rationally designed DNA mutations were successfully introduced into the wild-type *Salmonella typhimurium* *fliC* flagellin gene and can be found in the Appendix. Figure 2.5 shows a 1% agarose gel analysis of a typical sample of purified plasmid DNA. This analysis was performed on all purified plasmid DNA, prior to submission for DNA sequencing, to verify the presence and concentration of DNA. Three different *Salmonella* FliC flagellin variants with potential transition metal-binding sites were rationally designed by visual analysis of the protein structure to identify clusters of solvent-accessible residues that could be mutated to His residues. Two of these clusters were located in a D1-D2 interdomain cleft and one was located between the D2 and D3 domains on a surface region. Both of these regions were previously identified by the

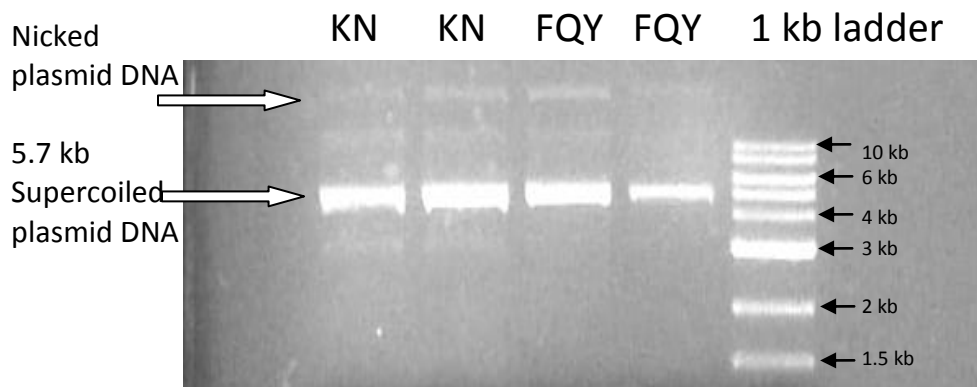


Figure 2.5

Example of agarose gel electrophoresis of purified pTH890 plasmid DNA encoding several mutations in the *fliC* gene. The 1% agarose gel was stained with ethidium bromide and imaged with a Kodak gel documentation system.

CASTp software as potential enzyme “active site-like” regions of the native flagellin protein. It has been observed that active site regions in multidomain enzymes are generally located at interfacial regions between adjacent domains⁹¹. These three flagellin variants were successfully generated from the WT FliC flagellin using site-directed mutagenesis of the *fliC* gene to replace two or three existing non-His residues located near one another in the folded protein structure with His residues, as shown in Figure 2.2. All three flagellin variants, QL, KNT and FQY, which were engineered for transition metal binding, retained the WT function in terms of high levels of *in vivo* expression, export and assembly into functional flagella fibers (Figure 2.6). These flagellin proteins were readily isolated and purified by mechanical shearing and centrifuge separation of the bacterial cells. The three FliC variant proteins were

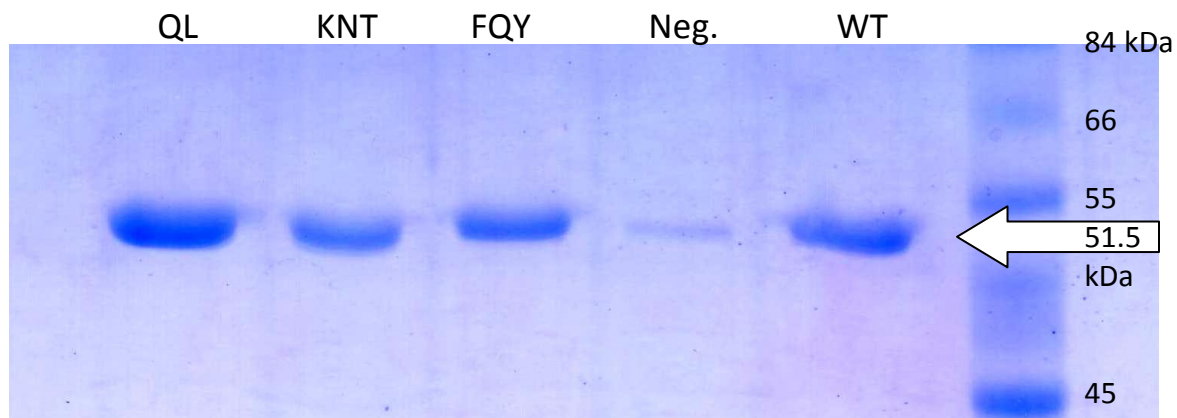


Figure 2.6

SDS-PAGE gel demonstrating export of flagellin variants. FliC proteins from SJW134 *Salmonella* cells with pTH890 plasmid were purified using mechanical shearing and separated on a 15% acrylamide SDS-PAGE gel to demonstrate export. The negative control (Neg.) is the extracellular protein isolated from the flagellin deficient SJW134 strain transformed with the pTrc99A vector. The small band visible in the negative control lane may be the flagellar hook associated protein, FlgE, which has a molecular weight of 50 kDa.

expressed and exported with approximately the relative same yields as WT and had essentially the same molecular mass as WT, which was indicated by SDS-PAGE analysis. The presence of protein bands under non-lytic conditions via mechanical shearing indicated that all three flagellin mutants were exported and assembled into flagella fibers.

2.3.2 *In vivo* Flagella Swarming Agar Motility Assay

Each of the three engineered QL, KNT and FQY flagellin variants were tested for physiological function with an *in vivo* swarming agar motility assay on 0.3% agar plates in the SJW 134 strain of *Salmonella* bacteria. This strain has deletions of both wild-type *fljC* and *fljB* flagellin genes, and will only exhibit motility when a functional flagellin gene is introduced via an expression plasmid, e.g. pTH890. While some motility function was observed for all three flagellin variants, variations were observed in their relative motilities, and a significant decrease in motility was observed for one of the variants. The KNT variant had approximately 10% of the wild-type motility, the FQY variant exhibited motility similar to wild-type, and the QL variant had approximately 70% of wild-type motility under normal assay conditions (Figure 2.7). Each of the flagellin mutants required at least two mutations before the final variant was generated. Thus, it is possible that the single or double mutations could lead to a change in motility behavior (Figure 2.8). Because the FQY triple mutation variant retained WT- like motility, it was not surprising that none of the mutations changed the motility pattern. The KNT variant had significantly reduced motility but the single histidine mutation did

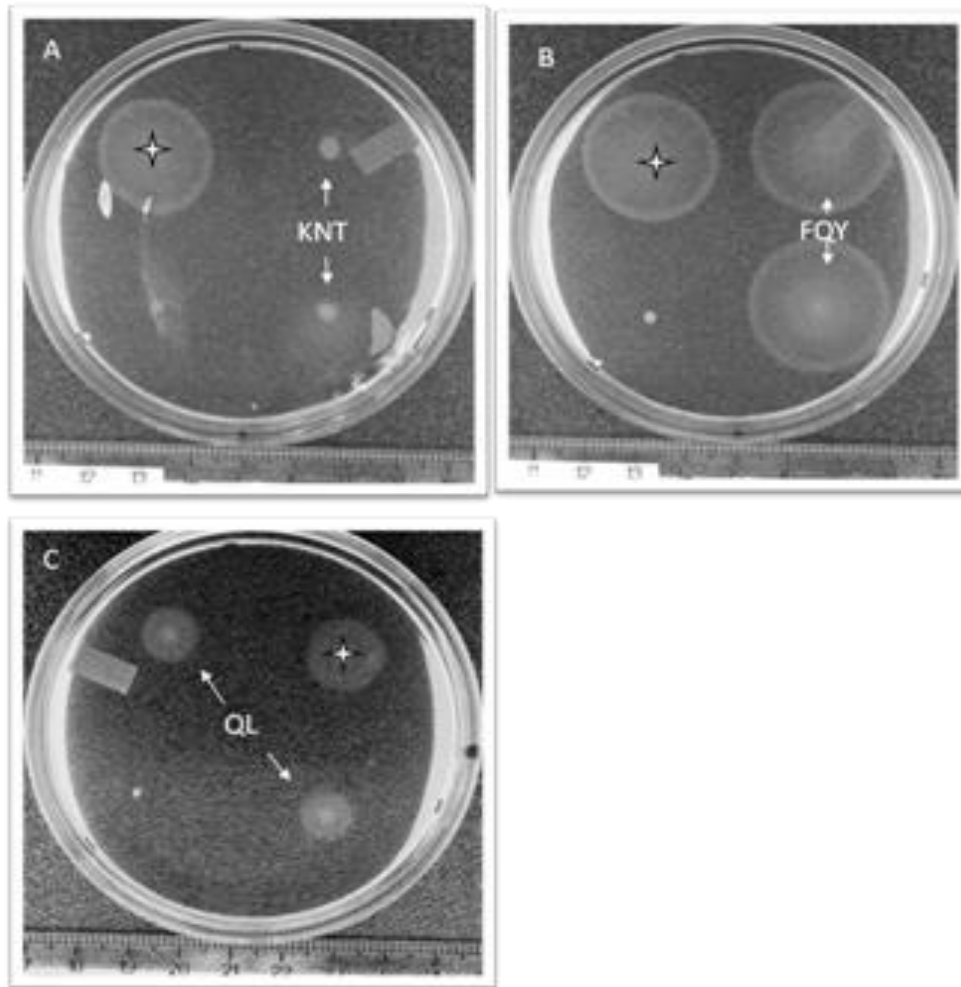


Figure 2.7

Motility agar assay of *S. typhimurium* flagellin variants. Using 0.3% agar LB, bacterial cultures of *S. typhimurium* strain SJW 134 were injected into the media at one point and incubated at 30 °C for six hours to observe any spreading of the cells due to flagellar motility, i.e., swimming. The four-pointed star symbol in each image represents the wild-type 100% motility positive control, which the other flagellin variants were compared to. The variants were grown in duplicate on each plate and are indicated with arrows. The final injection site in the lower left-hand corner is the non-motile strain SJW134 transformed with plasmid pTrc99A, which lacks the *fliC* gene.

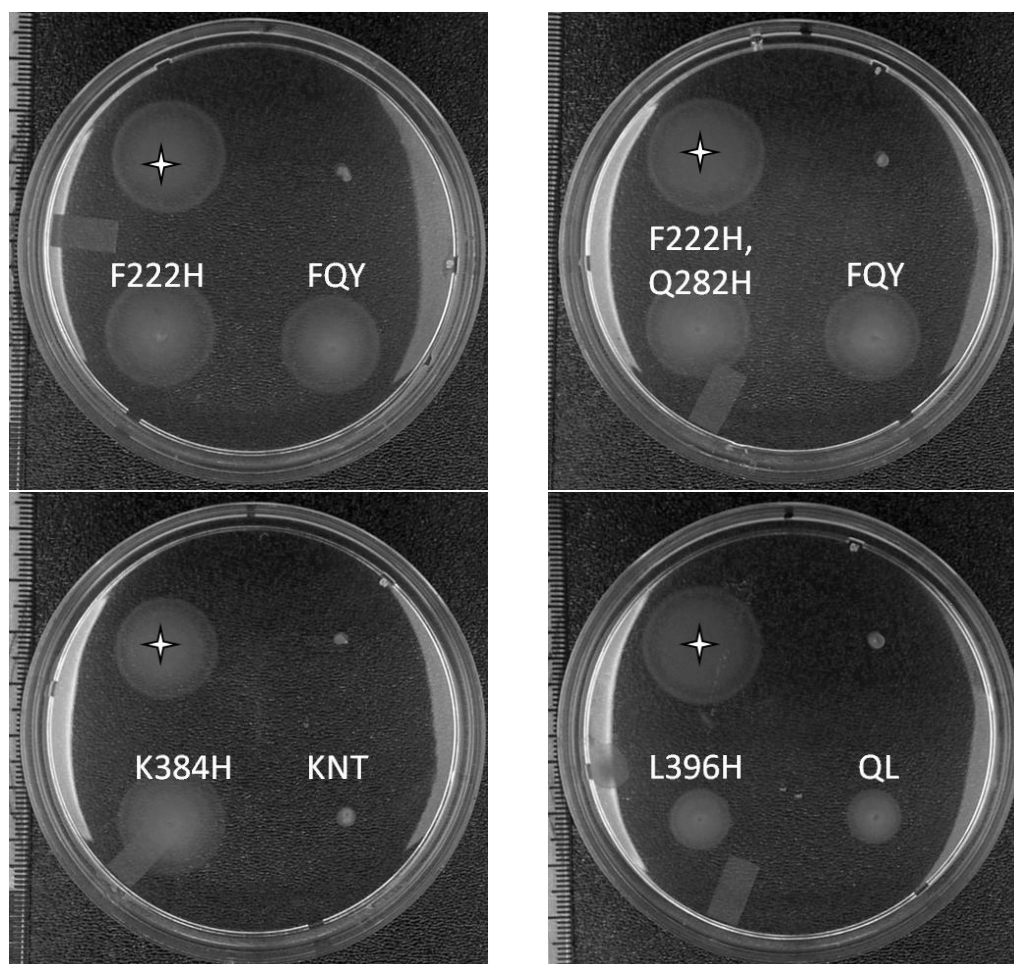


Figure 2.8

Motility of full and partial pTH890 variants in SJW134. Full and partial histidine substitution flagellin variants expressed in *S. typhimurium* were used to inoculate a point on 0.3 % LB agar to observe the relative swarming motilities. The four-pointed star indicates the WT positive control, to which all other samples were compared. Each of the swimming halos is labeled with the mutations present in the expressed flagellins. The upper right hand corner dot is the negative control SJW134 transformed with pTrc99A which lacks the *fliC* gene.

not appear to cause this change. The other two His residues were introduced simultaneously. Thus, the sequence changed from one His residue to three His residues in two rounds of mutations. After the other two mutations were introduced, the motility was greatly reduced. This result could indicate a change in the folded conformation of the protein or the stability of the flagella fiber. For the QL variant, the L396H mutation

was introduced first, and it appeared to reduce the motility of the bacteria drastically. It should be noted that at this site, there is already a single native His residue, His 388, and that the introduction of the second His residue resulted in reduced motility. The Q393H mutation, in addition to the L396H mutation, did not result in any additional decrease in motility.

