Mycobacterium Tuberculosis: Finding a Cure Through the Study of Membrane Proteins

By: April Jue

Abstract:

Mycobacterium Tuberculosis is a deadly pathogen that kills millions of people a year. Its genome has been sequenced and currently, its membrane proteins are being studied. Membrane proteins are difficult to express but there is hope that if we are successfully able to express and structure them that we will be able to understand their functions more clearly and develop new drugs that combat the pathogen.

During this internship, I concentrated on helping to express and purify the membrane proteins of *Mycobacterium Tuberculosis*.

Introduction:

Almost three million deaths a year are attributed to *Mycobacterium Tuberculosis*, the bacterium pathogen that causes tuberculosis. Currently this pathogen is becoming more resistant to many of the drugs that are used to treat it. Scientists thus need to develop new, more effective drugs for combating the pathogen. One of the keys to this challenge lies in the extensive study of the membrane proteins of *Mycobacterium Tuberculosis*.

Membrane proteins lie in the outer membrane of a cell and are responsible for carrying information such as hormonal signals and nutrients between the outside world and the inside of the cell. In addition they are able to convert the electrical and chemical membrane potentials into forms of energy. Drugs that would target these membrane proteins would most likely have very profound physiological effects since these proteins are responsible for such life-altering functions within the cell and its surroundings. It is

even suggested that in the future more than 90% of neuroscience drugs may be targeted towards membrane proteins.

The entire genome of *Mycobacterium Tuberculosis*, containing as many as 1200 open reading frames (ORFs) coding for membrane proteins, was sequenced and published in *Nature* June 1998. Using this sequenced genome, the reading frames coding for the membrane proteins can be cleaved and inserted into suitable plasmid vectors for cloning and expression. If the protein is successfully expressed, it can be grown in large-scale production and purified. The purification technique takes advantage of the 6X Histidine-tag located at the C-terminal end of the protein. The Histidine-tag is able to bind to the nickel charged column and thus retain the protein of interest. Another purification technique is ion-exchange chromatography that separates proteins based on their charge.

Once the protein is purified different techniques can be applied to try to determine the protein's structure. A combination of electron diffraction and X-ray crystallography, solution and solid-state NMR, and mass spectrometry can be used. Often the structures can be very difficult to determine. It is nevertheless an important task to complete because often a proteins biological function can be inferred from its structure. Knowledge of the different membrane proteins biological functions will be very useful in developing new drugs that fight or even help prevent tuberculosis.

Protein Expression System Overview:

The plasmid that has been used is called pET 29b. The gene coding for the protein (the insert) is put into the plasmid in place of the Nde1-Xho1 sequence. In this

particular experiment, the plasmid is transformed into BL21 (DE3) cells lacking the Lon and ompT proteases that degrade proteins during the purification process. Ordinarily, the gene expression is under strict control of a T7 promoter and so the cells are unable to express the protein without induction. When cells are induced with isopropylthio-β–D-galactoside (IPTG), the repressor protein is removed from the promoter region and T7 RNA polymerase is free to begin transcription followed by translation of the gene encoding for the protein.

Materials:

LB medium was used to grow the *E. Coli* cells in. Ten grams of NaCl, ten grams of bacto-tryptone, and five grams of bactoyeast extract were added to approximately 800 mL of deionized water and stirred until dissolved. The pH of the medium was adjusted to approximately pH 7-8 by pipetting the appropriate amount of 1.0 N NaOH into the solution. Deionized water was added to give a final volume of one liter (1000 mL). The container was covered with aluminum foil and autoclaved for 20 minutes.

DNA samples were run on agarose gel (1%) to check for the presence of the insert or plasmid in the solution. Half a gram of agarose (0.5 g) was added to fifty milliliters (50 mL) of a Tris Borate buffer that contained ethidium bromide. The solution was heated in a water bath or the microwave until the agarose had dissolved. Then the flask was allowed to cool down until it could be comfortably held. The solution was poured into the gel holder containing the combs. Any bubbles were pushed to the side and the solution was allowed to polymerize. Right before use, the combs were removed and the Tris Borate buffer was poured over the gel.

