# Structural Investigation of FtsK and Close Homologs

Theodore Klimek, Hideki Aihara, John Lee

Submitted to the University of Minnesota Digital Conservancy: Sunday, May 17, 2012

FtsK is an integral membrane DNA translocase of the Fts family found within *Escherichia coli* and with close homologs across many bacterial species. Some of those homologs (Tra family) are also involved in a variety of type IV secretion systems, commonly involved in bacterial conjugation and often in infection. Our work focused on the cloning, expression, purification, and crystallization of FtsK and several homologs, TrwB, TraD, and TraG, for the purpose solving the atomic resolution structure of these proteins. We prepared a variety of constructs in bacterial T7 based expression vectors for the purpose of expressing these proteins in *E. coli.* For the proteins that showed positive expression, the protein was solubilized with detergent and then purified through nickel-NTA affinity and size-exclusion chromatography (SEC). Results are shown through SEC reports and SDS-PAGE gels that indicate protein purity (by gel) and homogeneity and oligomeric state (by SEC). Many bacterial membrane based proteins, especially those of FtsK and its relatives are accesible targets for antibiotic intervention in humans as they are unique to bacterial systems. For this reason we hope that the procedures followed here may enable further development of our knowledge of these systems, that we might maintain a leg up on rapidly evolving pathogenic species.

## I. INTRODUCTION

FtsK is an essential DNA transporter found in the gram-negative bacteria such as *Escherichia coli*. Encoded by the *ftsK* gene, FtsK is a 147 kDa protein with a high level of homology with other DNA transport proteins, such as transfer (Tra) proteins and TrwB found in *E. coli* and other gram negative bacteria. The Fts family of proteins is involved in cell invagination and septum formation during cell division (Dorazi and Dewar 2000). The details of FtsK function at the septum are not clear, however, it is known that the membrane domain is embedded in the inner plasma membrane, probably co-localized with other divisosome proteins. The C-terminal cytoplasmic domain ($FtsK_C$) readily oligomerizes as a homo-hexamer and has been shown via biochemical and structural studies to accommodate double stranded DNA (Strick and Quessada-Vial 2006; Massey et al. 2006). These two domains are connected by a linking domain $FtsK_L$. The cytoplasmic domain is composed of three sub-domains, α, β, and γ. The α and β domains provide an ATP-powered motor function and the γ domain recognizes FtsK-orienting polar sequences (KOPS) of daughter chromosomes to align their termini and assign one to each daughter cell. In essence FtsK is believed