A major goal of the His substitution mutations was to determine if any transition metal ions, e.g., Zn^{2+} , could be bound by the introduced His residue imidazole side chains. Because metal ions are known to stabilize the folded structure of many globular proteins, it is possible that a change in motility could be observed in either the presence or absence of metal ions. To investigate the possibility of metal-ion dependent motility, 0.3% motility agar plates were inoculated with *Salmonella* expressing WT or one of the three flagellin variants. This was followed by addition nearby of a solution of a transition metal ion or the metal ion chelator, EDTA, that should bind up any stray transition metal ions (Figure 2.9). The agar plates were then incubated for several hours and the bacterial inoculates were observed for changes in motility. When EDTA was added to the agar near the KNT variant inoculation site, a substantial increase in motility was observed. Motility was regained in the direction where the EDTA was inoculated, which indicates that the change in motility may be due to the presence of the EDTA metal chelator. Because this metal-dependent motility for this KNT flagellin variant could be due to a variety of factors, additional agar motility experiments were performed with the KNT flagellin variant. Purified DI water was inoculated at one site to see if a disturbance in the agar was responsible for the change. A 20 mM, pH 8.0 Tris

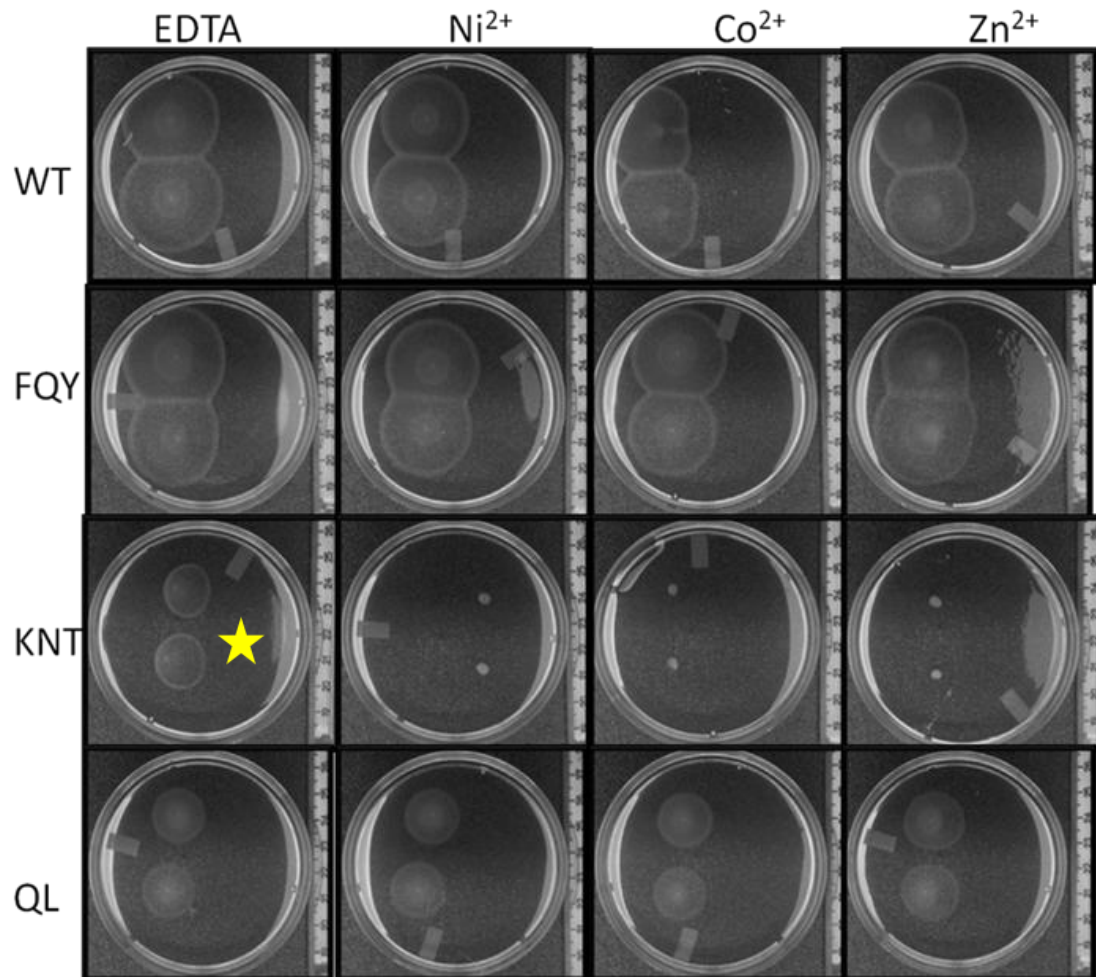


Figure 2.9

Motility of wild-type and flagellin variants in the presence of metals or EDTA. *S. typhimurium* SJW134 bacteria containing the pTH890 plasmid of interest were inoculated in 0.3% motility agar to observe the swimming pattern. The 0.1 mM metal or EDTA solution was injected into the agar approximately 10 mm from the bacterial inoculation site. The star indicates the motility plate assay that demonstrated a substantial amount of change in motility for the KNT variant. Plates were imaged six hours after inoculation and incubation at 30 °C.

buffer solution and a 20 mM, pH 5.9 Tris buffer solution were used at two different sites to determine if a change in the pH of the surrounding regions could also cause a change in motility. A third site was left with no additional inoculations as a control and a fourth region was inoculated with 0.1 M EDTA. Of all five possible substance inoculations, EDTA was the only one that induced the motility phenotype (Figure 2.10). These results suggest that by removing the metal ions from the solution surrounding the bacteria, the flagella are able to regain motility in media with 0.3% agar. If metal ions are causing a constraint on the conformation of the flagella fiber, this could lead to the motility phenotype observed. This could lock the bacteria in stationary tumble mode rather than

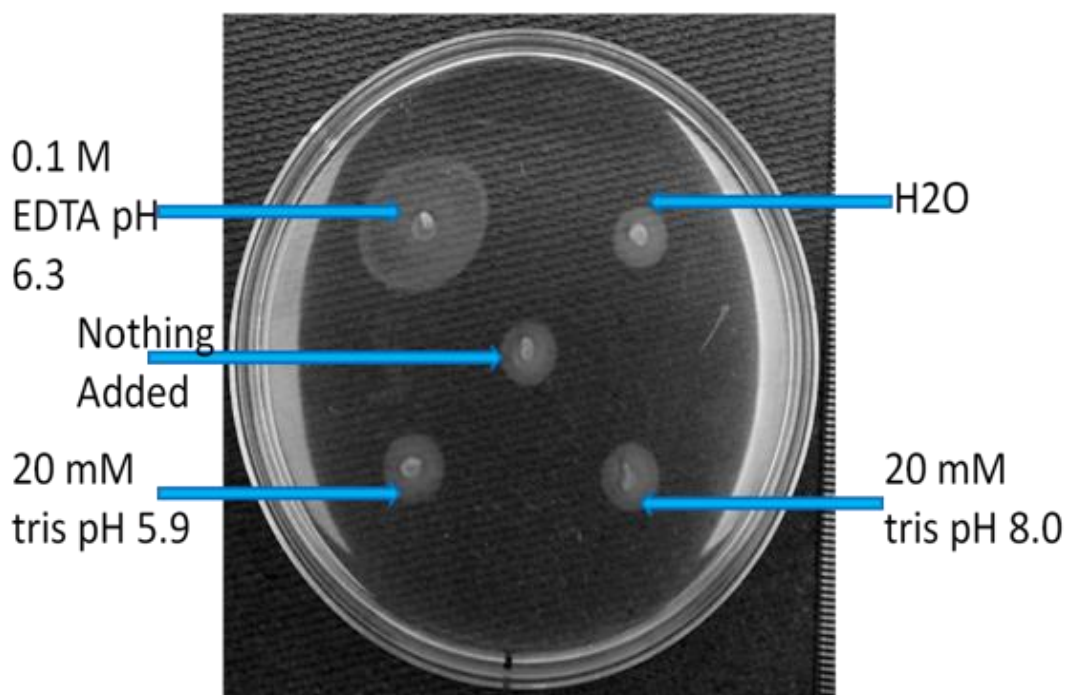


Figure 2.10

KNT flagellin variant motility as a function of pH, in the presence of EDTA and water. The KNT flagella variant exhibited a change in motility phenotype in the presence of EDTA, which was again demonstrated here. To rule out the possibility that the change in motility was due to changes in viscosity or pH, various solutions were inoculated next to the culture in motility media. The arrows indicate the approximate locations where test solutions were injected.

running, it could cause the flagella fiber to be in the straight supercoiled form which reduces motility⁹⁴, or it could cause the bacterium's flagella to stick together, as was observed previously when metal binding was taking place⁶⁸. The presence of the trace metal ions found in the media did not prevent the flagellin export and assembly, because the flagella can be mechanically sheared off as seen in Figure 2.11 and visualized with transmission electron microscopy (Section 2.3.5). Although the mechanism by which the His substitution mutations interfere with normal function is unclear, it was apparent from these results that the addition of the metal ion chelator EDTA reversed this effect.

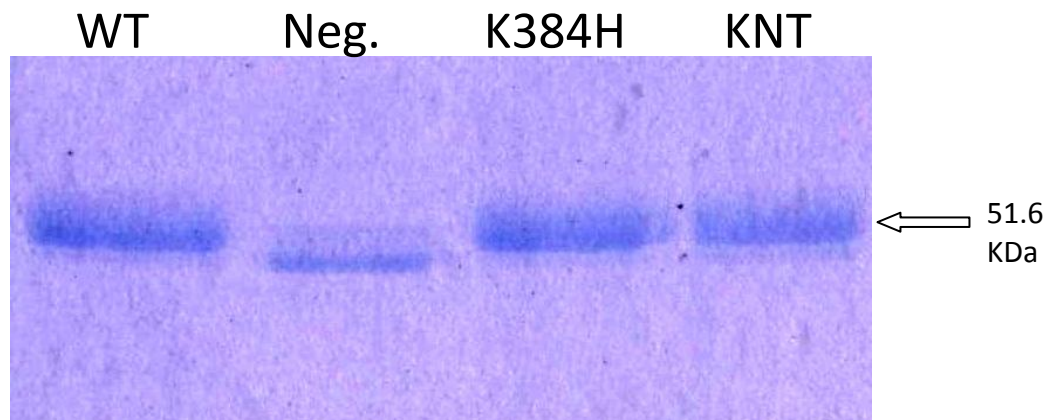


Figure 2.11

Mechanical shearing of SJW134 cells used in partial mutant motility assay. Overnight cultures used to inoculate motility gels in Figure 2.8 were subject to mechanical shearing and then separated on a 12.5% acrylamide SDS-PAGE gel. The negative control (Neg.) is SJW134 transformed with the pTrc99A empty vector. The protein band visible for the negative control is likely the FliD hook protein, which has a molecular weight of 50 KDa. Arrow marks indicate where the FliC protein is located.

2.3.3 Trypsin Digest Analysis of Flagellin Variant Stability

Limited proteolysis of flagellin proteins with trypsin, a serine protease, is often used to assay the folded state, stability and self-assembly of flagellin^{65,70,71}. Monomeric flagellins are much more susceptible to protease digestion than polymerized flagellins that comprise a flagella fiber. Thus, a trypsin digest assay should reveal the degree to which a sample of purified flagellin is dissociated into soluble monomers. SDS-PAGE analysis indicated that the polymerized flagella assembled from the three FliC variants were as resistant to ~20 hour trypsin digestion as the WT flagella fibers (Figure 2.12). The flagellins monomers were obtained through heat depolymerization of the polymerized flagella fibers. When the monomers were subjected to limited trypsin digestion, they were degraded into smaller protein fragments in a time-dependent pattern that has been observed previously^{70,71,95}, yielding smaller stable core fragments at various times, e.g. the F41 fragment. There are a total of 42 possible trypsin cleavage sites in wild-type FliC flagellin, as determined by ExPASy tools PeptideCutter

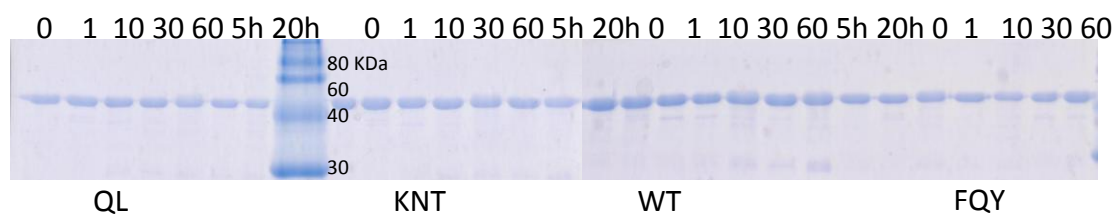


Figure 2.12

SDS-PAGE analysis of trypsin digests of polymeric WT and metal-site variant flagella fibers. Polymeric flagella were subject to limited proteolysis with 1:300 trypsin to FliC at room temperature for up to 20 hours. Gel lanes are labeled with the time at which the sample was taken in minutes or in hours with the addition of the 'h'. Gels were run on 15% acrylamide cross-linking using the SERVA Recombinant SDS PAGE Protein Marker as a molecular weight standard.

(http://web.expasy.org/peptide_cutter/) using the FliC protein ID, P06179. In monomeric FliC flagellin, the disordered D0 domain is the first to be cleaved when subjected to trypsin, which reduces the native ~51 KDa protein to a partial fragment approximately 40 KDa in size (comprised of the D1-D3 domains), termed the F41 fragment (PBD 1IO1). This F41 fragment results from removal of the N-terminal sequence region from residues 1 to 53 and the C-terminal sequence from residues 451-494. Further digestion of flagellin with trypsin results in proteolysis of the D1 domain region of the F41 fragment to form a smaller 27 KDa fragment, comprised primarily of the outer D2/D3 domains, referred to as the F27 fragment. This digestion step removes additional D1 domain residues up to Lys 179 and Arg 422 on the N-terminal and C-terminal sequence regions, respectively. The KNT variant has one less trypsin digestion site than WT, because trypsin selectively cleaves at Lys residues and in this variant, native residue Lys 384 was mutated to a His residue. However, this is not one of the major cleavage sites that results from limited proteolysis digestion to form the F41 or F27 fragments. All three of the FliC variant monomers yielded similar digestion fragments in a time-dependent pattern similar to WT monomers (Figure 2.13). This result indicated that all three of the flagellin variants were folded in a manner similar to the native protein, rather than existing in a disordered, random-coil state. However, when compared to WT monomers, the QL variant monomer appeared to have a less stable tertiary protein structure, as the D1-D3 domains, which correspond to the F41 fragment, were more rapidly digested. The F41 fragment, which was stable in the WT for up to ten minutes, was digested after approximately one minute in the QL variant.

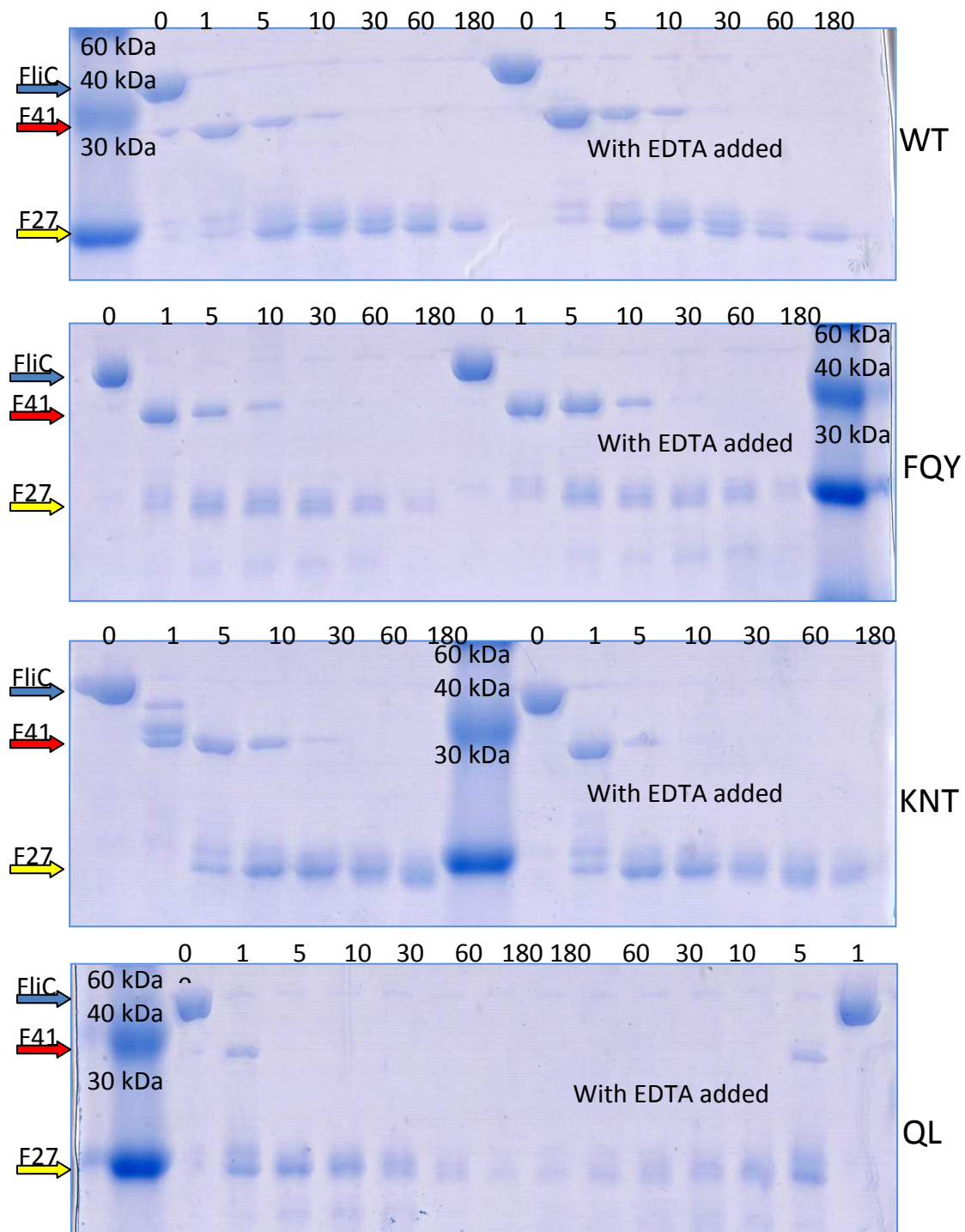


Figure 2.13

SDS-PAGE analyses of trypsin digest of flagellin monomers with and without EDTA. Gels are 15% acrylamide loaded with FliC protein subjected to limited proteolysis with trypsin. Gel lanes are labeled with the time at which each sample was collected after the start of trypsin digestion. EDTA was added to the protein samples on the right side of each gel, as indicated. Arrows are used to indicate the location of the full length FliC, the F41 fragment and the F27 fragment. The SERVA Recombinant SDS-PAGE Protein Marker was used as the molecular weight standard.