Protein samples were run on SDS Polyacrylamide Tricine Gel to check for protein expression. The resolving gel was made by mixing 1.2 mL of 49.5 % Acrylamide/N,N'-methylenebisacrylamide (AA/Bis), 2.0 mL of gel buffer (1.0 M Tris pH 8.3, 2% SDS), 2.0 mL of Glycerol:Water (56:44), and 0.75 mL of water. Then 60 μL of Ammonium Persulfate (APS) 10% and 12 μL of TEMED were added to the solution to catalyze the polymerization of the gel. The resolving gel was poured into the gel holder. A layer of water was added on top to help the gel settle evenly. The stacking gel was made by mixing 0.25 mL of 49.5% AA/Bis, 0.25 mL of gel buffer, and 2.0 mL of water. Then 25 μL of APS 10% and 5 μL of TEMED were added to the solution to catalyze the polymerization of the gel. Once the resolving gel solidified, the water was poured off and the comb was placed into the gel holder. The stacking gel was poured to the top and allowed to polymerize. Before use, the gel holder was placed in Tricine buffer and the comb was removed.

Agar plates were also used to grow colonies of cells on. Seven and half grams (7.5 g) of bacto-agar were added to half a liter (0.5 L) of LB medium. The solution was autoclaved. After cooling, 500 μL of kanamycin was added and the solution was poured into the plates. Any bubbles that appeared were heated. As the solution in the plates cooled, it polymerized.

An 8X Binding solution was used in protein purification. Forty milliliters (40 mL) of Tris pH 7.9, 58.44 g NaCl, and 0.688 g Imidizole were added to a tube and deionized water was added to a final volume of 250 mL.

An 8X Washing solution was used in protein purification. Forty milliliters (40 mL) of Tris pH 7.9, 58.44 g NaCl, and 8.25 g Imidizole were added to a tube and deionized water was added to a final volume of 250 mL.

A 4X Elution buffer was used in protein purification. Twenty milliliters (20 mL) of Tris pH 7.9, 29.22 g NaCl, and 68.8 g Imidizole were added to a tube and deionized water was added to a final volume of 250 mL.

A 2X Strip buffer was used in protein purification. Ten milliliters (10 mL) of Tris pH 7.9, 14.61 g NaCl, and 125 mL of 0.5 M EDTA were added to a tube and denionized water was added to a final volume of 250 mL.

A 2X Charge buffer was used in protein purification. $NiSO_4$ (5.26 g) was added to 200 mL of water.

Methods:

Plasmid Propagation and Purification:

E. Coli cells containing a plasmid vector with insert were placed in 5 mL of LB medium and 5uL kanamycin to grow overnight. The following day 10 mL of LB medium was inoculated with 10 μL of kanamycin and 500 μL of overnight culture. The sample was allowed to grow for 3 hours on the shaker at 250 rpm.

In order to transfer the plasmid from the cloning cells to the expression cells, the plasmid must be purified. We used QuickSpin Plasmid Purification Kit. (Qiagen Inc., USA) According to their procedure, the samples were spun down 3-4 minutes at 8000 rpm. The supernatant was discarded and the pellet was resuspended in 250 μL of Buffer P1 and transferred to a microcentrifuge tube. Two hundred fifty microliters (250 μL) of

Buffer P2 was added and the tube was gently inverted 4-6 times. Next, 350 μL of Buffer N3 was added and the tube was immediately inverted 4-6 times. The samples were placed in the centrifuge for 10 minutes at 14,000 rpm. The supernatant was transferred to a QIAprep column and centrifuged for 1 minute at 12,000 rpm. The flowthrough was discarded and the QIAprep spin column was washed by adding 0.5 mL Buffer PB and centrifuging 1 minute at 12,000 rpm. Seven hundred fifty microliters (750 μL) of Buffer PE were added and the samples were centrifuged 1 minute at 12,000 rpm. The flowthrough was discarded and the samples centrifuged for an additional minute at 12,000 rpm. The QIAprep column was placed in a clean 1.5 mL microcentrifuge tube and 50 μL Buffer EB was added to the column. The solution sat for 1 minute and then was centrifuged for 2 minutes at 12,000 rpm. The liquid was transferred to a clean aliquot and refrigerated. It was run on an agarose gel to check for the presence and purity of the plasmid.

If the plasmid was present, insert expression screening was done to check for the presence of the insert in the plasmid. Colonies were transferred from plates into microtubes filled with 200 μL of water. Holes were poked on top and samples were placed in a 100° C water bath for 5 minutes. Samples were spun down at 12,000 rpm for 2 minutes. A PCR mix was made using the supernatant (DNA template) from the spun-down cells (5 μL), a reverse primer (2 μL), a forward primer (2 μL), a Master Mix (10 μL) (Eppendorf Inc., USA) and water (6 μL). A PCR reaction was run for 2 hours 30 minutes. The samples were run on agarose gel to check for the correct weight of the isnert..