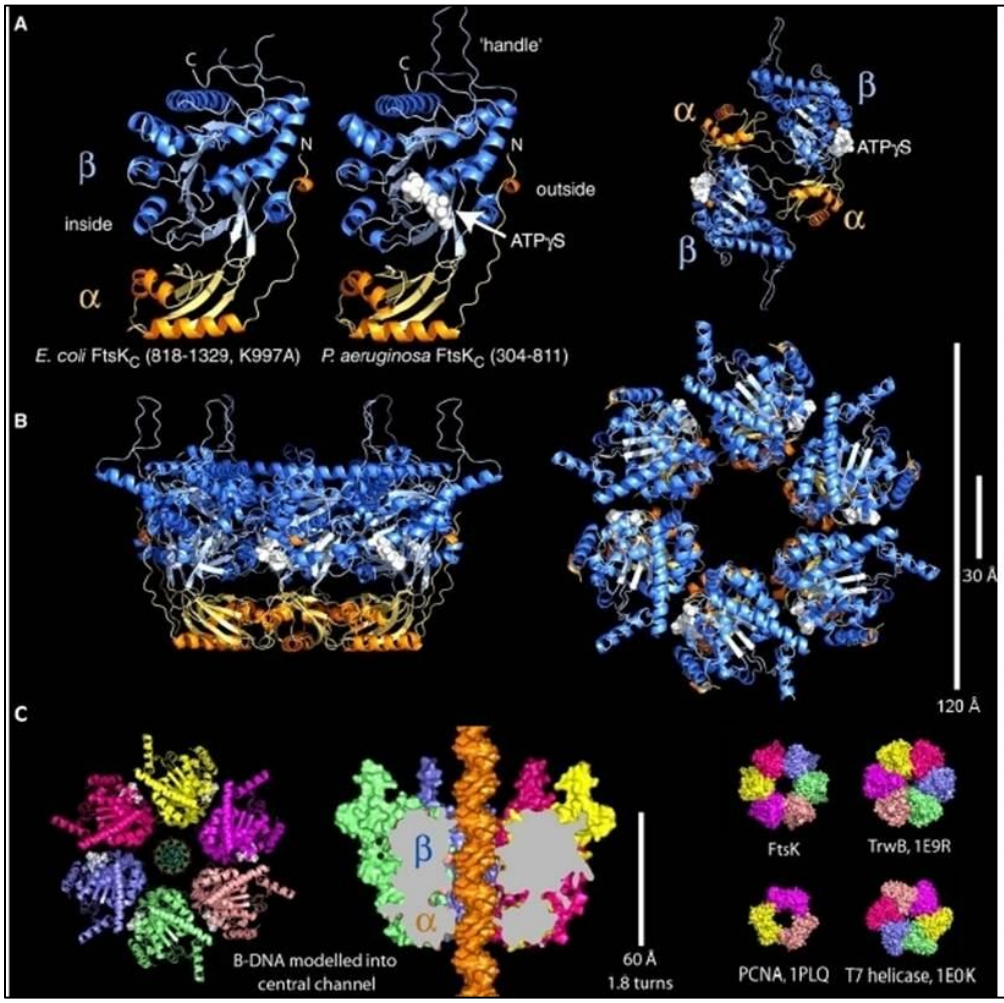


**Figure 1. Structures of $FtsK_C$**

**(A)** Crystal structures of $FtsK_C$ from *E. coli* (left) and *Pseudomonas aeruginosa* (middle). According to the hexameric structure (B) the center of the channel is to the left of the structures shown. Each shows two domains, α and β. Image on the right shows a common head-tail dimer that occurred in crystals.

**(B)** Crystal structure of hexameric $FtsK_C$ viewed from the side (left) and top (right). Hexamere is in form of a pore-like cylindrical structure with an inner channel of ~30 Å and outer diameter of ~120 Å.

**(C)** Central channel can accommodate dsDNA as modeled from top (left) and from cut-away side view (DNA orange) (middle). For comparison similar protein channels, TrwB, T7, and PCNA are shown alongside FtsK. Only PCNA and FtsK have a wide enough channels to accommodate dsDNA.

Image adapted from (Strick and Quessada-Vial 2006)

to be responsible for last-minute chromosome shuffling between the two daughter cells to assure the proper division of genetic material. FtsK has been shown to translocate DNA at a rate of up to 7 kbp per second, making it the fastest known molecular motor (Strick and Quessada-Vial 2006).

While FtsK is known to be an integral membrane protein, it is not known whether this domain forms a hexameric pore structure like many of its closest homologs or an anchor structure in the plasma membrane (Bigot et al. 2007). If the structure is that of a pore it is possible that FtsK performs some form of conjugative chromosomal resolution after the fusion of the septum, between the two divided cells. However, experiments have pointed in favor of an anchor function that localizes the protein at the septum to allow $FtsK_C$ to perform its role in the closing septum cytoplasmic space (Bigot et al. 2007).

The FtsK homologs we investigated: TrwB, TraD, and TraG - also fall under the category of coupling proteins sometimes referred to as Type 4 coupling proteins or T4CP's (Gomis-Rüth and Coll 2001). T4CP's mediate interactions in bacterial cells during DNA conjugation. Conjugation is the process of DNA exchange between bacterial cells that allows for the rapid dissemination of genetic traits throughout a bacterial species. This in turn is the source of the rapid appearance and spread of antibiotic-resistant bacterial strains. Conjugation is also utilized by mammalian and plant pathogens, such as *A. tumefaciens*, for the transfer of foreign DNA into a target cell. A single strand of the conjugative plasmid is passed from one cell to another and then the single strand in each cell is replicated and thus the genetic information shared. The single stranded DNA is chaperoned by a series of proteins known as the relaxosome. This protein-ssDNA complex is then facilitated by a T4CP at the inner membrane and shuttled to through the inner membrane into the periplasmic space where its transfer through the outer membrane and into the neighboring cell is completed by a Type 4 Secretion System (T4SS). These coupling proteins share a high homology with FtsK. Both contain an N-terminal domain (located in the inner membrane) linked (via a linking domain) to a hexameric C-terminal domain in the cytoplasm with an ATPase DNA translocase function (de Paz et al. 2010).

Our goal is to crystallize and solve the atomic resolution crystal structure of FtsK via X-ray crystallography. It is well established that it is difficult to crystalize and solve the crystal structure of integral membrane proteins, with such proteins comprising only about ~400 unique structures out of more than 80,000 depositions in the protein data bank. We conducted protein expression, purification, and crystallization for FtsK from *Thermus aquaticus* and three close homologs of the same FtsK/SpoIII family, TrwB from *E. coli*, TraD from *E. coli*, and TraG from *Agrobacterium tumefaciens* (de Paz et al. 2010).

The atomic resolution structure of any of these DNA channels would greatly add to our understanding of integral DNA channels. The detailed structural data can aid in the development of a novel class of structure guided antibiotics that can that can target DNA conjugative and conjugative-like systems. Such antibiotics would be especially useful as DNA conjugative channels are not found in human cells and are absolutely necessary for the viability of many bacteria. These targeted antibiotics could provide an effective attack in the fight against bacterial infections.