The QL variant has mutations in the D2 domain, which may play a role in the observed differences in trypsin digestion patterns by causing the protein to be more disordered or in a conformation that is more accessible to trypsin. The transition metal chelator EDTA was added to the monomeric samples to determine if the removal of metal ions from the solution would cause any change in the trypsin digest pattern. No change was apparent in the time-dependent trypsin digestion patterns for the WT, FQY and QL variants. However, the KNT variant exhibited a dramatic shift in its stability or resistance to trypsin digestion as a function of the presence or absence of the transition metal ion chelator, EDTA. Without EDTA, the KNT variant was degraded in a similar time-dependent fashion to WT. However, in the presence of 1 mM EDTA, the F41 fragment was digested much more quickly than WT. For WT with and without EDTA and for KNT without EDTA, the F41 fragment was present after 10 minutes with trypsin. However, when the EDTA was present in the trypsin digest with KNT, the F41 fragment was only stable for about five minutes. This result suggests that removal of any free transition metal ions from the protein results in destabilization of the KNT flagellin protein variant or allows easier access to a trypsin cleavage site. Because the Lys 384-His mutation rendered this potential trypsin cleavage site inert, there should actually be one less site for trypsin to act on in the D2 domain. As discussed previously, metal ions are known to stabilize the tertiary structure of many proteins. Thus, the apparent destabilization of the KNT variant upon the removal of metal ions is another possible indication of transition metal ion binding by the KNT mutant that was not observed in WT flagellin or the other two variants. This result is also consistent with the previous

result where this variant exhibited metal-dependent function in agar motility assays (Figure 2.9).

2.3.4 Circular Dichroism (CD) Spectroscopic Analysis of Flagellin Secondary Structure

CD spectroscopic analysis can yield estimates of the overall composition of alpha helix, beta sheet, and random coil secondary structures for a protein in solution. CD analysis of the monomeric flagellin proteins yielded similar Mean Residue Ellipticity (MRE) spectra to those previously observed for monomeric *Salmonella* flagellin proteins^{60,70,96,97} (Figure 2.14). Secondary structures such as alpha helix, beta sheets,

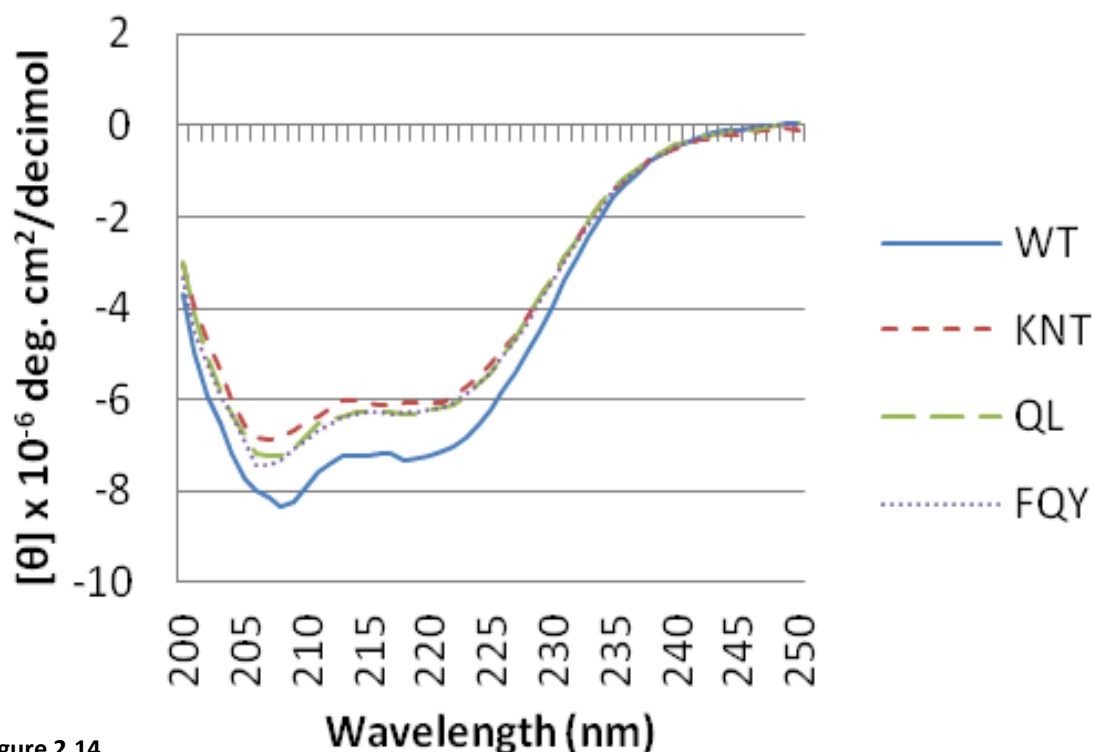


Figure 2.14

CD spectra of the three flagellin variants compared with wild-type. The FliC protein concentration was corrected for in units of Mean Residue Ellipticity. CD spectra were measured over the wavelength range of 200 to 250 nm for the flagellin variants and wild-type.

random coil and beta turns can be determined by signature peaks. If the protein was comprised mainly of beta-turns, there would be a negative peak at 218 nm and a positive peak at 195 nm. If the protein was a random coil, it would have low ellipticity below 210 nm with a negative peak at 195 nm. Alpha helix proteins have a double negative peak with one minimum at 222 nm and the other at 208 nm and then becoming positive at 195 nm⁹⁸. The flagellin CD spectrum more closely resembles an alpha-helical protein spectrum, which has been observed previously for this protein^{60,70,96,97}. The shape of the WT flagellin CD spectrum indicated that it was properly folded and that the three flagellin variants were also folded and had similar secondary structures as WT, despite the presence of two or three point mutations in the different domain regions. The slight shift in peak intensity observed could be due to errors in determining protein concentration either by the UV₂₈₀ reading or the inherent error in the extinction coefficient. According to CD secondary structure analysis software K2D2⁹⁹ and K2D3¹⁰⁰ (<http://www.ogic.ca/projects/k2d2/>), there was very little difference in the predicted secondary structures between WT and the FliC variants (Table 2.8). The differences in the composition of the secondary structures varied greatly between each of the two predictions, as well as compared to what has previously been predicted from CD spectroscopy data. Therefore, it is difficult to determine the percent composition of each type of secondary structure in flagellin.

Table 2.8
Predictions of Percent Alpha-Helix and Beta-Sheet of *Salmonella* FliC Protein Based on CD Spectroscopy Data Using K2D2 and K2D3 Prediction Software

	K2D2		K2D3		Uratani <i>et al.</i> ⁹⁷	
	<u>α-Helix</u>	<u>β-Sheet</u>	<u>α-Helix</u>	<u>β-Sheet</u>	<u>α-Helix</u>	<u>β-Sheet</u>
WT	56.8%	5.1%	9.8%	29.8%	20%	30%
KNT	50.2%	7.4%	5.2%	34.6%		
FQY	50.2%	7.4%	5.8%	33.9%		
QL	50.2%	7.4%	5.8%	34.0%		

2.3.5 Transmission Electron Microscope Images

By utilizing transmission electron microscopy (TEM), bacterial cells can be visualized with their attached flagella fibers to better determine if the flagellin protein was exporting and assembling. TEM imaging was performed on the WT and three metal-site flagellin variants with live *Salmonella* cells. The TEM images indicated that the WT and all three variants produced detectable extracellular flagella fibers (Figure 2.15). Furthermore, the TEM images of the flagella fibers formed *in vivo* from the three variants did not indicate any obvious differences compared to wild-type flagella. This result demonstrated that the reduced motility phenotype observed for the KNT and QL variants in 0.3% motility agar was probably not due to lack of export and assembly of these mutated flagellins to form flagella fibers. These results are also consistent with the similar amounts of flagellin proteins incorporated into extracellular flagella fibers for wild-type and each of the three variants via SDS-PAGE analysis (Figure 2.6), following isolation of extracellular flagellins by mechanical shear, as discussed in Section 2.2.4 *Salmonella* FliC Flagellin Expression and Purification. By visual inspection, the flagella fibers produced by *Salmonella* cells expressing the WT flagellin and the QL, KNT and FQY flagellin variants all exhibited a similar wavy or curly morphology pattern. These results

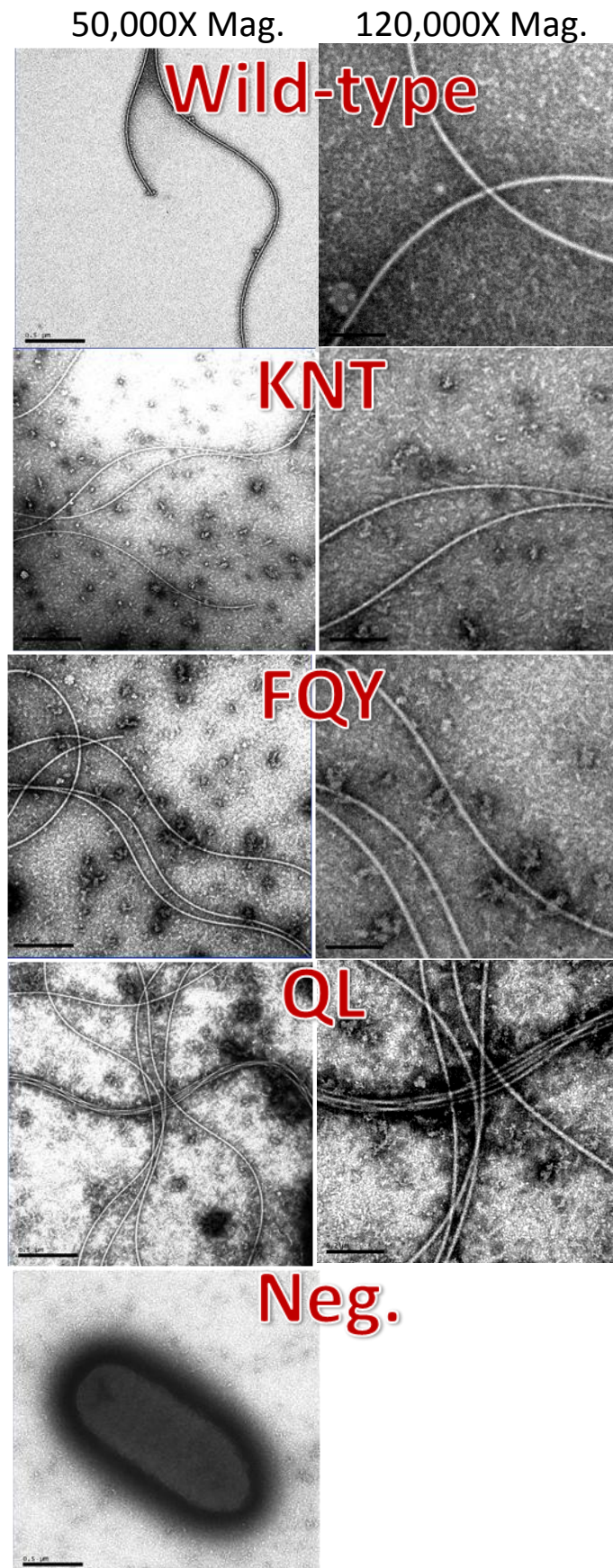


Figure 2.15

TEM images of bacteria and flagella. TEM images of flagella formed *in vivo* from the three flagellin variants and wild-type. The name of each variant is labeled on the corresponding images with the magnification for each image shown at the top of the figure. Neg. is *Salmonella* SJW134 with pTRc99A empty vector. Scale bar on the left is 0.5 μ m, scale bar on the right is 0.2 μ m.

indicate that the reduced motility observed in the KNT variant was not due to formation of the straight supercoiled form of flagella filaments.

2.3.6 Analysis of Flagellin Metal binding via ICP Metal Analysis

A key question of this study was whether or not any of the three rationally-designed metal binding sites in the FQY, QL, or KNT flagellin variants were able to bind transition metal ions, such as Zn^{2+} . Purified, recombinant CA2 and both the monomeric and the polymerized forms of each of the three variants and WT were incubated with Zn^{2+} ions and then dialyzed to remove an unbound, excess metal ions. The samples were then sent to The Center for Applied Isotope Studies at the University of Georgia for analysis of 20 different types of metals using ICP-ES. Raw ICP-ES data can be found in the Appendix. Buffer samples with 0, 5 or 10 parts per million (ppm) of zinc sulfate were submitted in addition to the protein samples to correct for the buffer at the suggestion of Rebecca Auxier, who operates the ICP-ES at the Center for Applied Isotope Studies. The ICP-ES analysis of these three buffer controls generated the equation:

$$y = 4.4229x - 2.5974 \quad (\text{Eq. 2})$$

where y is the corrected concentration for Zn^{2+} in ppm and x is the raw Zn^{2+} in ppm. This equation is a correction factor to adjust the measured ppm of Zn^{2+} to the actual value. The corrected ppm concentration value of Zn^{2+} was then converted to molarity (M) using the zinc atomic weight of 65.38 g/mol. Flagellin protein concentrations (M) were

determined using UV₂₈₀ absorbance values measured before the samples were submitted. These two concentration values were then used to determine the metal:protein ratio. Purified, recombinant CA2, a known zinc metalloenzyme, was used as a positive control for zinc binding. The data indicates that there are two zinc ions for every CA2 protein (Table 2.9) and the previous thought was that CA2 should only bind one zinc ion per protein molecule. However, a second metal binding site has recently been recognized in CA2, in several published crystal structures¹⁰¹⁻¹⁰³. WT FliC was used as a comparison for zinc ions bound by the three mutated FliC proteins. The ICP-ES data indicated that both monomeric WT (WTm) and polymeric WT (WTp) FliC bind approximately one zinc ion per protein. Currently, there are no known native metal binding sites on FliC, although the one native His residue, His 388, represents a potential candidate for a zinc ligand, and numerous other acidic residues, i.e., 37 Asp and 17 Glu residues (Table 2.1), or it may also interact electrostatically with metal cations. In fact, the Glu 386 residue is located only two residues away from the native His residue, which

Table 2.9
Determining Metal:Protein Ratio With Zinc from ICP-ES Data

Protein Sample^a	Protein Concentration (M)	Zn²⁺ Concentration (M)	Zn²⁺: Protein Molar Ratio
WTm	5.31X 10 ⁻⁶	6.44X 10 ⁻⁶	1.21
WTp	5.70X 10 ⁻⁶	7.59X 10 ⁻⁶	1.33
FQYm	5.42X 10 ⁻⁶	2.75X 10 ⁻⁶	0.51
FQYp	7.10X 10 ⁻⁶	2.27X 10 ⁻⁶	0.32
KNTm	7.49X 10 ⁻⁶	1.05X 10 ⁻⁵	1.40
KNTp	5.64X 10 ⁻⁶	1.09X 10 ⁻⁵	1.95
QLm	9.22X 10 ⁻⁶	3.20X 10 ⁻⁵	3.47
QLp	5.92X 10 ⁻⁶	1.15X 10 ⁻⁵	1.95
CA2	4.46X 10 ⁻⁶	8.36X 10 ⁻⁶	1.87

^aProteins are named with either 'p' for the polymeric flagella form or 'm' for the monomeric form.

together could potentially bind to a metal ion. Overall, the WT FliC protein is negatively charged with a theoretical pI of 4.79, so the interaction with zinc ions could be electrostatic. Figure 2.16 highlights all of the Asp and Glu residues and the single His residue within the wild-type flagellin protein. There are two major regions where these residues are located; one near the native histidine and the other in the D3 domain. If there is an electrostatic metal ion interaction for WT flagellin, it could occur at either of these two locations.