Protein Expression and Purification:

To transform the plasmid from the cloning cells to the expression cells, samples were placed on ice and allowed to thaw. Two μL of ligation reaction (purified plasmids) were added to the cells. The samples were left on ice for 5 minutes. The tubes were heated for 30 seconds in 42° C water bath and then placed on ice for 2 minutes. One-hundred and fifty μL room temperature SOC medium was added to each tube. The samples were placed on a shaker at 250 rpm for 1 hour. Then 70 μL of transformation sample was spread onto each LB agar plate. The plates were allowed to culture overnight. The following day, cell samples were taken from the plate and put into 5 mL LB medium containing 5 μL kanamycin. This sample was cultured overnight. The following day, plasmid DNA was purified and run on gel again to check for the plasmid.

To grow expression cells, cell samples were placed in 5 mL of LB medium and 5uL kanamycin to grow overnight. The following day 10 mL of LB medium was inoculated with 10 μL kanamycin and 500 μL of overnight culture. Tubes were placed on the shaker at 250 rpm. Once optical density at a 600 nm wavelength (OD_{600}) reached between 0.4-0.6, the solution was induced with 1 μL IPTG. Prior to induction, a 1 mL aliquot of uninduced sample was taken from the solution. The induced solution was allowed to culture for 3 more hours at 250 rpm.

To harvest the expression cells, the samples were placed in a centrifuge and spun down at 5000 rpm for 10 minutes (1 mL aliquot) or 15 minutes (large induced sample). The supernatant was discarded and the cells were resuspended in 1 mL water. At this point, the induced sample was transferred to a 1.5 mL centrifuge tube and stored at -20°

C. The uninduced aliquot was centrifuged at 5000 rpm for another 10 minutes. The supernatant was discarded and the cells were resuspended in 100 μL of 2% SDS.

To sonicate the induced sample, and thus separate the soluble and insoluble parts of the cell, the sonic dismembrator was set at 5. The tubes were placed on ice. The tip of the dismembrator was placed in the middle of a cell sample and the sample was sonicated for 20 seconds. Then the sample was allowed to sit for 20 seconds. This procedure was repeated 5 times for each tube. The tip of the sonic dismembrator was washed with water when the cell sample was changed. The samples were centrifuged at 10,000 rpm for 10 minutes. The supernatant (soluble portion) was separated into a separate tube. The insoluble portion was resuspended in 1 mL of water and centrifuged at 10,000 rpm for 10 minutes. The supernatant was discarded and the sample resuspend in 100 μL of 2% SDS.

To prepare the samples for the SDS Tricine gel, 15 uL of the uninduced sample and 15 uL of 2X Loading Buffer were put in a tube. Fifty uL of the soluble induced sample and 50 uL of the 2X Loading buffer were put in another tube. Fifty uL insoluble induced sample, 20 uL urea, and 70 uL 2X Loading Buffer were put in a third tube. Holes were poked in each tube and then the tubes were put in the microwave for 15 seconds. The samples were run on SDS Tricine gel to check for protein expression.

If protein expression was present, the protein was grown in large-scale production. Overnight cultures of cells were prepared in LB medium and kanamycin. The following day, one liter (1 L) of LB medium was inoculated with 5 mL of overnight culture and 1 mL of kanamycin. The cells were cultured, induced, and harvested as previously described. The samples were run on SDS Tricine gel to verify protein expression.

To purify the protein, an affinity column was set up. To prepare the column, nickel charged resin (Novagen, Inc., USA) was poured into a column. The column settled by allowing water to drip out. Thirty mL $NiSO_4$ charge buffer was run through the column in order to charge it. The cell samples were sonicated, frozen for 15 minutes, and sonicated once again. Then the samples were spun down for 10 minutes at 14,000 rpm. The supernatant was removed and the pellet was resuspended in water. The supernatant (loading fraction) was run through the column and collected in a tube. Next, 5 mL Binding Buffer was run through the column and collected in separate tube. Then10 mL Wash Buffer was run through the column and collected. Next, 10 mL Elution Buffer was run through column and collected. Finally 10 mL strip buffer was run through column. To wash the column, water and SDS were run through the column. Each of the collected fractions was run on a SDS Tricine gel to check for the purified protein. If the protein appeared in all fractions, the fractions were put in a dialysis tube that lowered the concentration of Imidizole.

Another way of purifying the protein was by using an ion-exchange column. The column was made with DEAE sepharose resin (BioRad Inc., USA). The sample was run through the column and collected in a tube. Then 20 mM TrisHCl, 500 mM NaCl buffer was run through the column and collected in a separate tube. Each of the fractions was run on a SDS Tricine gel to check for the purified protein.