This stage of our work consisted, primarily, of screening for FtsK and its related homologs as candidates for structure determination. The criteria evaluated were milligram quantity expression level, stability, and homogeneity after purification. To this end, much of our work has been focused on protein production in *E. coli* and subsequent purification via nickel affinity and size-exclusion chromatography. We have been able to identify several constructs as suitable candidates for crystallization experiments and these have been screened for crystallization in a variety of conditions (>1000 crystallization conditions). Although we have yet to find the proper conditions to yield crystals of FtsK or the other DNA translocases, we are well under way in the foundational work necessary to reach this goal.

## II. MATERIALS AND METHODS

The following protocol was largely adapted from Newby/Straud methods (Newby et al. 2009).

**Table 1. Constructs**

List of protein constructs by protein, organism, expression vector and residue number (from N-terminal).

| Protein | Organism | Vector | Construct (residues) |
|---|---|---|---|
| FtsK | T. aquaticus | SumoPro | 1-137 |
| | | | 1-180 |
| | | | 1-189 |
| | | | 1-203 |
| | | | 1-230 |
| TrwB | E. coli | SumoPro | full |
| | | | 1-70 |
| | | | 1-71 |
| | | | 1-72 |
| TraD | E. coli | SumoPro | 1-134 |
| | | | 1-557 |
| | | pET 19.3C | full |
| TraG | A. tumefaciencs | SumoPro | full |
| | | pET 19.3C | full |

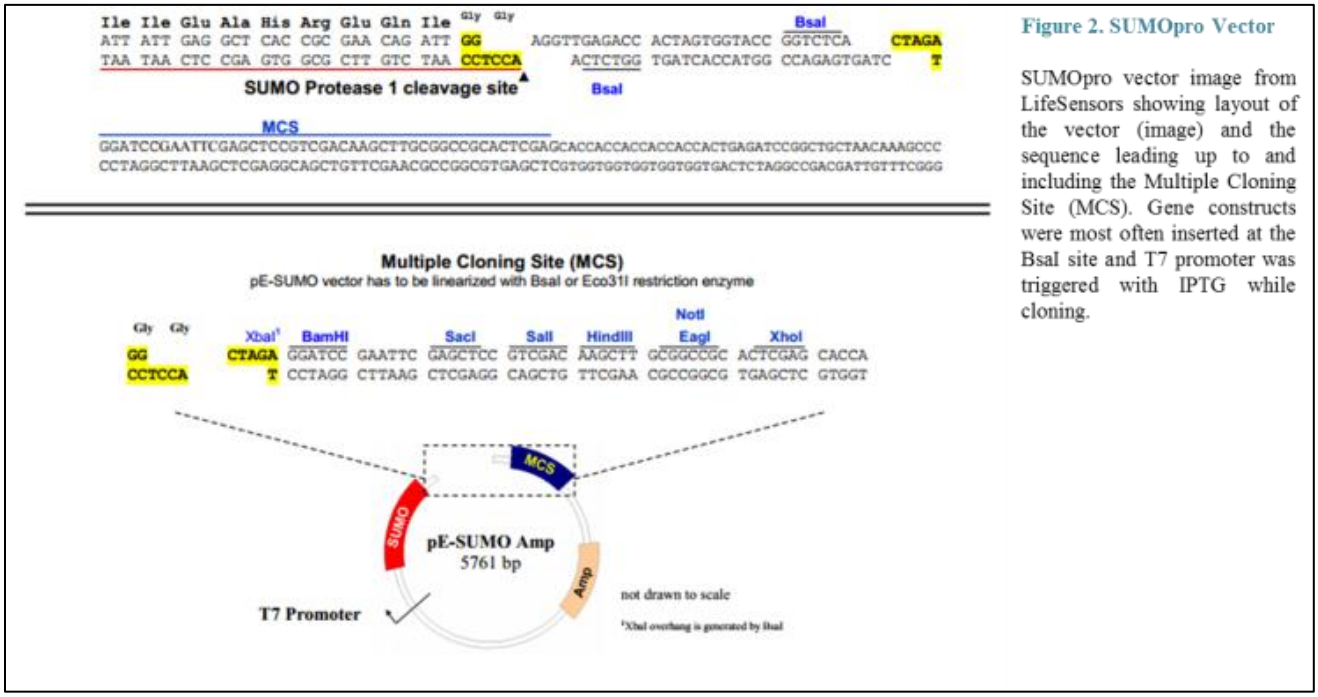


**Figure 2. SUMOpro Vector**

SUMOpro vector image from LifeSensors showing layout of the vector (image) and the sequence leading up to and including the Multiple Cloning Site (MCS). Gene constructs were most often inserted at the BsaI site and T7 promoter was triggered with IPTG while cloning.

1) Constructs and Expression

i. Genes corresponding to the protein of interest were cloned into the multiple cloning site of the SUMOpro-Amp or pET19.3C expression vector in *E. coli* (Table 1 and Fig. 2).
ii. *E. coli* cells [Bl21(DE3)] were transformed via 42°C heat shock with the transformant vector, spread on LB Agar media plates with 100µM Ampicillin, and grown o/n at 37°C.
iii. Single colony selected from plate and used to inoculate 50ml starter culture. Starter cultures grown o/n at 37°C on shaker at 220RPM.
iv. 1L LB were inoculated using 10ml of starter culture each.
v. Liquid colonies were grown o/n at 37°C on shaker at 220RPM to 0.5-1.0 $OD_{600}$ and expression initiated by adding 0.5µM IPTG. Grown another 3-5hrs at 37°C or 16 hrs at 18°C.