The most surprising result was that the monomeric form of the QL flagellin variant, (QLm), appeared to bind a substantial amount of Zn^{2+} compared to all of the

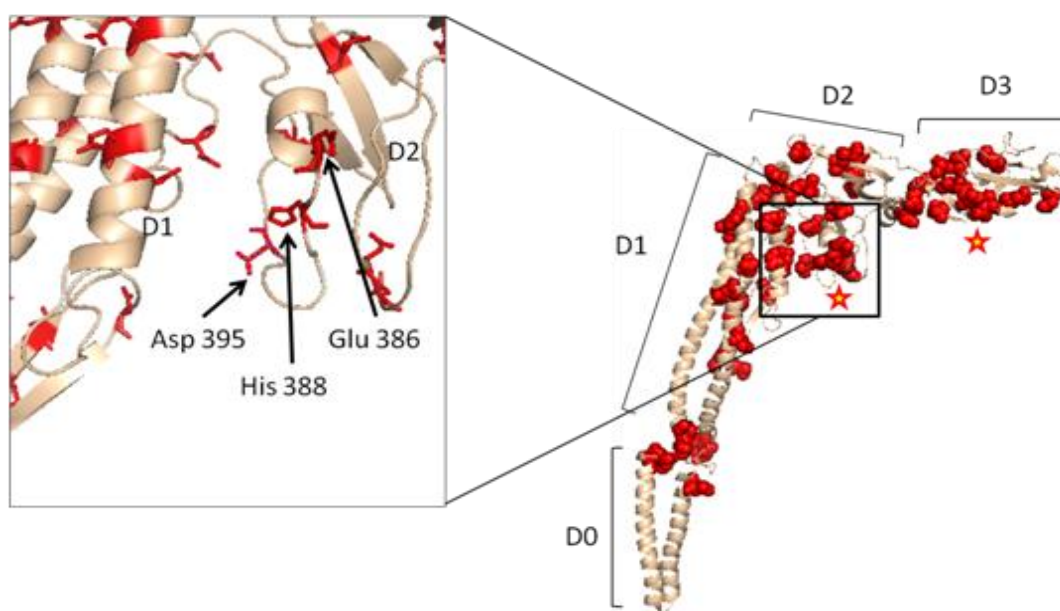


Figure 2.16

Ribbon structure of the wild-type FliC flagellin protein highlighting all of the His, Asp and Glu residues. All His, Glu and Asp residues are indicated with spheres. Regions with clusters of His, Asp and Glu residues are indicated with a star. One of those locations is in the D3 domain near where the D2/D3 pocket is located. The other is located where the one native His is located. The enlarged image shows the close proximity of both an Asp and a Glu residue to the native His.

other samples. Although this is preliminary data, it appears that Q_{Lm} binds more than three zinc ions per protein monomer, which is 2.26 zinc ions per protein more than its WT_m counterpart. Both the polymeric QL (Q_{Lp}) and KNT (KNT_p) bind approximately 0.6 more zinc ions per protein than the WT_p. The KNT variant monomer (KNT_m) binds approximately 0.2 zinc ions more per protein than WT_m. The differences observed between the monomeric and the polymeric QL and KNT variants implies that there is a shift in metal ion binding favorability, depending on the state of flagellin oligomerization, which was not observed between the monomeric and polymeric WT samples. Because the polymerized version of KNT binds more zinc than the monomer, it may be a possible explanation for the reduced motility phenotype that was rescued with the addition of the metal ion chelator EDTA. Both monomeric and polymeric forms of the FQY variant bound significantly less zinc ions than either form of WT FliC. This leads to the possibility that there is a slight change in folding that reduces any native metal ion interaction it may have. Two of the three mutations, Phe and Tyr to His, changed nonpolar residues to more polar residues, which could alter the folding of the D2/D3 domain. Because the D3 domain is one of the regions where there are many negatively charged residues located, it is possible that this repositions the negatively charged residues in a conformation that is less favorable for metal binding. This possible change in conformation, however, does not alter the ability of the protein to export, assemble and function in motility. If the metal:protein ratio observed for WT FliC is due to electrostatic interactions, the predicted pI value for the FQY variant increased to 4.94, which may indicate the protein was slightly less negative.

Although the flagellin proteins were not incubated with any other metal ion, there is a possibility that the proteins could interact with trace metals other than zinc that are present in the buffer solution. Therefore, the proteins were also analyzed for interaction with the transition metals, Co, Cu, Ni, Fe and Mn (Table 2.10). Although ratios of metal ion to protein were very small, there was one sample that appeared to possibly have some slight interaction. The KNTp variant, which has the motility phenotype, exhibited a small degree of association with Cu. It was only about one Cu ion for every seven FliC monomers; however it was the highest observed trace metal ion to protein ratio observed. CA2 also bound approximately one Cu ion for every five proteins. This result is consistent with the possibility for the KNT FliC variant, which was modeled to have the same metal ion binding geometry as CA2, to interact with the Cu²⁺ ion.

Table 2.10
Determining Metal:Protein Ratio of Trace Metals from ICP-ES Data

	<u>Co:Protein</u>	<u>Cu:Protein</u>	<u>Ni:Protein</u>	<u>Fe:Protein</u>	<u>Mn:Protein</u>
WTm	-0.001	0.010	0.010	0.014	0.003
WTp	0.002	0.006	0.022	0.002	0.001
FQYm	0.001	0.043	0.030	0.010	0.003
FQYp	0.004	0.021	0.018	0.019	0.001
KNTm	0.006	0.015	0.017	0.018	0.002
KNTp	0.006	0.151	0.017	0.041	0.006
QLm	0.002	0.075	0.010	0.014	0.003
QLp	0.005	-0.042	0.012	0.012	0.000
CA2	-0.006	0.213	0.005	-0.006	-0.007

Metal:protein ratios were calculated via M to M concentration determined using the metal ions molecular weight.

The most notable trace metal:protein interactions are indicated with bold type.

2.3.7 4-NPA Esterase Assay

WT and each flagellin variant were assayed for ester hydrolysis activity using 4-NPA as a colorimetric substrate and CA2 as a positive control. The protein samples were incubated with zinc prior to assay. Both the monomeric and polymeric forms of the flagellin proteins were used to determine if there was any activity. The absorbance of the samples was measured over a total interval of 15 minutes and compared with the change in absorbance observed for CA2. In the end, there appeared to be no detectable esterase enzymatic activity for the mutated flagellins (Table 2.11), whereas the CA2 exhibited a high rate of ester hydrolysis.

Table 2.11
4-NPA Hydrolysis Assay Percent Change of FliC Mutants Compared to CA2 and WT

<u>Protein Name^a</u>	<u>Percent Activity (%)</u> <u>with Standard</u> <u>Deviation^b</u>	<u>Average Overall</u> <u>Change at Abs₃₄₈^c</u>	<u>Average V_{max}</u> <u>(Abs₃₄₈ s⁻¹)^c</u>
CA2	99.9 ± 8.8	5.6 x 10 ⁻²	5.6 x 10 ⁻⁵
FQYm	-1.7 ± 1.8	-1.7 x 10 ⁻³	-8.3 x 10 ⁻⁷
FQYp	-0.8 ± 2.1	-2.0 x 10 ⁻³	-1.2 x 10 ⁻⁶
KNTm	-2.0 ± 1.4	-1.8 x 10 ⁻³	1.7 x 10 ⁻⁷
KNTp	1.2 ± 1.8	-8.7 x 10 ⁻⁴	-5.0 x 10 ⁻⁷
QLm	0.4 ± 0.4	-5.0 x 10 ⁻⁴	1.2 x 10 ⁻⁷
QLp	1.8 ± 0.8	-5.7 x 10 ⁻⁴	-1.7 x 10 ⁻⁷
WTm		-7.0 x 10 ⁻⁴	-1.7 x 10 ⁻⁷
WTp		-1.6 x 10 ⁻³	-1.2 x 10 ⁻⁶

^aNames ending in 'm' are monomers and ending in 'p' are polymers

^bValues calculated as total change from beginning to end of assay. Percents are calculated with WT being the negative control and CA2 being the positive control.

^cValues are corrected for by the change observed in just the buffer.

2.4 Bacterial Flagellin Engineering Main Conclusions

Using rational design with the assistance of 3D computer modeling software, locations for mutations in the FliC protein were selected to design a potential active site modeled after CA2. After introducing the selected histidine mutations into the FliC protein, it could be concluded that the protein was still exported and assembled into functional fibers. The reduced motility observed for the KNT variant was rescued upon the addition of the metal chelator EDTA. There was also an observed change in the trypsin digest pattern for KNT and QL. Both the QL and KNT variants have a less stable F41 fragment; however, the KNT F41 fragment is less stable upon removal of free metal ions with the chelator EDTA. The altered trypsin digest pattern and the EDTA rescue of motility were initial indicators that there may be a role of metal ion interaction for the KNT variant.

Based on the preliminary ICP-ES data, a metal binding site was successfully engineered into *Salmonella* flagellin using a rational design approach. Of the three sites that were designed, two appear to have some affinity for the transition metal ion zinc. The putative metal binding sites were chosen based on the known properties of many multidomain enzymes such that their active sites usually form a concave pocket and are located between protein domains. This allows for some conformational rearrangement to occur during substrate binding, catalysis, and product release. The recombinant CA2 that was used as a positive control for the ICP-ES data indicated that it bound approximately two zinc ions for every protein which is consistent with previously published CA2 data. The two FliC variants that appear to have successful metal binding

sites reside in the largest predicted pocket of the FlhC flagellin protein located between the D1 and D2 domains. It appears that the QL variant binds transition metals, such as the zinc (II) ion, more prevalently in the monomeric state as compared to all the other mutants. The polymerized form of the KNT variant had more prevalent zinc binding than the monomeric form and both forms exhibited higher levels of zinc binding than WT. The observation that the KNT polymerized flagellin binds more metal than the monomer suggests that metal binding may play a role in the EDTA rescued motility phenotype observed. This observation, combined with the altered trypsin digest of the monomer when EDTA is present, indicates that there is indeed some transition metal interaction for the KNT flagellin variant. If metal binding causes this nearly non-motile phenotype in motility agar, this could give some insight into why the flagellin protein has evolved to have only one native His and no Cys residues, which are the most commonly observed metal ion ligand. The TEM images also suggested that the reduced motility for the KNT variant was not due to flagella binding together or lack of flagella. The KNT variant was also observed to have some association with the trace copper ions that were in solution, suggesting a possible metal ion preference. There appears to be less metal ion association for the FQY variant as compared to WT. This could be due to a change in the protein fold of the D2/D3 domain caused by the introduction of the His mutations.

Although there is preliminary evidence for metal binding by at least two of the three FlhC protein mutants, there appears to be no significant esterase activity with 4-NPA as a substrate.

CHAPTER 3

ENGINEERING THIOREDOXIN OF FLITRX TO PERFORM NUCLEOPHILIC HYDROLYSIS OF 4-NPA ESTER SUBSTRATE

3.1 Background for Protein Engineering of FliTrx Catalytic Site

The FliTrx system is comprised of the *E. coli* thioredoxin protein, TrxA, genetically inserted as a fusion in the middle of the partially deleted D2-D3 domain region of the *E. coli* FliC flagellin protein. Part of the middle, hypervariable region of the wild-type *fliC* gene was deleted, and the *trxA* gene was inserted at regions corresponding to FliC residues Gly 243 and Ala 352⁶³ (Invitrogen, Carlsbad, CA). The multiple cloning site (MCS) was removed and replaced with a single RsrII restriction site and a non-native Cys residue was mutated to a Ser residue⁶⁹. The system was initially designed for screening libraries of random peptides for interactions with other proteins, e.g. monoclonal antibodies. Using the non-motile GI826 *E. coli* strain from Invitrogen, which has deletions of the *fliC* flagellin and *motB* flagellar motor protein from the genome, the FliTrx protein is exported and assembles into flagella fibers, but the flagella do not rotate, enabling more stable binding interactions of displayed peptides with other ligands, e.g. antibodies.

E. coli thioredoxin is a 108 residue protein that has previously been engineered to perform a variety of different non-native functions; some of these thioredoxin variants were rationally designed using the crystal structure of the folded protein^{43,78}. Bolon and Mayo⁷⁸ made several point mutations to thioredoxin and modified the overall

function of the protein to perform histidine-mediated nucleophilic hydrolysis of the ester substrate, *p*-nitrophenyl acetate (4-NPA). The mutations were designed using the program ORBIT (Optimization of Rotamers By Iterative Techniques)¹⁰⁴. The algorithm of this protein design software functions by identifying both the potential locations of catalytic sites in the native protein structure and predicting appropriate substitution mutations of these target residues that are necessary to generate a new ligand binding site. With predicted mutations generated for the top ten potential active sites, a total of two thioredoxin variants were engineered via standard molecular biology protocols, using the software predictions, and tested for the desired function. The two thioredoxin variants, named PZD1 and PZD2, that were generated by this procedure, successfully performed hydrolysis of 4-NPA ester substrate at a rate above background. Residue Tyr 70 was mutated to Ala in both variants (Y70A) to create more space for substrate binding in the active site. The ORBIT software also predicted that mutating thioredoxin residue Asp 26 to Ile (D26I) would increase the thermostability, and therefore, was also included in both of the variants. The PZD1 thioredoxin variant also contained an additional Phe 12 His (F12H) mutation, while the PZD2 thioredoxin variant contained two other additional mutations, Phe 12 Ala (F12A) and Leu 17 Ala (L17A). If thioredoxin is folded properly as an internal fusion partner in *E. coli* flagellin, and the previously described esterase active site region is solvent exposed, then the same mutations originally introduced by Bolon and Mayo⁷⁸ into thioredoxin could potentially result in the same catalytic 4-NPA esterase activity in the FliTrx system upon mutagenesis.

3.2 Methods and Procedures

3.2.1 Media Prep

Rich Media (RM) was prepared by dissolving 20 g of casamino acids, and 10 ml of 100% glycerol in 890 ml of water. The solution was then autoclaved for 20 minutes on liquid cycle. Once cooled, 100 ml of 10X M9 salts, 1 ml of 1 M MgCl_2 , and 1 ml of 100 mg/ml ampicillin (Amp) were added. The media was stored at 4°C for up to 2 weeks.

RM with glucose (RMG)-Amp agar plates were prepared by dissolving 20 g of casamino acids, and 15 g of agar in 875 ml water. The solution was then autoclaved for 20 minutes on a liquid cycle. Once it cooled to approximately 55 °C, 100 ml of M9 salts, 1 ml of 1 M MgCl_2 , 25 ml of 20% glucose (Sigma-Aldrich, St. Louis, MO) and 1 ml of 100 mg/ml Amp were added. It was mixed well and then the plates were poured. Plates were good for up to two weeks when stored at 4 °C in the dark.