Results:

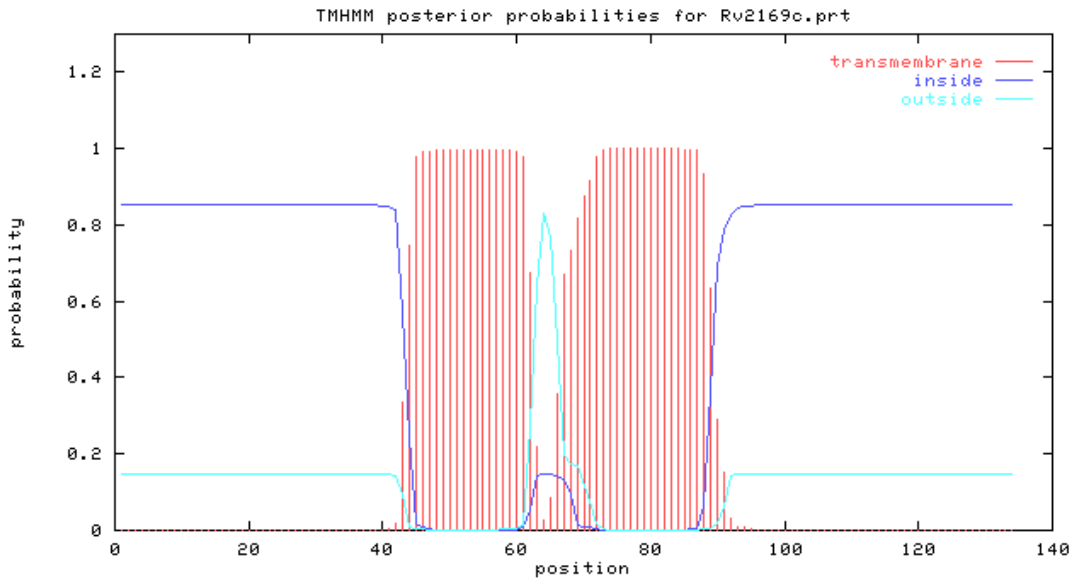
Expression Trials:

Gene Name	DNA size (kb)	Protein M.W. (Da)	pET 29b plasmid	Transformation	Tried	Expressed	Large Scale
Rv1171	438	15168.6	+	+	+	+	+
Rv1342c	360	13369.8	+	+	+	+	+
Rv0011c	279	10418.2	+	+	+	+	+
Rv0514	297	10273.6	+	+	+	-	-
Rv0870c	387	13673.6	+	+	+	-	-
Rv3857	195	6748.47	+	+	-	-	-
Rv3789	363	13361.8	+	-	-	-	-
Rv2169c	402	14556.7	+	+	+	+	+
Rv2076	249	9029.31	+	+	-	-	-
Rv0460c	237	8063.45	+	+	-	-	-
Rv1305	243	8039	+	-	-	-	-
Rv0882	282	9612.04	+	-	-	-	-

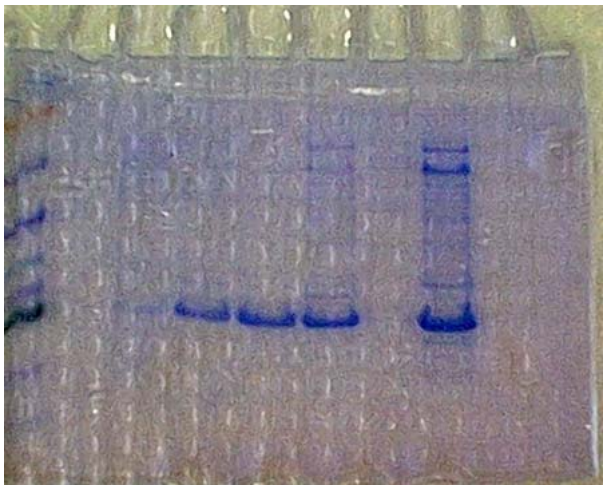
All of the genes we dealt with were successfully inserted into the pET 29b plasmid. However, only three-quarters were transformed into the BL21 (DE3) cells for expression. Out of 6 genes that we tried to express, we were successful in expressing 4. All of the genes that expressed were successfully expressed in large-scale production as well.

Purification Trials:

Rv2169c is the unnamed membrane protein that we attempted to purify. This protein is composed of 134 amino acids and has a molecular weight is 14456.7 Daltons. It has an isoelectric point of 11.7109. There is not a lot known about this protein, but it is believed to have one transmembrane segment as seen below.