2) Harvest and Membrane Preparation

i. The cultures were transferred to tubes and centrifuged at 5,000XG for 25 minutes at 4°C.
ii. Discarded supernatants and resuspended all pellets in 50ml of lysis buffer (0.5M NaCl, 50µM Tris 7.4).
iii. Resuspended cells were lysed by sonication (2min sonication, 2min on ice x 5cycles)
iv. The homogenate was then centrifuged at 8000XG for 20min at 4°C, and the supernatant was retained.
v. Supernatant ultracentrifuged at 130,000XG for 60min at 4°C with Ti45 rotor. Pellet retained (lipid portion of the pellet is visible as a translucent clear solid compared to other tan cell debris).
vi. Retained supernatant and added to Ni-NTA column.

3) Solubilization

i. Membrane pellet resuspended in Membrane Resuspension Buffer (50mM Tris pH 7.4, 300mM NaCl).
ii. While stirring, the detergent (dodecyl B-D maltoside) is added to a final concentration of 1%. Mixture transferred to beaker and stirred at 4°C for ~1hour.
iii. Solubilized mixture is transferred to an ultracentrifuge tube and centrifuged for 60min at 130,000XG at 4°C.
iv. Retained supernatant and added to Ni-NTA column

4) IMAC Purification and Tag Cleavage

i. ~5ml of NiNTA slurry was added to a column and equilibrated with excess wash buffer (50mM TRIS pH 7.4, 0.3M NaCl, 0.1% DDM).
ii. Supernatant from solubilization was run through Ni-NTA to isolate tagged protein.
iii. Sample was washed with 15ml wash buffer + 40mM imidazole
iv. Sample was eluted with 15ml wash buffer + 300mM imidazole
v. The elution was added to a dialysis membrane (10kD cutoff) with 1mg of Ulp Protease
vi. Elution was dialyzed in 500ml of wash buffer at 4°C. Complete cleavage of the SUMO-Tag was observed within 1-2 hours.

vii. Sample was run through column containing 5ml of NiNTA, equilibrated with wash buffer.
NOTE: Any portion of sample containing a histidine-tag, the cleaved SUMO tag (which contains a 6Xhis tag), or the Ulp1 protease (which also contains 6Xhis tag), bound to the column while the cut DNA translocase flowed through the column.
viii. Flow-through was collected and concentrated to ~1ml for size-exclusion chromatography.

5) Size-exclusion Chromatography

i. Equilibrated Superdex200 10/300 size-exclusion column with 20mM TRIS pH7.4, 300mM NaCl, 0.05%DDM.
ii. 0.5ml of the sample injected into the column.
iii. Fractions corresponding to protein peaks collected and saved for SDS gel

6) Gel Electrophoresis

i. Fractions were run on a gel (5-20% SDS-PAGE gels, from BioRad), in SDS-PAGE buffer and stained with Coomassie in order to visualize sample purity
ii. The SEC chromatograph and stained gel lead to evaluation of the purity, quantity, and stability of the membrane protein.

7) Crystallization

i. Concentrated the purified protein to 10mg/ml with a centrifugal concentrator (Millipore, Ultrafree).
ii. The sample was screened against commercially available crystallization screens (96 condition format) using crystallization robot.

## III. RESULTS

Our crystallization trials have been unsuccessful to date. Therefore results reported will focus on protein expression, harvest, and purification results. Within this portion of the protocol we achieved strong protein yields and purity for some proteins. SEC printouts are reports from the Superdex200 SEC system.