3.2.2 Plasmid

The pFliTrx plasmid was purchased from Invitrogen (Grand Island, NY) and had the random peptide loop removed⁶⁹ (Figure 3.1). It uses the P_L promoter from bacteriophage λ , which is tightly regulated and induced by tryptophan. The bacteriophage λ cI repressor blocks transcription by binding to the operator region

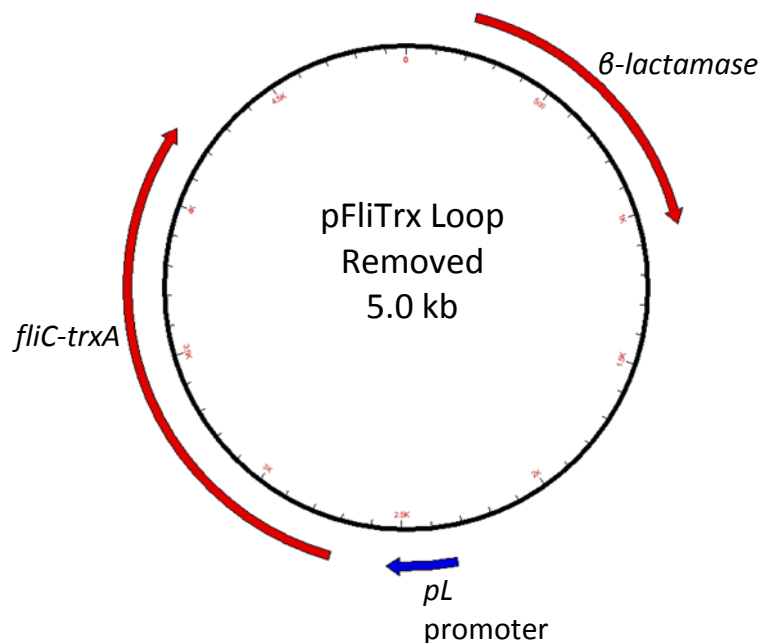


Figure 3.1

Computer generated image of the pFliTrx plasmid. The *fliC-trxA* gene encodes the FliTrx protein. Genes are labeled accordingly. Image was generated using Discovery Studio Gene (Version 1.5).

upstream from the P_L promoter. In the *E. coli* G1826 bacteria genome, the *trp* repressor regulates the expression of the *cl* repressor. When growth media is tryptophan free, the *cl* repressor protein is transcribed and binds to the P_L promoter operator region and prevents the transcription of the downstream gene. Induction of gene transcription can be achieved by adding tryptophan to the growth media, which shuts down the production of the *cl* repressor gene product.

3.2.3 Site-Directed Mutagenesis of Thioredoxin *trxA* Gene in the pFliTrx Plasmid

Mutations were introduced into the *trxA* thioredoxin gene in the pFliTrx plasmid, based on previous mutations described by Bolon and Mayo⁷⁸. Two successful variants (PZD1 and PZD2) were chosen as model targets to engineer into the FliTrx protein with

the loop removed. DNA sequencing to verify removal of the loop is described in the Appendix. Primers were designed using the Agilent Technology Quikchange Primer Design web-based software

(<https://www.genomics.agilent.com/collectionsubpage.aspx?pagetype=tool&subpagetype=toolqcpd&pageid=15>) and ordered from Integrated DNA Technologies (Coralville, IA)

and are listed in Table 3.1. Nearest neighbor temperatures were calculated using the calculator at Northwestern

(<http://www.basic.northwestern.edu/biotools/oligocalc.html>). Site-directed mutagenesis was performed using Phusion™ High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA) in 200 µl PCR tubes (Midwest Scientific, Valley Park, MO). PCR conditions are described in Tables 3.2 and 3.3, which are based on the suggested protocol, which

Table 3.1

FliTrx Site-Directed Mutagenesis Primers

<u>Primer name</u>	<u>Primer sequence</u>	<u>Nearest neighbor temp.</u>
D26I	5' GGGGCGATCCTCGTC ATT TTCTGGGCAGAGTG 3' 5' CACTCTGCCCAGAAA AAT GACGAGGATCGCCCC 3'	68
Y70A	5' GCACTGCGCCGAAA GCT GGCATCCGTGG 3' 5' CCACGGATGCC AGC TTTCGGCGCAGTGC 3'	70
F12A before L17H	5' AATTATTCACCTGACTGACGACAGT GCT GACACGGATGTACTC 3' 5' GAGTACATCCGTGTC AGC ACTGTCGTCAGTCAGGTGAATAATT 3'	69
L17H	5' TGACACGGATGTAC CAC AAAGCGGACGGG 3' 5' CCCCCTCCGCTTT GTG TACATCCGTGTCA 3'	65
F12H	5' AATTATTCACCTGACTGACGACAGT CAT GACACGGATGTACTC 3' 5' GAGTACATCCGTGTC ATG ACTGTCGTCAGTCAGGTCAATAATT 3'	68

The bolded and highlighted base pairs signify the residue being mutated. Mutations are blocked by which variant they are in; D26I and Y70A are in both, F12A and L17H are in PZD2 and F12H is in PZD1.

Table 3.2
PCR Reaction Volumes and Concentrations with Phusion™ Polymerase for Site-Directed Mutagenesis

Item	Volume (μL)	Concentration
Phusion™ 5X HF Buffer	5	5X
DNA template	1	5-100 ng
Forward Primer	0.5	125 ng
Reverse Primer	0.5	125 ng
dNTP	0.5	10 mM
Sterile water	17.5	
NEB Phusion™ Polymerase	0.25	2 U/μL
Total	25.25	

Table 3.3
PCR Thermocycler Settings for Site-Directed Mutagenesis on pFliTrx Using Phusion™ Polymerase

Cycle step	Temp.	Time	Cycles
Initial denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	25
Annealing	NN-3 °C	30 s	
Extension	72 °C	2m	
Final extension	72 °C	5 min	1
	4 °C	Hold	
Final extension	72 °C	5 min	1
	4 °C	Hold	

can be found at <http://www.neb.com/nebecomm/products/protocol631.asp>. After completion, the PCR DNA-template DNA mixture was incubated with 0.75 μL DpnI restriction enzyme at 37 °C for 1 hour to digest the methylated template DNA. Using 50 μl of electro-competent GI826 *E. coli*, 1 μl of the PCR product was gently added and the cells were incubated on ice for 15 minutes. After the incubation time, the cells were electroporated using a Bio-Rad MicroPulser™ (Hercules, CA), using the EC1 bacteria setting. Shocked cells were immediately suspended in 800 μl of warm SOC and shaken in

an incubator at 200 rpm, 30 °C for 1 hour. Cells were then plated on fresh RMG-Amp plates and incubated overnight at 25-30 °C. Isolated colonies were then picked from the center of the colony on the transformation plate and used to streak another RMG-Amp plate, which was grown overnight under the same conditions. Using a sterile needle, cells were picked from an isolated colony on the streak plate and used to inoculate the colony PCR reaction tube. Conditions for colony PCR can be found in Tables 3.4 and 3.5.

Table 3.4
Colony PCR Reagent Volumes and Concentrations

Item	Volume (μL)	Concentration
10x Taq buffer	2.5	10x
Bacteria		
Forward Primer	0.5	125 ng
Reverse Primer	0.5	125 ng
dNTP	0.5	10 mM
Sterile water	20.875	
Taq Polymerase	0.125	2 U/μL
Total	25.25	

Table 3.5
Colony PCR Thermocycler Settings

Cycle step	Temp.	Time	Cycles
Initial denaturation	95 °C	30 s	1
Denaturation	95 °C	30 s	30
Annealing	65 °C	60 s	
Extension	68 °C	5 min	
Final extension	68 °C	5 min	1
	4 °C	Hold	

PCR reaction products were analyzed by gel electrophoresis, using a 1% agarose gel with 0.5 µg/ml ethidium bromide and imaged using a UV₃₀₂ light source and a Kodak gel imaging station to determine if the plasmid was present. Colonies with plasmid were used to inoculate 7 ml of RM-Amp media in a 34 ml Pyrex® glass culture tube (Corning, Tewksbury, MA) in a CEL-GRO® tissue culture rotary incubator (Lab-Line, Maharashtra, India) rotating at speed setting 7 (approx. 100 rpm) at 30 °C. The plasmid DNA was then purified from the overnight culture using the QIAprep® Spin Miniprep kit (Qiagen, Valencia, CA) with a final elution volume of 30 µl. Plasmids were then submitted for DNA sequencing at the University of Michigan Sequencing Core. DNA sequence results were aligned using ClustalW2 software in the ExPASy tools at EBI (European Bioinformatics Institute) available at (<http://www.expasy.ch/tools/>) and can be found in the Appendix.

3.3 FliTrx Engineering Conclusions

Thioredoxin has been used in several previous experiments for enzyme engineering, allowing modifications and introductions of new enzymatic activity^{43,78}. As a protein fusion with FliC in the FliTrx system, thioredoxin could potentially be modified using previously successful modifications to better determine if the FliTrx thioredoxin can be altered to perform the same activities as previously modified thioredoxins, e.g. histidine-mediated nucleophilic hydrolysis of *p*NPA⁷⁸. The FliTrx system allows for an exportable, purifiable and polymerizable fusion protein system of FliC and thioredoxin. FliTrx has been used in the past for many engineering feats such as nickel

sequestration⁶⁷ as well as various mineral and metal coated nanoparticles^{51,54,69}. By using a previously successful thioredoxin engineering project as a model, it was proposed to introduce those same mutations into the FliTrx protein. Of the two or three mutations that were required for each of the esterase variants, only one was successfully completed. Due to the difficulty in introducing the mutations into the FliTrx system, esterase activity was not able to be tested for.

CHAPTER 4

CATALYTIC ENGINEERING OF THE FLAGELLIN PROTEIN - FUTURE DIRECTIONS

4.1 FliTrx Engineering Future Directions

The thioredoxin modifications performed by Bolon and Mayo⁷⁸ introduced an esterase activate site by rational design. The same modifications were proposed to be introduced into the thioredoxin in the FliTrx fusion protein. Not all of the mutations were successfully introduced into FliTrx, thus additional mutagenesis would need to be performed on the pFliTrx plasmid. Because the GI826 strain for expression of FliTrx poses some issues, it would be beneficial to try to find a different strain that is not as sensitive to temperature and media. The bacterial system would have to allow for the export of flagellin to assemble into flagella and have any native flagellin genes removed in addition to disrupting the motor so the flagella are not rotated. The FliTrx gene can then be removed from the pFliTrx and inserted into a cloning vector that is easier to use. This modification would allow for a more stable system to work with pFliTrx.

In addition to this set of mutations for esterase activity by hydrolysis of 4-NPA, there are other previously published experiments that would be worth testing as well. For example, Benson *et al.*⁴³ successfully engineered an iron binding site with superoxide dismutase activity into thioredoxin. These mutations could be introduced into FliTrx protein in the thioredoxin region to also explore whether or not the same engineering experiments would be successful in the FliC-TrxA fusion system FliTrx.

4.2 Flagellin Engineering Future Directions

Flagellin engineering is a field of research that has much room for expansion. The flagellin protein is easily manipulated through genetics and can very easily be purified from bacterial cultures and studied. Due to limitations in resources, many of the additional experiments that would be ideal to fully determine if this project was successful could not be performed at this time.

Motility of *Salmonella* expressing the mutated FliC variants has been investigated in motility media but not in aqueous conditions. There are a variety of methods that have been used previously, such as dark-field microscopy¹⁰⁵ and fluorescent dyes^{35,106}, to visualize the flagella on the bacteria while also monitoring motility patterns. These techniques may determine what might be different about the flagella that are causing the reduced motility observed for the KNT and QL variants. While observing bacterial motility under the microscope, the addition of solutions such as metal ions and chelators could be used to determine if there are any changes in the flagella fibers or in the motility pattern of the bacteria.

The data presented in the previous chapters are preliminary results that suggest metal binding has been achieved in two of the three designed FliC mutants. ICP-ES can be used to determine if there is metal association between the protein variants and other transition metals such as Co^{2+} , Cu^{2+} , Ni^{2+} and Mn^{2+} by incubating the proteins with these specific metal ions. It could also be a way to determine if there is a preference of the type of metal bound by incubating the proteins with a cocktail of metal ions and

then using ICP-ES to determine what metals are more frequently associated, i.e., a competitive binding experiment.

Once it is determined that metal binding is taking place with the His metal binding FliC variant proteins, a next step would be determining the affinity in which these metal ions are bound. Isothermal titration calorimetry (ITC) can be used to determine the stoichiometry, equilibrium binding constant, and the thermodynamics of molecular binding events^{107,108}. ITC is very sensitive and can determine if there are any interactions, even those which may be weak. Zinc ion concentration can also be quantified through zinc binding fluorescent dyes. Using a number of quantifiable fluorescent dyes with different K_D values¹⁰⁹, the K_D value for the protein and zinc can be estimated through a competitive zinc binding assay with fluorescent quenching. Either ITC or fluorescent dye competitive binding assays are plausible ways to determine the affinity of the protein for metal ions such as Zn^{2+} .

Esterase activity, with esters such as 4-NPA, is not the primary catalytic activity of CA enzymes nor is it an activity that can be performed by all CA enzymes. The hydrolysis of CO_2 by CA enzymes is difficult to measure because it can happen at a rapid rate and requires measurement with precise instruments. By using stopped-flow spectroscopy with pH indicators, the steady state kinetic rate of CO_2 hydration can be determined¹¹⁰. It is possible for the mutated FliC proteins, which are modeled after CA2, to perform the hydrolysis of CO_2 and not the secondary esterase activity with 4-NPA.

Although it does not appear that significant enzyme activity was achieved with the esterase substrate 4-NPA, there are modifications that can be explored to improve

the activity. When designing the mutations in FliC, it was difficult to predict how the surrounding residues would respond to the local changes. Potentially, with a few more mutations in the shell of residues surrounding the metal binding site, the right organization could result in the desired catalytic activity. In CA2, there are additional residues in the surrounding active site environment that contribute to either the enzymatic activity or to substrate binding/coordination. First, the three primary His ligands that coordinate the zinc ion are in turn hydrogen bonded to other secondary side chain or peptide backbone ligands in the protein structure. Second, there is a hydrophobic patch formed by several nonpolar residues that is important for CO₂ substrate binding. Third, there is a fourth His residue located ~8 Å from the catalytic zinc ion that performs a critical role as a proton shuttle in the process of proton transport in and out of the active site. If CA2-like function is to be successfully engineered into FliC, it will be necessary to identify how the area surrounding any designed metal binding sites could be altered to better improve the chances at achieving activity. To determine which residues to mutate using a rational design approach, the 3D structure of the metal binding site with either X-ray crystallography or NMR would be required. Mutations for CA2-like activity in the FliC variants could also be explored through a directed evolution approach with random mutagenesis and screening for the desired activity.

Once metal binding and enzymatic activity is achieved, there are many different paths that can be taken to improve and modify function. There are many different types of enzymatic activities that use metal binding for their reactions and the FliC variants

could potentially be tailored to perform these various functions. In addition, the ability of the flagellin protein to be manipulated by way of the D3 domain while still retaining polymerizable function gives rise to the opportunity for chimeric proteins. Because the metal binding mutations are located in the D1/D2 domains, other proteins or enzymes of interest could be inserted into the D3 domain to add additional functions. This is versatile enough to even polymerize with other engineered flagellins to create a truly chimeric fiber that could have many potential applications as a biosensor or an industrial enzyme.