Nickel-charged Affinity Column Gel:

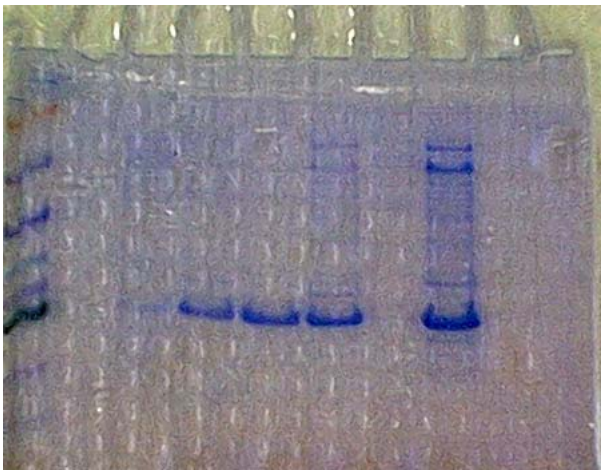


The table below shows which fraction was run in each column of the gel for the nickel-charged affinity column.

1	2	3	4	5	6	7	8	9	10
	Marker	Binding	Loading	Washing	Elution	Strip	Insoluble	Insoluble	

When the nickel-charged affinity column was used, the protein was supposed to bind to the column and then be eluted in the elution fraction. The protein correctly appeared between 110000 and 170000 Daltons. However, the above gel and table show that the protein started eluting in the Washing Buffer. The levels of Imidizole (60 mM) in this buffer must have been too high for this protein. Imidizole is a chemical that competes with the protein for the binding sites on the column and thus helps elute the protein from the column.

Ion-Exchange Column Gel:



The table below shows which fraction was run in each column of the gel for the ion-exchange column. The column binds negatively charged ions.

3	4	5	6	7
Marker	Concentrated Loading	Unconcentrated Loading	Tricine-HCl 20 mM	Tricine-HCl 20 mM, NaCl 500 mM

Since our protein was positively charged at pH 7.9, this protein was supposed to elute right through the ion-exchange column while the negative impurities were supposed to bind to the column. The above gel and table show that this column did help to purify the protein. The concentrated loading sample shows a very weak, but purified, band of protein. We concentrated it down even more in order to try to get a darker band, but we ran out of time and were unable to run another gel and see the results.

Discussion:

In the past, it has characteristically been very difficult to express membrane proteins. Our results support this. Out of the 12 genes and proteins that we worked with, we were only able to express 4 of them in 8 weeks. Attempting to express, much less structurally characterize, all of the membrane proteins of *Mycobacterium Tuberculosis* is obviously a very challenging task. About a third of each genome codes for membrane proteins, and yet today less than 1% of known protein structures are membrane proteins. Even if just fifty new membrane protein structures are determined during the course of this project, the number of known protein structures will triple.

One reason that these membrane proteins are so difficult to express and purify is because they are hydrophobic and need a lipid environment to express well and obtain a correct, native 3D structure. They often can't express well in foreign cells and even when they do, they aggregate to form inclusion bodies that are difficult to separate. They usually express in the insoluble portion of the cell but this is not advantageous because it is here that a lot of extraneous junk congregates to make them even harder to purify.

To help achieve successful membrane protein expression, many different vectors and bacterial host strains are available for use. Some vectors may carry a T7lac promoter while others may carry just a T7 promoter. Some may be designed for targeting proteins that contain an N-terminal His Tag and a protease recognition site. Others may only be able to transcribe RNA and not translate it. Thus, a vector that can express one protein well might not be a suitable choice for expressing another protein. In my experience during this internship, pET 29b was the only plasmid vector that was used for expression. Perhaps other vectors can be used to try to express proteins that are currently unable to be expressed. Bacterial host cells also possess different mutations that may or may not make them suitable hosts for a particular plasmid vector. Having so many choices and combinations of plasmids and host cells is what will help fuel this projects successful completion.

Acknowledgements:

I would like to thank Dr. Gao, our mentor, for giving me the opportunity to take part in this project. I would also like to thank Dr. Alla Korepanova for helping explain and guide me through the project and taking time to work with me in the lab. I thank Yuanzi Hua and the rest of the lab staff for working with me in the lab as well.

References:

1. Sambrook, J., Fritsch, E.F., and T. Maniatis, Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1989.
2. http://genolist.pasteur.fr/TubercuList/

3. http://magnet.fsu.edu/~changlin/MBPweb/
4. http://magnet.fsu.edu/~changlin/MBPweb/proposal/Final_1-17a.htm
5. http://magnet.fsu.edu/~changlin/MBPweb/proposal/Final_18-24a.html
6. http://magnet.fsu.edu/~changlin/MBPweb/proposal/Final_26-33a.html