**Standards:**

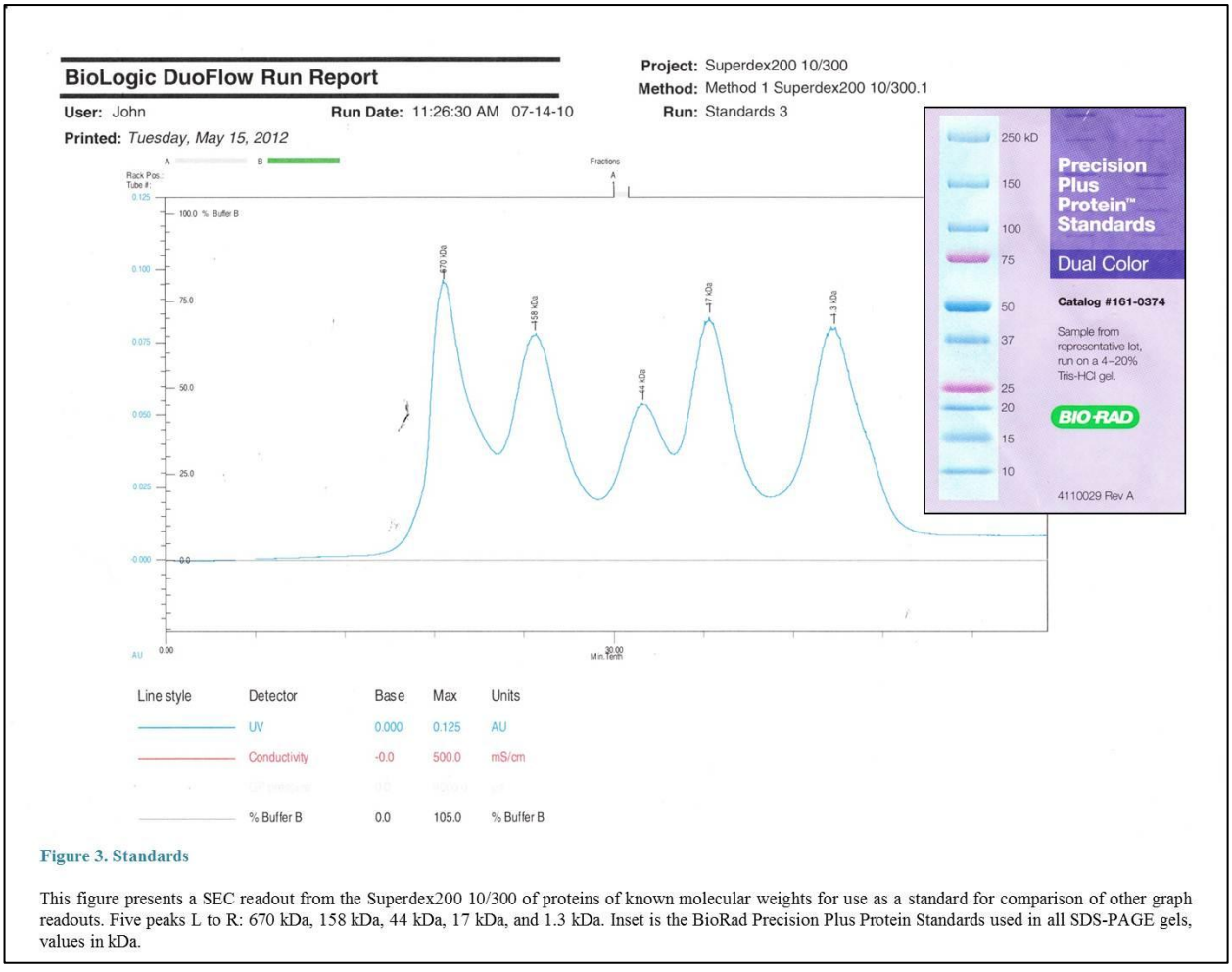


Figure 3. Standards

This figure presents a SEC readout from the Superdex200 10/300 of proteins of known molecular weights for use as a standard for comparison of other graph readouts. Five peaks L to R: 670 kDa, 158 kDa, 44 kDa, 17 kDa, and 1.3 kDa. Inset is the BioRad Precision Plus Protein Standards used in all SDS-PAGE gels, values in kDa.