Appendix A

DNA Sequencing Results

FliC QL mutations final DNA sequencing results and abbreviated alignment with wild-type FliC

```

CNGGGGCGNTCCGCTTTCGGTGGACTCCTGCGACAGCAACTGAGGATGTGAAAAATGTACAAGTTGCAATGCTGATTG
ACAGAGGCTAAAGCCGCATTGACAGCAGCAGGTGTACCGGCACAGCATCTGTTGTTAAGATGTCTTATACTGATAATAA
CGGTAAACTATTGATGGTGGTTTAGCAGTTAAGGTAGGCGATGATTACTATTCTGCAACTCAAATAAAGATGGTTCCA
TAAGTATTAATACTACGAAATACACTGCAGATGACGGTACATCCAAAACGCACTAAACAACTGGGTGGCGCAGACGGC
AAAACCGAAGTTGTTTCTATTGGTGGTAAACTTACGCTGCAAGTAAAGCCGAAGGTCACAACTTAAAGCACACCCTGA
TCATGCGGAAGCGGCTGCTACAACCACCGAAAACCCGCTGCAGAAAATTGATGCTGCTTTGGCACAGGTTGACACGTTAC
GTTCTGACCTGGGTGCGGTACAGAACCGTTTCAACTCCGCTATTACCAACCTGGGCAACACCGTAAACAACCTGACTTCT
GCCCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCGCAGATTCTGCAGCAGGCCGG
TACCTCCGTTCTGGCGCAGGCGAACCAGGTTCCGCAAAACGTCTCTTTACTGCGTTAAGGATCCGAATTCGAGCTCC
GTCGACAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCT
GATAAACAGAAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCCGTA
GCGCCGATGGTAGTGTGGGGTCTCCCATGCGAGAGTAGGGNAACTGCCAGGCATCAAATAAANCGAAGGCTCAGTCGA
AAGACTGGGGCCTTTCGTTTTATCTGTTGGTTTGTGCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCNNGGAGCGGAT
TTGAACGTTGCGAACACGCCGCNAGGGGGGGGGNNGGGNAAGGAACCCC

FliC      TACGCTGCAAGTAAAGCCGAAGGTCACAACCTTTAAAGCACAGCCTGATCTGGCGGAAGCG 1200
QL        TACGCTGCAAGTAAAGCCGAAGGTCACAACCTTTAAAGCACACCTGATCATGCGGAAGCG 413
          *****

```

FliC FQY mutations final DNA sequencing results and abbreviated alignment with wild-type FliC

GCCCANGCGTGTTCGGCCGACTCAGTTCACGGCGTGAAAGTCCTGGCGCAGGACAACACCCTGACCATCCAGGTTGGTG
 CAACGACGGTGAAACTATCGATATCGATCTGAAGCAGATCAACTCTCAGACCCTGGGTCTGGATACGCTGAATGTGCAAC
 AAAATATAAGGTCAGCGATACGGCTGCAACTGTTACAGGACATGCCGATACTACGATTGCTTTAGACAATAGTACTTTT
 AAAGCCTCGGCTACTGGTCTTGGTGGTACTGACCAGAAAATTGATGGCGATTTAAACATGATGATACGACTGGAAAATA
 TTACGCCAAAGTTACCGTTACGGGGGGGACTGGTAAAGATGGCTATTATGAAGTTTCCGTTGATAAGACGAACGGTGAGG
 TGA CTCTTGCTGGCGGTGCGACTTCCCCGCTTACAGGTGGACTACCTGCGACAGCAACTGAGGATGTGAAAAATGTACAC
 GTTGCAAATGCTGATTTGACAGAGGCTAAAGCCGCATTGACAGCAGCAGGTGTTACCGGCACAGCATCTGTTGTTAAGAT
 GTCTTATACTGATAATAACGGTAAACTATTGATGGTGGTTTAGCAGTTAAGGTAGGCGATGATTACTATTCTGCAACTC
 AAAATAAAGATGGTTCCATAAGTATTAATACTACGAAATACACTGCAGATGACGGTACATCCAAAACGCACTAAACAAA
 CTGGGTGGCGCAGACGGCAAAACCGAAGTTGTTTCTATTGGTGGTAAACTTACGCTGCAAGTAAAGCCGAAGGTCACAA
 CTTTAAAGCACAGCCTGATCTGGCGGAAGCGGCTGCTACAACCACCGAAAACCGCTGCAGAAAATTGATGCTGCTTTGG
 CACAGGTTGACACGTTACGTTCTGACCTGGGTGCGGTACAGAACCGTTTCAACTCCGCTATTACCAACCTGGGCAACACC
 GTAAACAACCTGACTTCTGCCGTAGCCGTATCGAANATTCCGACTACNCGACNGAANTTTCCAACATGTCTCGCGCGCN
 AAATTCTGCANCAGGCCGGTACCTCCNTTCTGGCGCAGGCGAACNAGNTCCCCAAAACGTCCCTCTCTTTACTGCNTTA
 AGGATCCNAATN

FQY.A	AGGTCAGCGATACGGCTGCAACTGTTACAGGA	<u>CAT</u>	GCCGATACTACGATTGCTTTAGACA	229
FliC	AGGTCAGCGATACGGCTGCAACTGTTACAGGATATGCCGATACTACGATTGCTTTAGACA			598

FQY.A	ATTTAAAA	<u>CA</u>	TGATGATACGACTGGAAAATATTACGCCAAAGTTACCGTTACGGGGGG	CA	349
FliC	ATTTAAAAATTTGATGATACGACTGGAAAATATTACGCCAAAGTTACCGTTACGGGGGG				718

FQY.A	AAAATGTACA	<u>C</u>	GTTGCAAATGCTGATTTGACAGAGGCTAAAGCCGCATTGACAGCAGCAG	529
FliC	AAAATGTACAAGTTGCAAATGCTGATTTGACAGAGGCTAAAGCCGCATTGACAGCAGCAG			898

(Mutation that is not underlined is an unintended change from GGA to GGG however this does not change the protein sequence)

FliC KNT mutations final DNA sequencing results and abbreviated alignment with wild-type FliC

K384H

CNCGGGCGNTCCGCTTCGGTGGACTCCTGCGACAGCAACTGAGGATGTGAAAAATGTACAAGTTGCAAATGCTGATTGA
 CAGAGGCTAAAGCCGATTGACAGCAGCAGGTGTACCGGCACAGCATCTGTTGTTAAGATGCTTATACTGATAATAAC
 GGTAAGAACTATTGATGGTGGTTTAGCAGTTAAGGTAGGCGATGATTACTATTCTGCAACTCAAAATAAGATGGTTCCAT
 AAGTATTAATACTACGAAATACACTGCAGATGACGGTACATCCAAAACGCACTAAACAACTGGGTGGCGCAGACGGCA
 AAACCGAAGTTGTTTCTATTGGTGGTAAACTTACGCTGCAAGTCACGCCGAAGGTCACAACCTTAAAGCACAGCCTGAT
 CTGGCGGAAGCGGCTGCTACAACCAACGAAAACCGCTGCAGAAAATTGATGCTGCTTTGGCACAGGTTGACACGTTACG
 TTCTGACCTGGGTGCGGTACAGAACCGTTTCAACTCCGCTATTACCAACCTGGGCAACACCGTAAACAACCTGACTTCTG
 CCCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCGCAGATTCTGCAGCAGGCCGGT
 ACCTCCGTTCTGGCGCAGGCGAACCAGGTTCCGCAAAACGTCCTCTCTTTACTGCGTTAAGGATCCGAATTCGAGCTCCG
 TCGACAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAAGCGGTCTG
 ATAAACAGAAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCCGTAG
 CGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGNAACTGCCAGGCATCAAATAAACGAAAGGCTCAGTCGAA
 AGACTGGGCCTTTCCGTTTTATCTGGTGTGGTGGTGAACGCTCTCTGAGTAGGACAAATCCGCCGGGANCCGGATTT
 GAACGTTGCGAACCACCGGCCCGGAAGGGNGCGGGCCAGGACNCCGCCNTNAACTGCCAGGCATCAANTTAAGCANAA
 GGCCATCCTGACGGATGGCCNTTTTNGCGTTTCTACAACTCTTTGTTNTTTTCTAAAANATNCAANATGNNTCCNC
 TCATGAAAAAAA

K384HPth890	TACGCTGCAAGT	CAC	GCCGAAGGTCACAACCTTAAAGCACAGCCTGATCTGGCGGAAGCG	412
FliC	TACGCTGCAAGTAAAGCCGAAGGTCACAACCTTAAAGCACAGCCTGATCTGGCGGAAGCG			1200
	*****	*	*****	

K384H + T116H/N120H

ATNGNCGGATACAATTTCCACAGGNAAACAGACCATGGGGGAATTCGAGCTCGGTACCCGGAGATCCTCTAGAAATAATT
 TTGTTTAACTTTAAGAAGGAGATATACATATGGCACAAGTCATTAATACAAACAGCCTGTCGCTGTTGACCCAGAATAAC
 CTGAACAAATCCCAGTCCGCTCTGGGCACCGCTATCGAGCGTCTGTCTTCCGGTCTGCGTATCAACAGCGCGAAAGACGA
 TCGCGCAGGTCAGGCGATTGCTAACCGTTTTACCGCGAACATCAAAGGTCTGACTCAGGCTTCCCGTAACGCTAACGACG
 GTATCTCCATTGCGCAGACCACTGAAGGCGCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTT
 CAGTCTGCTAACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCCACCAGCGCCTGCATGAAATCGA
 CCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCTGGCGCAGGACAACACCCTGACCATCCAGGTTGGTGCCA
 ACGACGGTGAAACTATCGATATCGATCTGAAGCAGATCAACTCTCAGACCCTGGGTCTGGATACGCTGAATGTGCAACAA
 AAATATAAGGTCAGCGATACGGCTGCAACTGTTACAGGATATGCCGATACTACGATTGCTTTAGACAATAGTACTTTTAA
 AGCCTCGGCTACTGGTCTTGGTGGTACTGACCAGAAAATTGATGGCGATTTAAAATTTGATGATACGACTGGAAAATATT
 ACGCCAAAGTTACCGTTACGGGGGGGACTGGTAAAGATGGCTATTATGAAGTTTCCGTTGATAAGACGAACGGTGAGGTG
 ACTCTTGCTGGCGGTGCGACTTCCCCGCTTACAGGTGGACTACCTGCGACAGCAACTGAGGATGTGAAAAAATGTACAAG
 TTGCAAATGCTGATTTGACAGAGGCTAAAGCCGCATTGACAGCAGCAGGGGTTACCGGCACAGCATCTGNTGTTAAGATG
 TCTTATACTGATAATAACGGTAAACTATTGATGGGGGTTAGCAGTTAAGGTAGGCGAATGATTACTATTCTGCAACTC
 AAAATAAAANANGGTTCCNTAAGNTNTTAATACTACGAAATACC

k+n/t.1	CTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATC	CACCAGCGCCTG	CAT	GAAATCGA	480
FliC	CTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATC	CACCCAGCGCCTG	AACGAAATCGA	371	

```

mflitrX      -----NTTANATCNNGCG----- 13
pflitrX      AACTTACTGGAATTACCCTTTCTACGGAAGCAGCCACTGATACTGGCGGAACCTAACCCAG 3360
thioredoxin  -----ATGTTACACCAACAACG----- 17
                                     :. :.*

mflitrX      -----AAATCACCGGGN 25
pflitrX      CTTCAATTGAGGGTGTTTATACTGATAATGGTAATGATTACTATGCGAAAATCACCGG-- 3418
thioredoxin  -----AAACCAACACGCC 30
                                     **. .**.*

mflitrX      NGGNNGATAACGATGGG-----TATGAGCGATAAAATTATTCACCTGACTGACGACAGT 79
pflitrX      -TGGTGATAACGATGG-----TATGAGCGATAAAATTATTCACCTGACTGACGACAGT 3470
thioredoxin  AGGCTTATTCTGTGGAGTTATATATGAGCGATAAAATTATTCACCTGACTGACGACAGT 90
          * . **:. * .*** *****

mflitrX      TTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTCGATTTCTGGGCAGAGTGG 139
pflitrX      TTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTCGATTTCTGGGCAGAGTGG 3530
thioredoxin  TTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTCGATTTCTGGGCAGAGTGG 150
          *****

mflitrX      TGCG----- 143
pflitrX      TGCGGTCCAGTGTGCTGGGCCCAGCCGCCAGATCTGAGCTCGCGGCCGCGATATCGCTA 3590

```

```

thioredoxin  TGCG----- 154
            ****

mflitr      ----GTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGC 199
pflitr      GCTCGACCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGC 3650
thioredoxin ----GTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGC 210
            *:*****

mflitr      AAAC TGACCGTTGCAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGC 259
pflitr      AAAC TGACCGTTGCAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAAGCTGGC 3710
thioredoxin AAAC TGACCGTTGCAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGC 270
            *****

mflitr      ATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTG 319
pflitr      ATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTG 3770
thioredoxin ATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTG 330
            *****

mflitr      GGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCCTCTGCCGCC 379
pflitr      GGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCCTCTGCCGCC 3830
thioredoxin GGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCGTAA----- 384
            *****

mflitr      AGTTCTCCAACCGCGGTCAAAC TGCGGAGATGATGGCAAAACAGAAGTGGTCGATATT 439
pflitr      AGTTCTCCAACCGCGGTCAAAC TGCGGAGATGATGGCAAAACAGAAGTGGTCGATATT 3890
thioredoxin -----

mflitr      GATGGTAAACATACGATTCTGCCGATTTAAATGGCGGTAATCTGCAAACAGGTTTGACT 499
pflitr      GATGGTAAACATACGATTCTGCCGATTTAAATGGCGGTAATCTGCAAACAGGTTTGACT 3950
thioredoxin -----

mflitr      GCTGGTGGTGAGGCTCTGACTGCTGTTGCAAATGGTAAAACACGGATCCGCTGAAAGCG 559
pflitr      GCTGGTGGTGAGGCTCTGACTGCTGTTGCAAATGGTAAAACACGGATCCGCTGAAAGCG 4010
thioredoxin -----

mflitr      CTGGACGATGCTATCGCATCTGTAGACAAATTCGGTTCTTCCCTCGGTGCGGTGCAAAAC 619
pflitr      CTGGACGATGCTATCGCATCTGTAGACAAATTCGGTTCTTCCCTCGGTGCGGTGCAAAAC 4070
thioredoxin -----

mflitr      CGTCTGGATTCCGCGGTTACCAACCTGAACAACACCCTACCAACCTGTCTGAAGCGCAG 679
pflitr      CGTCTGGATTCCGCGGTTACCAACCTGAACAACACCCTACCAACCTGTCTGAAGCGCAG 4130
thioredoxin -----

mflitr      TCCCGTATTCAGGACGCCGACTATGCGACCGAAGTGTCATATGTCGAAAGCGCAGATC 739
pflitr      TCCCGTATTCAGGACGCCGACTATGCGACCGAAGTGTCATATGTCGAAAGCGCAGATC 4190
thioredoxin -----

mflitr      ATCCAGCAGGCCGGTAAC TCCGTGTTGGCAAAAGCTAACCAGGTACCGCAGCAGGTTCTG 799
pflitr      ATCCAGCAGGCCGGTAAC TCCGTGTTGGCAAAAGCTAACCAGGTACCGCAGCAGGTTCTG 4250
thioredoxin -----

mflitr      TCTCTGCTGCAGGGTTAATCGTTGTAACCTGATTAAC TGAGACTGACGGCAACGCCNAAT 859
pflitr      TCTCTGCTGCAGGGTTAATCGTTGTAACCTGATTAAC TGAGACTGACGGCAACGCCAAAT 4310

```

```

thioredoxin  -----

mflitr      TGCTGATGCGCTGCGCTTATCAGGCCTACAAGTTGAATTGCAATTTATTGAATTTGCAC  919
pflitr      TGCTGATGCGCTGCGCTTATCAGGCCTACAAGTTGAATTGCAATTTATTGAATTTGCAC  4370
thioredoxin -----

mflitr      ATTTTGTAGGCCGGATAAGGCGTTTACGCCGCATCCGGCAACATAAAGCGCAATTTGTC  979
pflitr      ATTTTGTAGGCCGGATAAGGCGTTTACG-CGCATCCGGCAACATAAAGCGCAATTTGTC  4429
thioredoxin -----

mflitr      AGCAACGTGCTTCCCCGCCACCGGCGGGGTTTTTTTCTGCCTGGAATTTACCTGTAACCC  1039
pflitr      AGCAACGTGCTTCCCG-CCACCGGCGGGGTTTTTTTCTGCCTGGAATTTACCTGTAACCC  4488
thioredoxin -----

mflitr      CCAAT-AACCCCTCATTNC-CCCACTAATCGTCNGATN-AAAACCCTGCA-AAACGGATA  1095
pflitr      CCAAATAACCCCTCATTTCAACCACTAATCGTCCGATTAAAAACCCGTCAGAAACGGATA  4548
thioredoxin -----

mflitr      ATCATGCCGATAACTCATATANCGCAGGGCTGTTTATCGTGANTNCCCGGGGATCCTCTA  1155
pflitr      ATCATGCCGATAACTGCTATAACGCAGGGCTGTT-----TGAATTCCCGGGGATCCTCTA  4603
thioredoxin -----

mflitr      AAGTCNACCTGCAGNCATGCAAC-----  1178
pflitr      GAGTCGACCTGCAGGCATGCAAGCTTGCGACTGGCCGTCGTTTTACAACGTCGTGACTGG  4663
thioredoxin -----

```

pflitr is the original pFliTrx sequence from Ivitrogen.
 mflitr is the modified pFliTrx with the loop removed
 Thioredoxin is the wild-type from *E. coli*. GenBank: K02845.1

FliTrx loop removed with D26I mutation and abbreviated alignment with FliTrx loop removed