**FtsK:**

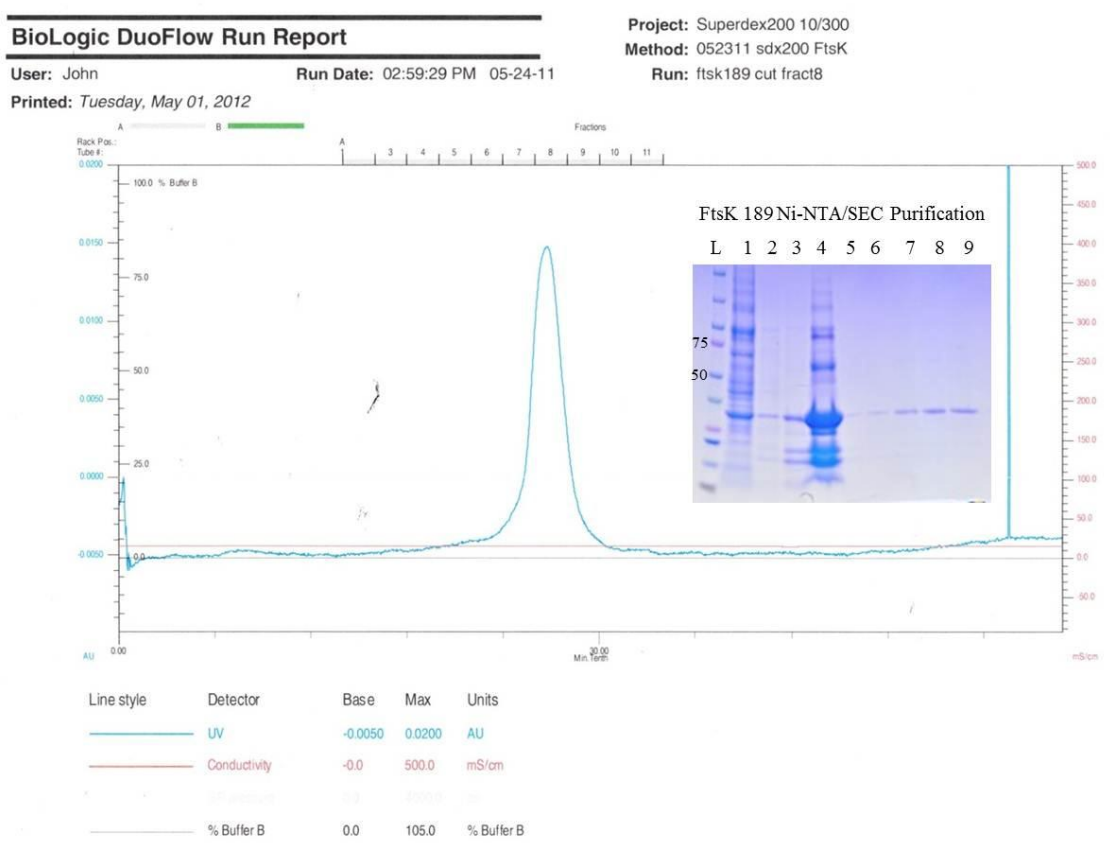


**Figure 4. FtsK 189 SEC Report and SDS-PAGE Gel Image**

This figure presents both the SEC report graph and a SDS-PAGE gel (inset) for Ulp1 cut FtsK isolate protein (189 residue construct). Protein was expressed in *E. coli*, harvested, solubilized, purified via Ni-NTA affinity column, and SEC purified (Superdex200 10/300). **Report graph**: One Gaussian UV absorbance peak can be seen (~150AU, <158 kDa) within fractions 7, 8, and 9 that corresponds with FtsK. **Gel Image:** cut TrwB ~30kDa, and SUMO-tag ~20kDa. **Lane Key:** 1 = Soluble Fraction after sonification and solubilization, 2 = Flow through (Ni-NTA) 3 = 40mM imidazole wash, 4 = 300mM imidazole elution, 7 = SEC fraction #7, 8 = #8, 9 = #9, Cut FtsK protein runs at ~30kDa. Ladder: Precision Plus Protein Stadards, Bio Rad.

**TrwB:**

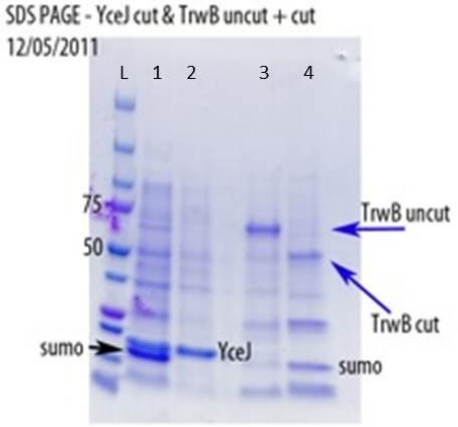


**Figure 5. TrwB full SDS-PAGE Gel Image**

This figure presents an SDS-PAGE gel for Ulp1 cut TrwB isolate protein (full residue construct). Protein was expressed in *E. coli*, harvested, solubilized, and purified via Ni-NTA affinity column. Lane Key: 1 + 2 ignore, separate protein, 3 = Flow through (Ni-NTA), 4 = 40mM imidazole wash. Ladder: Precision Plus Protein Stadards, Bio Rad.

**TraD:**

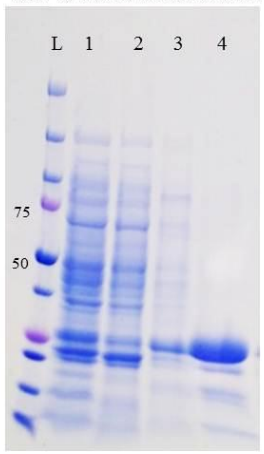


**Figure 6. TraD 134 SDS-PAGE Gel Image**

This figure presents an SDS-PAGE gel for Upl1 cut TraD isolate protein (134 residue construct). Protein was grown in *E. coli*, harvested, solubilized, and isolated via Ni-NTA affinity column. Lane Key: 1 = Soluble Fraction after sonification and solubilization, 2 = Flow through (Ni-NTA), 3 = 40mM imidazole wash, 4 = 300mM imidazole elution. Ladder: Precision Plus Protein Stadards, Bio Rad.

**TraG:**

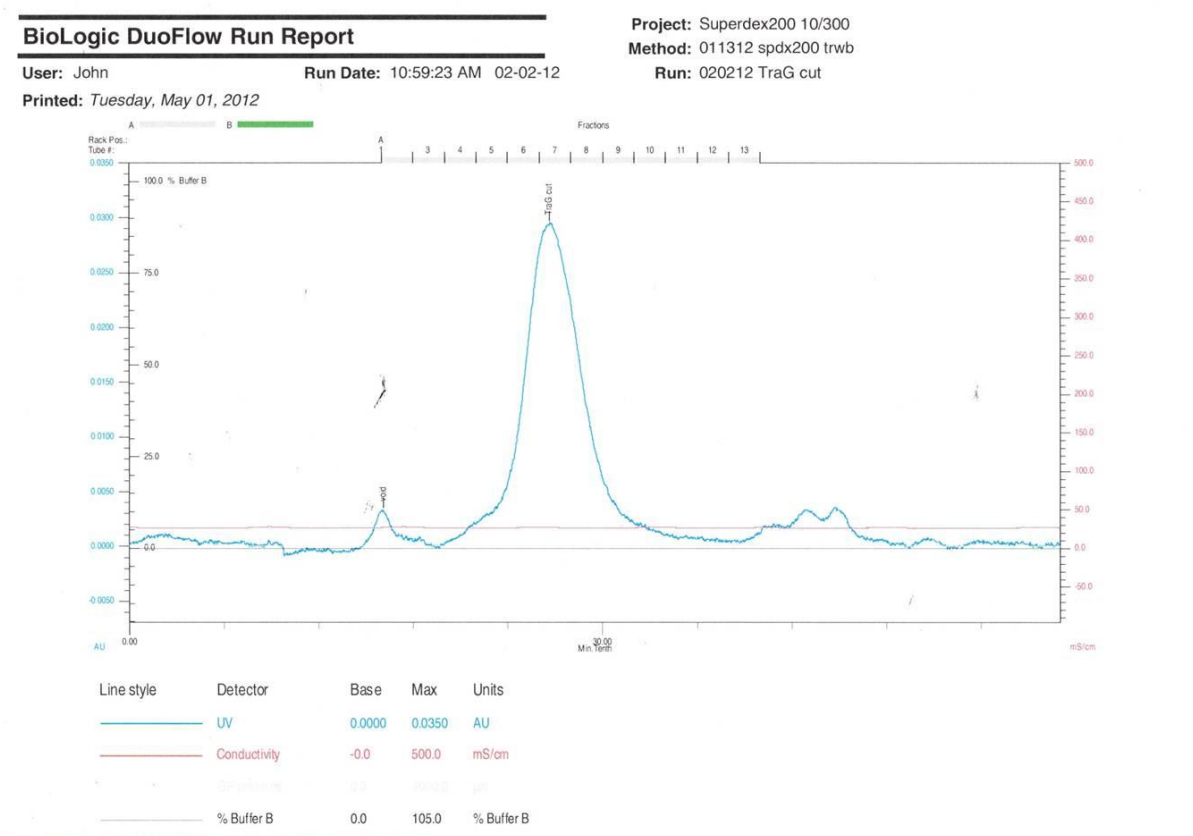


**Figure 7. TraG full SEC Report and SDS-PAGE Gel Image**

This figure presents both the SEC report graph and a SDS-PAGE gel (inset) for uncut and Ulp1 cut TraG isolate protein (full residue construct). Protein was expressed in *E. coli*, harvested, solubilized, purified via Ni-NTA affinity column, and SEC purified (Superdex200 10/300). **Report graph**: SEC trial of cut TraG from fractions 5 and 6 of previous SEC and protease digest (lanes 5+6 of gel). One Gaussian UV absorbance peak can be seen (~300AU, near 158 kDa) within fractions 5-9 that corresponds with TraG. The smaller void peak is visible at near 670 kDa. **Gel Image:** TraG uncut at ~75kDa, cut TraG ~70kDa, and SUMO-tag ~20kDa. **Lane Key:** 1 = fraction 2 (void), 2 = fraction #4, 3 = fraction #5, 4 = fraction #6, 5 = cut #5, 6 = cut #6, 7 = fraction #4, 8 = fraction #5, 9 = fraction #6. Ladder: Precision Plus Protein Stadards, Bio Rad.

## IV. DISCUSSION

The data presented in this work assesses several criteria of sample quality necessary for crystallization of the protein. The first is expression. We must obtain milligram quantities of protein in order for us to move forward to crystallization trials. The next is purity. We sought to purify a sample to greater than 95%. Other metrics that are important for this work were stability, proper oligomeric state, and homogeneity. Several proteins studied in this work produced acceptable purity, yield, stability, and homogeneity.

We used two main methods to evaluate the proteins selected for this work. SDS-PAGE gel analysis was used to assess the purity of the sample. Size exclusion chromatography (SEC) was used to both purify and assess the stability and homogeneity of the proteins. Through SEC, we were able to further purify our protein from contaminant proteins that were not separated out through affinity chromatography, and we also used it as a basis to evaluate the stability and homogeneity of the sample. The relative size of the protein based on the chromatogram and the Gaussian nature of the peak was used as a qualitative measure of purity and homogeneity. It is our current understanding that if a sample is pure and well ordered, it would form a sharp single peak corresponding to the relative size of the protein. Overlapping peaks indicated contamination of a protein similar in size to the protein of interest or disorder in the sample being evaluated. The size of the protein was also assessed by its run-through time, an indicator of aggregative state.