NGAGATC NNGCGAATCCCGGTGGTGATACGATGGTATGAGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACG
 GATGTACTCAAAGCGGACGGGGCGATCCTCGTCATTTTCTGGGCAGAGTGGTGCGGTCCGTGCAAATGATCGCCCCGAT
 TCTGGATGAAATCGCTGACGAATATCAGGGCAAACCTGACCGTTGCAAACCTGAACATCGATCAAACCCCTGGCACTGCGC
 CGAAATATGGCATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTG
 TCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCCTCTGCCGCCAGTTCTCAACCGCGGTCAAACCTGGGCGG
 AGATGATGGCAAACAGAAGTGGTCGATATTGATGGTAAACATACGATTCTGCCGATTAAATGGCGGTAATCTGCAA
 CAGGTTTGACTGCTGGTGGTGAGGCTCTGACTGCTGTTGCAAATGGTAAACACGGATCCGCTGAAAGCGCTGGACGAT
 GCTATCGCATCTGTAGACAAATCCGTTCTCCCTCGGTGCGGTGCAAACCGTCTGGATTCCGCGGTTACCAACCTGAA
 CAACACCACTACCAACCTGTCTGAAGCGCAGTCCCGTATTCAGGACGCCGACTATGCGACCGAAGTGCCAATATGTCGA
 AAGCGCAGATCATCCAGCAGGCCGGTAACTCCGTGTTGGCAAAGCTAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTG
 CAGGGTTAATCGTTGTAACCTGATTAACCTGAGACTGACGGCAACGCCAAATTGCCTGATGCGCTGCGCTTATCAGGCCTA
 CAAGTTGAATTGCAATTTATTGAATTTGCACATTTTTGTAGGCCGGAATAAAGGCGTTTTACGCCGCATCCGGCAACAT
 AAAGCGCAANTTTGTCAGCANC

D26Iftx2	CGGATGTACTCAAAGCGGACGGGGCGATCCTCGTCA	TTTTCTGGGCAGAGTGGTGCGGTC	140
FliTrx	CGGATGTACTCAAAGCGGACGGGGCGATCCTCGTCGATT	TTTTCTGGGCAGAGTGGTGCGGTC	178
	*****	*****	

Appendix B

Sequencing Primers

FliC sequencing primers

Table A.1: FliC sequencing primers

<u>Primer name</u>	<u>Primer sequence</u>	<u>Used to sequence</u>
pTH890_MCS_For	CAATTAATCATCCGGCTCGT	N120H/T116H
FliC_For_1	TCCATCCAGGCTGAAATCA	Y190H, F222H, Q282H,
FliC_For_2	TAAGACGAACGGTGAGGTGA	K384H, L396H, Q393H

^aPrimers were synthesized by IDT

FliTrx sequencing primer

Table A.2: FliTrx sequencing primer

<u>Primer name</u>	<u>Primer sequence</u>	<u>Used to sequence</u>
FliTrx_Forward	GCTTCAATTGAGGGTGTTTATACT ^a	All FliTrx mutations

^aPrimer is different than the one provided by Invitrogen with pFliTrx™ Peptide Display Vector kit

^bPrimer synthesized by Invitrogen

Appendix C

ICP-ES Protein-Metal Analysis Data

	5/16/12	ICP		ppm			
#	Element	<u>FQ1M</u>	<u>KNTM</u>	<u>FQP</u>	<u>KNTP</u>	<u>QLP</u>	<u>QLM</u>
1	Al3082	0.0148	0.0331	0.0057	0.0297	0.0068	0.0228
2	B_2496	0.0074	0.0141	0.0097	0.0074	0.0041	0.0052
3	Ba4934	0.0007	0.0011	0.0006	0.0012	0.0004	0.0007
4	Ca3158	0.3078	0.3127	0.684	0.805	1.043	0.6352
5	Cd2288	0.0038	0.0038	0.0038	0.0152	0.0038	0.007
6	Co2286	0.0004	0.0027	0.0016	0.0019	0.0019	0.0012
7	Cr2677	0.0056	0.0034	0.0001	0.0079	-0.0003	0.0011
8	Cu3247	0.0148	0.0072	0.0097	0.0542	-0.0157	0.044
9	Fe2599	0.003	0.0074	0.0074	0.013	0.0041	0.0074
10	K_7664	580.6	574.8	564.2	572.1	571.5	572.3
11	Mg2790	-0.0293	-0.0444	0.0411	0.0335	0.0461	0.0461
12	Mn2576	0.0008	0.0008	0.0004	0.0017	0	0.0017
13	Mo2020	0.0326	0.0143	0.0075	0.008	0.0075	0.0143
14	Na5889	3.461	11.23	5.851	6.121	6.415	4.684
15	Ni2316	0.0095	0.0073	0.0073	0.0057	0.0041	0.0052
16	P_2149	0.7242	0.8499	0.4755	0.91	0.444	0.6727
17	Pb2203	0.027	0.0443	0.0347	0.0482	0.0405	0.0424
18	Si2881	0.0753	0.0776	0.0593	0.0753	-0.0046	0.0616
19	Sr4215	0.0012	0.0012	0.0028	0.0027	0.0036	0.0023
20	Zn2139	0.7483	1.255	0.7171	1.287	1.324	2.664

#	Element	<u>WTM</u>	<u>WTP</u>	<u>CA</u>	<u>B0</u>	<u>B05</u>	<u>B10</u>
1	Al3082	-0.0148	-0.0068	-0.0183	0.0034	0.0125	0.0103
2	B_2496	-0.0004	0.0097	0.0019	0.0007	-0.0004	0.003
3	Ba4934	-0.0002	0.0001	-0.0005	-0.0001	0.0007	0.0004
4	Ca3158	0.6117	1.034	0.8855	0.4077	0.6563	0.5481
5	Cd2288	0.0038	0.007	0.0087	0.007	0.0038	0.0087
6	Co2286	-0.0004	0.0008	-0.0016	0.0012	0.0023	0.0019
7	Cr2677	-0.0054	-0.0031	-0.0074	-0.0028	-0.0005	-0.0021
8	Cu3247	0.0034	0.0021	0.0605	-0.0207	0.0097	0.0326
9	Fe2599	0.0041	0.0007	-0.0015	-0.0015	0.003	0.0007
10	K_7664	565.2	572.1	573.5	573.1	557.8	552.4
11	Mg2790	0.0059	0.0712	0.0436	-0.062	-0.0042	0.026
12	Mn2576	0.0008	0.0004	0	0	0.0008	0.0004
13	Mo2020	0.0054	0.0044	0.0033	0.0038	0.0106	0.0143
14	Na5889	3.044	3.931	6.197	1.869	3.747	5.831
15	Ni2316	0.003	0.0073	0.0014	0.0052	0.0041	0.0052
16	P_2149	0.8556	0.6098	0.8099	0.5098	0.5584	0.8271
17	Pb2203	0.0347	0.0289	0.0096	0.0328	0.0347	0.0366
18	Si2881	-0.0639	-0.0799	-0.1073	-0.0411	0.0616	-0.0251
19	Sr4215	0.002	0.0034	0.0022	0.0013	0.0017	0.0014
20	Zn2139	0.9894	1.016	1.115	0.5682	1.758	2.827

Appendix D**Recombinant DNA Approval****WESTERN MICHIGAN UNIVERSITY**

Recombinant DNA Biosafety Committee

Recombinant DNA Biosafety Committee**Project Approval Certification****For rDNA Biosafety Committee Use Only**

Project Title: Engineering and Display of Enzymes and Proteins on Bacterial Flagella Fibers

Principal Investigator: Brian Tripp

IBC Project Number: 12-BTa

Date Received by the rDNA Biosafety Committee: October 31, 2011

☒ Reviewed by the rDNA Biosafety Committee

☒ Approved

☐ Approval not required

Kim Esau

Vice Chair of rDNA Biosafety Committee Signature

12/01/2011
Date

Revised 11/2011 WMU RDBC
All other copies obsolete

Walwood Hall, Kalamazoo MI 49008-5456
PHONE: (616) 387-8293 FAX: (616) 387-8276

BIBLIOGRAPHY

- 1 Brannigan, J. A., and Wilkinson, A. J. Protein engineering 20 years on. *Nature Reviews Molecular Cell Biology* **3**, 964-970 (2002).
- 2 Liese, A., and Villela Filho, M. Production of fine chemicals using biocatalysis. *Current Opinion in Biotechnology* **10**, 595-603 (1999).
- 3 Benner, S. A., and Sismour, A. M. Synthetic biology. *Nature Reviews Genetics* **6**, 533-543 (2005).
- 4 Macnab, R. M. How bacteria assemble flagella. *Annual Review of Microbiology* **57**, 77-100 (2003).
- 5 Malapaka, R. R. V., Adebayo, L. O., and Tripp, B. C. A deletion variant study of the functional role of the *Salmonella* flagellin hypervariable domain region in motility. *Journal of Molecular Biology* **365**, 1102-1116 (2007).
- 6 Yonekura, K., and Maki-Yonekura, S. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**, 643-650 (2003).
- 7 Lu, Y., Berry, S. M., and Pfister, T. D. Engineering novel metalloproteins: Design of metal-binding sites into native protein scaffolds. *Chemical Reviews* **101**, 3047-3080 (2001).
- 8 Orengo, C. A., Michie, A. D., Jones, D. T., Swindells, M. B., and Thornton, J. M. CATH: A hierarchic classification of protein domain structures. *Structure* **5**, 1093-1108 (1997).

- 9 Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *Journal of Molecular Biology* **247**, 536-540 (1995).
- 10 Kuchner, O., and Arnold, F. H. Directed evolution of enzyme catalysts. *Trends in Biotechnology* **15**, 523-530 (1997).
- 11 Chen, R. Enzyme engineering: rational redesign versus directed evolution. *Trends in Biotechnology* **19**, 13-14 (2001).
- 12 Tainer, J. A. J., and Roberts, V. A. V. Enzyme motifs in antibodies. *Nature* **348**, 589-589 (1990).
- 13 Wade, W. S., Koh, J. S., Han, N., Hoekstra, D. M., and Lerner, R. A. Engineering metal coordination sites into the antibody light chain. *Journal of the American Chemical Society* **115**, 4449-4456 (1993).
- 14 McGrath, M. E., Haymore, B. L., Summers, N. L., Craik, C. S., and Fletterick, R. J. Structure of an engineered, metal-actuated switch in trypsin. *Biochemistry* **32**, 1914-1919 (1993).
- 15 Willett, W. S., Gillmor, S. A., Perona, J. J., Fletterick, R. J., and Craik, C. S. Engineered metal regulation of trypsin specificity. *Biochemistry* **34**, 2172-2180 (1995).
- 16 Lu, Z., DiBlasio-Smith, E. A., Grant, K. L., Warne, N. W., LaVallie, E. R., Collins-Racie, L. A., Follettie, M. T., Williamson M. J., and McCoy, J.M. Histidine patch thioredoxins. *Journal of Biological Chemistry* **271**, 5059-5065 (1996).

- 17 Vita, C., Roumestand, C., Toma, F., and Ménez, A. Scorpion toxins as natural scaffolds for protein engineering. *Proceedings of the National Academy of Sciences USA* **92**, 6404–6408 (1995).
- 18 Pierret, B., Virelizier, H., and Vita, C. Synthesis of a metal-binding protein designed on the alpha/beta scaffold of charybdotoxin. *International Journal of Peptide Protein Research*. **46**, 471-479 (1995).
- 19 Mueller, H. N., and Skerra, A. Grafting of a high-affinity Zn(II)-binding site on the beta-barrel of retinol-binding protein results in enhanced folding stability and enables simplified purification. *Biochemistry* **33**, 14126-14135 (1994).
- 20 Brinen, L. S., Willett, W. S., Craik, C. S., and Fletterick, R. J. X-ray structures of a designed binding site in trypsin show metal-dependent geometry. *Biochemistry* **35**, 5999-6009 (1996).
- 21 Hughes, K. T., and Aldridge, P. D. Putting a lid on it. *Nature Structural and Molecular Biology* **8**, 96-97 (2001).
- 22 Yonekura, K., Maki, S., Morgan, D. G., DeRosier, D. J., Vonderviszt, F., Imada, K., Namba, K. The bacterial flagellar cap as the rotary promoter of flagellin self-assembly. *Science* **290**, 2148-2152 (2000).
- 23 Samatey, F. A., Imada, K., Nagashima, S., and Vonderviszt, F. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. *Nature* **410**, 331-337 (2001).

- 24 Maki-Yonekura, S., Namba, K., and Yonekura, K. Conformational change of flagellin for polymorphic supercoiling of the flagellar filament. *Nature Structural and Molecular Biology* **17**, 417-423 (2010).
- 25 Namba, K., Yamashita, I., and Vonderviszt, F. Structure of the core and central channel of bacterial flagella. *Nature* **342**, 648-654 (1989).
- 26 Iino, T. Polarity of flagellar growth in *Salmonella*. *Microbiology* **56**, 227-239 (1969).
- 27 Beatson, S. A., Minamino, T., and Pallen, M. J. Variation in bacterial flagellins: From sequence to structure. *Trends in Microbiology* **14**, 151-155 (2006).
- 28 Kuwajima, G. Construction of a minimum-size functional flagellin of *Escherichia coli*. *Journal of Bacteriology* **170**, 485-488 (1988).
- 29 Homma, M., Fujita, H., Yamaguchi, S., and Iino, T. Regions of *Salmonella typhimurium* flagellin essential for its polymerization and excretion. *Journal of Bacteriology* **169**, 291-296 (1987).
- 30 Smith, K. D., Andersen-Nissen, E., Hayashi, F., Strobe, K., Bergman, M. A., Rassouljian Barrett, S. L., Cookson, B. T., and Aderem, A. Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. *Nature Immunology* **4**, 1247-1253 (2003).
- 31 Silverman, M., and Simon, M. Flagellar rotation and the mechanism of bacterial motility. *Nature* **249**, 73-74 (1974).
- 32 Berg, H. C., and Anderson, R. A. Bacteria swim by rotating their flagellar filaments. *Nature* **245**, 380-382 (1973).

- 33 Yonekura, K., Maki-Yonekura, S., and Namba, K. Growth mechanism of the bacterial flagellar filament. *Research in Microbiology* **153**, 191-197 (2002).
- 34 Macnab, R. M. Bacterial flagella rotating in bundles: A study in helical geometry. *Proceedings of the National Academy of Sciences USA* **74**, 221-225 (1977).
- 35 Turner, L., Ryu, W. S., and Berg, H. C. Real-time imaging of fluorescent flagellar filaments. *Journal of Bacteriology* **182**, 2793–2801 (2000).
- 36 Flores, H., Lobaton, E., Stefan Méndez-Diez, Tlupova, S., and Cortez, R. A study of bacterial flagellar bundling. *Journal of Bacteriology* **67**, 137-168 (2005).
- 37 Berg, H. C. The rotary motor of the bacterial flagella. *Annual Review of Biochemistry* **72**, 19-54 (2003).
- 38 Matsuura, S., Kamiya, R., and Asakura, S. Transformation of straight flagella and recovery of motility in a mutant *Escherichia coli*. *Journal of Molecular Biology* **118**, 431-440 (1978).
- 39 Kamiya, R., and Asakura, S. Flagellar transformations at alkaline pH. *Journal of Molecular Biology* **108**, 513-518 (1976).
- 40 Kamiya, R., and Asakura, S. Helical transformations of *Salmonella* flagella *in vitro*. *Journal of Molecular Biology* **106**, 167-186 (1976).
- 41 Thompson, A. J., and Gray, H. B. Bio-inorganic chemistry. *Current Opinion in Chemical Biology* **2**, 155-158 (1998).
- 42 Dudev, T., and Lim, C. Metal binding affinity and selectivity in metalloproteins: Insights from computational studies. *Annual Review of Biophysics* **37**, 97-116 (2008).