In both the SDS-PAGE and SEC, the actual size of membrane proteins is difficult to assess. Lipid and detergent micelles bound to the protein (often necessary for maintaining stability) often cause membrane proteins to run smaller than expected on protein gels. In addition, the proteins used for standardization of the Superdex200 are soluble proteins while our membrane proteins are less cooperative. Finally the matrix of the column may interact with our isolate protein causing a false upward shift in molecular weight (Newby et al. 2009).

The Ftsk189 construct (Fig. 4) demonstrates a high level of promise for further experiments. The single Gaussian peak of the chromatogram (~150AU) suggests high homogeneity, while the size (near 158kDa) is much larger than expected and indicative of a hexameric state. The SDS-PAGE gel shows similar purity in the fractions taken from SEC as the contaminant bands are faint. Again the protein appears slightly larger than would be expected for a monomer, (~30kDa) reflecting classic membrane protein behavior.

The full TrwB construct (Fig. 5) demonstrates a fairly high level of purity before SEC purification. There are only a few contaminating bands present beside the expected TrwB ((~50kDa) and SUMO-tag (~20kDa) bands before and after being Ulp1 cut. The contaminating bands are of disparate sizes from the desired bands, and thereby likely to be separable through SEC. The protein again runs at a larger size than would be expected for TrwB in a monomeric state.

The TraD 134 construct (Fig. 6) demonstrates high purity with weak contaminating bands. The protein ran slightly larger than expected (~25kDa) but again this may be expected and does not negate the usefulness of the purified protein for crystallization.

Lanes 5 and 6 of the SDS-PAGE gel image of TraG (Fig. 7) were assessed by SEC. The chromatogram shows a homogenous Gaussian peak (~300AU, near 158kDa) with the exception of a slight shoulder approaching the void volume. The smaller void peak (~50AU, >670kDa) may be indicative of a larger oligomeric state or other contaminant proteins. The SDS PAGE gel image indicates purity and clear disparity between cut TraG (~70kDa) and SUMO-tag bands (~20kDa) and those of contaminant proteins. The TraG bands are again larger than expected.

Having attained a protein of acceptable purity, oligomeric state, yield (in milligram amounts), and stability, we proceeded to crystallization trials. When crystallization attempts were unsuccessful (usually, we screened over 1000 conditions), we returned to the growth and purification phases to modify the processes and produce a more pure and stable protein structure. Within the protocols outlined above we hope other researchers may find tools to increase the quality and efficiency of the membrane protein isolation process for crystallization towards crystallographic trials. The importance of these proteins as targets for antibiotic applications makes it important for human health and our ongoing fight against rapidly evolving pathogenic species.

## V. REFERENCES

Bigot, Sarah, Viknesh Sivanathan, Christophe Possoz, François-Xavier Barre, and François Cornet. 2007. "FtsK, a Literate Chromosome Segregation Machine." *Molecular Microbiology* 64 (6) (May 18): 1434–1441. doi:10.1111/j.1365-2958.2007.05755.x.

Dorazi, R, and S J Dewar. 2000. "Membrane Topology of the N-terminus of the Escherichia Coli FtsK Division Protein." *FEBS Letters* 478 (1-2) (July 28): 13–18.

Gomis-Rüth, F X, and M Coll. 2001. "Structure of TrwB, a Gatekeeper in Bacterial Conjugation." *The International Journal of Biochemistry & Cell Biology* 33 (9) (September): 839–843.

Massey, Thomas H, Christopher P Mercogliano, James Yates, David J Sherratt, and Jan Löwe. 2006. "Double-stranded DNA Translocation: Structure and Mechanism of Hexameric FtsK." *Molecular Cell* 23 (4) (August): 457–469. doi:10.1016/j.molcel.2006.06.019.

Newby, Zachary E R, Joseph D O'Connell, Franz Gruswitz, Franklin A Hays, William E C Harries, Ian M

Harwood, Joseph D Ho, et al. 2009. “A General Protocol for the Crystallization of Membrane Proteins for X-ray Structural Investigation.” *Nature Protocols* 4 (5) (April 9): 619–637. doi:10.1038/nprot.2009.27.

de Paz, Héctor D, Delfina Larrea, Sandra Zunzunegui, Christoph Dehio, Fernando de la Cruz, and Matxalen Llosa. 2010. “Functional Dissection of the Conjugative Coupling Protein TrwB.” *Journal of Bacteriology* 192 (11) (June): 2655–2669. doi:10.1128/JB.01692-09.

Strick, Terence R, and Audrey Quessada-Vial. 2006. “FtsK: a Groovy Helicase.” *Nature Structural & Molecular Biology* 13 (11) (November): 948–950. doi:10.1038/nsmb1106-948.