- 43 Benson, D. E., Wisz, M. S., and Hellinga, H. W. Rational design of nascent metalloenzymes. *Proceedings of the National Academy of Sciences USA* **97**, 6292– 6297 (2000).
- 44 Lu, Y. Metalloprotein and metallo-DNA/RNAzyme design: Current approaches, success measures, and future challenges. *Inorganic Chemistry* **45**, 9930-9940 (2006).
- 45 McCall, K. A., Chih-chin, H., and Fierke, C. A. Function and mechanism of zinc metalloenzymes. *The Journal of Nutrition* **130**, S1437-1446S (2000).
- 46 Gregory, D. S., Martin, A. C. R., Cheetham, J. C., and Rees, A. R. The predictions and characterizaton of metal-binding sites in proteins. *Protein Engineering*. **6**, 29-35, (1993).
- 47 Tu, C., Tripp, B. C., Ferry, J. G., and Silverman, D. N. Bicarbonate as a proton donor in catalysis by Zn(II)- and Co(II)-containing carbonic anhydrases. *Journal of the American Chemical Society* **123**, 5861-5866 (2001).
- 48 Williams, R. J. P. The biochemistry of zinc. *Polyhedron* **6**, 61-69 (1987).
- 49 Cheang, U. K., Roy, D., Lee, J. H., and Kim, M. J. Fabrication and magnetic control of bacteria-inspired robotic microswimmers. *Applied Physics Letters* **97**, 3 (2010).
- 50 Diekmann, S., [Weston](#), J., [Anders](#), E., [Boland](#), W., Schönecker, B., [Hettmann](#), T., [von Langen](#), J., [Erhardt](#), S., [Mauksch](#), M., [Bräuer](#), M., [Beckmann](#), C., [Rost](#), M., [Sperling](#), P., and [Heinz](#), E. Metal-mediated reactions modeled after nature. *Reviews in Molecular Biotechnology* **90**, 73-94 (2002).

- 51 Kumara, M. T., Muralidharan, S., and Tripp, B. C. Generation and characterization of inorganic and organic nanotubes on bioengineered flagella of mesophilic bacteria. *Journal of Nanoscience and Nanotechnology* **7**, 2260-2272 (2007).
- 52 Hesse, W. R., Luo, L., Zhang, G., Mulero, R., Cho, J., and Kim, M. J. Mineralization of flagella for nanotube formation. *Materials Science and Engineering C-Materials for Biological Applications* **29**, 2282-2286 (2009).
- 53 Deplanche, K., Woods, R. D., Mikheenko, I. P., Sockett, R. E., and Macaskie, L. E. Manufacture of stable palladium and gold nanopartilces on native and genetically engineered flagella scaffolds. *Biotechnology and Bioengineering* **101**, 873-880 (2008).
- 54 Kumara, M. T., Tripp, B. C., and Muralidharan, S. Self-assembly of metal nanoparticles and nanotubes on bioengineered flagella scaffolds. *Chemistry of Materials* **19**, 2056-2064 (2007).
- 55 Westerlund-Wikström, B., Tanskanen, J., Virkola, R., Hacker, J., Lindberg, M., Skurnik, M., and Korhonen, T. K. Functional expression of adhesive peptides as fusions to *Escherichia coli* flagellin. *Protein Eng.* **10**, 1319-1326 (1997).
- 56 Tanskanen, J., Korhonen, T. K., and Westerlund-Wikström, B. Construction of a multihybrid display system: Flagellar filaments carrying two foreign adhesive peptides. *Applied and Environmental Microbiology* **66**, 4152-4156 (2000).
- 57 He, X. S., Rivkina, M., Stocker, B. A., and Robinson, W. S. Hypervariable region IV of *Salmonella* gene fliC_d encodes a dominant surface epitope and a stabilizing factor for functional flagella. *Journal of Bacteriology* **176**, 2406-2414 (1994).

- 58 Wei, H.-J., Chang, W., Lin, S.-C., Liu, W.-C., Chang, D.-K., Chong, P., and Wu, S.-C. Fabrication of influenza virus-like particles using M2 fusion proteins for imaging single viruses and designing vaccines. *Vaccine* **29**, 7163-7172 (2011).
- 59 Westerlund-Wikström, B. Peptide display on bacterial flagella: principles and applications. *International Journal of Medical Microbiology* **290**, 223-230 (2000).
- 60 Luke, J. M., Carnes, A. E., Sun, P., Hodgson, C. P., Waugh, D. S., and Williams, J. A. Thermostable tag (TST) protein expression system: Engineering thermotolerant recombinant proteins and vaccines. *Journal of Biotechnology* **151**, 242-250 (2010).
- 61 Song, L., Zhang, Y., Yun, N. E., Poussard, A. L., Smith, J. N., Smith, J. K., Borisevich, V., Linde, J. J., Zacks, M. A., Li, H., Kavita, U., Reiserova, L., Liu, X., Dumuren, K., Balasubramanian, B., Weaver, B., Parent, J., Umlauf, S., Liu, G., Huleatt, J., Tussey, L., and Paessler, S. Superior efficacy of a recombinant flagellin:H5N1 HA globular head vaccine is determined by the placement of the globular head within flagellin. *Vaccine* **27**, 5875-5884 (2009).
- 62 Miao, E., Andersen-Nissen, E., Warren, S., and Aderem, A. TLR5 and Ipaf: Dual sensors of bacterial flagellin in the innate immune system. *Seminars in Immunopathology* **29**, 275-288 (2007).
- 63 Lu, Z. J., Murray, K. S., Van Cleave, V., LaVallie, E. R., Stahl, M. L., and McCoy, J. M. Expression of thioredoxin random peptide libraries on the *Escherichia coli* cell-surface as functional fusions to flagellin - A system designed for exploring protein-protein interactions. *Bio-Technology* **13**, 366-372 (1995).

- 64 Klein, Á. G., Tóth, B. Z., Jankovics, H., Muskotál, A. I., and Vonderviszt, F. A polymerizable GFP variant. *Protein Engineering Design and Selection* **25**, 153-157 (2012).
- 65 Szabó, V., Muskotál, A. I., Tóth, B. z., Mihovilovic, M. D., and Vonderviszt, F. Construction of a xylanase variant capable of polymerization. *PLoS ONE* **6**, e25388 (2011).
- 66 Tripp, B. C., Lu, Z., Bourque, K., Sookdeo, H., and McCoy, J. M. Investigation of the 'switch-epitope' concept with random peptide libraries displayed as thioredoxin loop fusions. *Protein Engineering* **14** (2001).
- 67 Dong, J., Liu, C., Zhang, J., Xin, Z.-T., Yang, G., Gao, B., Mao, C.-Q., Liu, N.-L., Wang, F., Shao, N.-S., Fan, M., and Xue, Y.-N. Selection of novel nickel-binding peptides from flagella displayed secondary peptide library. *Chemical Biology and Drug Design* **68**, 107-112 (2006).
- 68 Woods, R. D., Takahashi, N., Aslam, A., Pleass, R. J., Aizawa, S.-I., and Sockett, R. E. Bifunctional nanotube scaffolds for diverse ligands are purified simply from *Escherichia coli* strains coexpressing two functionalized flagellar genes. *Nano Letters* **7**, 1809-1816 (2007).
- 69 Kumara, M., Srividya, N., Muralidharan, S., and Tripp, B. Bioengineered flagella protein nanotubes with cysteine loops: Self-assembly and manipulation in an optical trap. *Nano Letters* **6**, 2121-2129 (2006).

- 70 Vonderviszt, F., Kanto, S., Aizawa, S., and Namba, K. Terminal regions fo the flagellin are disordered in solution. *Journal of Molecular Biology* **209**, 127-133 (1989).
- 71 Muskotál, A., Seregélyes, C., Sebestyén, A., and Vonderviszt, F. Structural basis for stabilization of the hypervariable D3 domain of *Salmonella* flagellin upon filament formation. *Journal of Molecular Biology* **403**, 607-615 (2010).
- 72 Ezaki, S., Tsukio, M., Takagi, M., and Imanaka, T. Display of heterologous gene products on the *Escherichia coli* cell surface as fusion proteins with flagellin. *Journal of Fermentation and Bioengineering* **86**, 500-503 (1998).
- 73 Smith, K. S., Jakubzick, C., Whittam, T. S., and Ferry, J. G. Carbonic anhydrase is an ancient enzyme widespread in prokaryotes. *Proceedings of the National Academy of Sciences USA* **96**, 15184–15189 (1999).
- 74 Pastorekova, S., Parkkila, S., Pastorek, J., and Supuran, C. T. Carbonic anhydrases: Current state of the art, therapeutic applications and future prospects. *Journal of Enzyme Inhibition and Medicinal Chemistry* **19**, 199-229 (2004).
- 75 Maren, T. H., and Conroy, C. W. A new class of carbonic anhydrase inhibitor. *Journal of Biological Chemistry* **268**, 26233-26239 (1993).
- 76 Lindskog, S. Structure and mechanism of carbonic anhydrase. *Pharmacology and Therapeutics* **74**, 1-20 (1997).
- 77 Tashian, R. E. Genetic variation and evolution of the carboxylic esterases and carbonic anhydrases of primate erythrocytes. *American Journal of Human Genetics* **17**, 257–272 (1965).

- 78 Bolon, D. N., and Mayo, S. L. Enzyme-like proteins by computational design. *Proceedings of the National Academy of Sciences USA* **98**, 14274-14279 (2001).
- 79 Pocker, Y., and Stone, J. T. The catalytic versatility of erythrocyte carbonic anhydrase III kinetic studies of the enzyme-catalyzed hydrolysis of p-nitrophenyl acetate. *Biochemistry* **6**, 668-678 (1967).
- 80 Auld, D. S. Zinc coordination sphere in biochemical zinc sites. *BioMetals* **14**, 271-313 (2001).
- 81 Steiner, H., Jonsson, B.-H., and Lindskog, S. The catalytic mechanism of carbonic anhydrase hydrogen-isotope effects on the kinetic parameters of the human C isoenzyme. *European Journal of Biochemistry* **59**, 253-259 (1975).
- 82 Tu, C., Silverman, D. N., Forsman, C., Jonsson, B. H., and Lindskog, S. Role of histidine 64 in the catalytic mechanism of human carbonic anhydrase II studied with a site-specific mutant. *Biochemistry* **28**, 7913-7918 (1989).
- 83 Alexander, R. S., Nair, S. K., and Christianson, D. W. Engineering the hydrophobic pocket of carbonic anhydrase II. *Biochemistry* **30**, 11064-11072 (1991).
- 84 Höst, G., Mårtensson, L.-G. R., and Jonsson, B.-H. Redesign of human carbonic anhydrase II for increased esterase activity and specificity towards esters with long acyl chains. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1764**, 1601-1606 (2006).
- 85 Okrasa, K., and Kazlauskas, R. J. Manganese-substituted carbonic anhydrase as a new peroxidase. *Chemistry – A European Journal* **12**, 1587-1596 (2006).

- 86 Drakopoulou, E., ZinnJustin, S., Guenneugues, M., Gilquin, B., Menez, A., and Vita, C. Changing the structural context of a functional beta-hairpin - Synthesis and characterization of a chimera containing the curaremimetic loop of a snake toxin in the scorpion alpha/beta scaffold. *Journal of Biological Chemistry* **271**, 11979-11987 (1996).
- 87 Sheng, X., Ting, Y. P., and Pehkonen, S. O. The influence of ionic strength, nutrients and pH on bacterial adhesion to metals. *Journal of Colloid and Interface Science* **321**, 256-264 (2008).
- 88 West, M., Burdash, N. M., and Freimuth, F. Simplified silver-plating stain for flagella. *Journal of Clinical Microbiology* **6**, 414-419 (1977).
- 89 Guzzo, A., Diorio, C., and DuBow, M. S. Transcription of the *Escherichia coli* fliC gene is regulated by metal ions. *Applied and Environmental Microbiology* **57**, 2255-2259 (1991).
- 90 Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Research* **34**, W116-W118 (2006).
- 91 Petsko, G. A., and Ringe, D. *Protein Structure and Function*. (New Science Press Ltd., 2004).
- 92 Tsai, S. P., Hartin, R. J., and Ryu, J. Transfromation in restriction-deficient *Salmonella-typhimurium* LT2. *Journal of General Microbiology* **135**, 2561-2567 (1989).

- 93 Iyer, R., Barrese, A. A., Parakh, S., Parker, C. N. and Tripp, B. C. Inhibition profiling of human carbonic anhydrase II by high-throughput screening of structurally diverse, biologically active compounds. *Journal of Biomolecular Screening* **11**, 782-791 (2006).
- 94 Iino, T., and Mitani, M. A mutant of *Salmonella* possessing straight flagella. *Journal of General Microbiology* **49**, 81-88 (1967).
- 95 Yoshioka, K., Aizawa, S.-I., and Yamaguchi, S. Flagellar filament structure and cell motility of *Salmonella typhimurium* mutants lacking part of the outer domain. *Journal of Bacteriology* **177**, 1090-1093 (1995).
- 96 Vonderviszt, F., Sonoyama, M., Tasumi, M., and Namba, K. Conformational adaptability of the terminal regions of flagellin. *Biophysics Journal* **63**, 1672-1677 (1992).
- 97 Uratani, Y., Asakura, S., and Imahori, K. A circular dichroism study of *Salmonella* flagellin: Evidence for conformational change on polymerization. *Journal of Molecular Biology* **67**, 85-98 (1972).
- 98 Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* **1**, 2876-2890 (2006).
- 99 Perez-Iratxeta, C., and Andrade-Navarro, M. K2D2: Estimation of protein secondary structure from circular dichroism spectra. *BioMed Central Structural Biology* **8**, 25 (2008).

- 100 Louis-Jeune, C., Andrade-Navarro, M. A., and Perez-Iratxeta, C. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins: Structure, Function, and Bioinformatics* **80**, 374-381 (2011).
- 101 Lloyd, M. D., Thiyagarajan, N., Ho, Y. T., Woo, L. W. L., Sutcliffe, O. B., Purohit, A., Reed, M. J., Acharya, K. R., and Potter, B. V. L. First crystal structures of human carbonic anhydrase II in complex with dual aromatase-steroid sulfatase inhibitors. *Biochemistry* **44**, 6858-6866 (2005).
- 102 Lloyd, M. D., Pederick, R. L., Natesh, R., Woo, R. W. L., Purohit, A., Reed, M. J., Acharya, K. R., and Potter, N. L. Crystal structure of human carbonic anhydrase II at 1.95 angstrom resolution in complex with 667-cournate, a novel anti-cancer agent. *Biochemical Journal* **385**, 715-720 (2005).
- 103 Cozier, G. E., Leese, M. P., Lloyd, M. D., Baker, M. D., Thiyagarajan, N., Acharya, K. R., and Potter B. V. L. Structures of human carbonic anhydrase II/Inhibitor complexes reveal a second binding site for steroidal and nonsteroidal inhibitors. *Biochemistry* **49** (2010).
- 104 Dahiyat, B. I., and Mayo, S. L. *De novo* protein design: fully automated sequence selection. *Science* **278**, 82-87 (1997).
- 105 Macnab, R. M. Examination of bacterial flagellation by dark-field microscopy. *Journal of Clinical Microbiology* **4**, 258-265 (1976).
- 106 Grossart, H.-P., Steward, G. F., Martinez, J., and Azam, F. A simple, rapid method for demonstrating bacterial flagella. *Applied and Environmental Microbiology* **66**, 3632-3636 (2000).

- 107 Wilcox, D. E. Isothermal titration calorimetry of metal ions binding to proteins: An overview of recent studies. *Inorganica Chimica Acta* **361**, 857-867 (2008).
- 108 Zidane, F., Matéos, A., Cakir-Kiefer, C., Miclo, L., Rahuel-Clermont, S., Girardet, J.-M., and Corbier, C. Binding of divalent metal ions to 1-25 β -caseinophosphopeptide: An isothermal titration calorimetry study. *Food Chemistry* **132**, 391-398.
- 109 Thompson, R. B., Peterson, D., Mahoney, W., Cramer, M., Maliwal, B. P., Suh, S. W., Frederickson, C., Fierke, C., and Herman, P. Fluorescent zinc indicators for neurobiology. *Journal of Neuroscience Methods* **118**, 63-75 (2002).
- 110 Alber, B. E., Colangelo, C. M., Dong, J., Stalhandske, C. M. V., Baird, T. T., Tu, C. K., Fierke, C. A., Silverman, D. N., Scott, R. A., and Ferry, J. G. Kinetic and spectroscopic characterization of the gamma-carbonic anhydrase from the methanoarchaeon *Methanosarcina thermophila*. *Biochemistry* **38**, 13119-13128 (1999).