GENERATION OF A CYSTEINE-LESS BACKBONE YCF1 PROTEIN FOR ELECTRON PARAMAGNETIC RESONANCE SPECTROSCOPY

By

BRIANA BEATRIZ AGUILAR

A Thesis Submitted to The W.A Franke Honors College

In Partial Fulfillment of the Bachelor's Degree
with Honors in

Biochemistry (B.S.)

THE UNIVERSITY OF ARIZONA

M A Y 2 0 2 3

Approved by:

Dr. Thomas Tomasiak
Department of Chemistry and Biochemistry

Abstract

There are a lot of research applications to YCF1 including research treatments for various diseases. One of those illnesses, cystic fibrosis, has a protein that does not work well known as the cystic fibrosis transmembrane conductance regulator, CFTR. YCF1 and CFTR are homologs, thus the research conducted on understanding cysteine modification and conformation space could be utilized for understanding how to improve current treatment options for cystic fibrosis. The goal of the research is to have a multiple cysteine-less mutated YCF1's that are purified and ready for their confirmational space to be examined. The following was utilized for purification: primer design for mutagenesis, Q5 mini-prep, QIAprep spin miniprep kit, yeast growth, and nickel purification. The primary technique being used to examine the conformational space is EPR, and Cryo-EM could further confirm details about the orientation. The end result of the lab was purified proteins that were thoroughly purified and ready for EPR spin labels to be attached.

Background

The yeast cadmium factor 1 protein, known as YCF1, is part of a class of ATP Binding Cassette (ABC) transporters[2]. These transporters are one of the largest protein superfamilies across bacteria, yeast, and mammalian organisms[10]. The transporters can aid in the transport of amino acids and lipids, there is a high affinity for hydrophobic molecules and consequently, it can transport a diverse range of drugs[1]. Within the ABC transporter family, there are subfamilies identified through their topology (Fig 3.). ABCC

has many clinical implications for illnesses[12]. One of those families is related to the most common, fatal genetic disease in the United States, cystic fibrosis[5]. This disease affects the lungs which lead to the thickening of mucus membranes. Cystic fibrosis transmembrane conductance regulator, CFTR, is a protein that is vital for salts to be passed through membranes[8]. After ribosomes read the RNA for the translation process, the protein can then move throughout the cell toward the cell surface, in a process called trafficking[4]. At the cell's surface, the CFTR protein functions as a chloride channel[13]. For an individual with cystic fibrosis membrane the Cl^- ions, which are key for the regulation of tissue homeostasis, get trapped on one side of the membrane[13]. Meaning one of three things, the protein either does not mature, there is not enough quantity of the protein, or none of the protein was produced[9]. CFTR modulators are used as a treatment option for cystic fibrosis[5]. There are correctors, such as Ivacaftor, and potentiators, such as lumacaftor. There is currently one medication out on the market, Orkambi, that costs on average $22,000 USD per month[13]. As this is such an expensive treatment option monthly, it is vital to continue research in this field to explore other potential treatment options. The reason why YCF1 was primarily studied is that YCF1 and CFTR are homologs and belong to the same ABCC family[4]. YCF1 is a transporter that undergoes multiple conformational changes in a dynamic process, thus analyzing the conformational space is important to understand the transporter cycle[11]. That way, elucidating these transporter cycles is crucial for the further development of therapeutics for diseases like cystic fibrosis.

Various methods could be used in future experiments to evaluate the conformational space of YCF1 as cryogenic electron microscopy (cryo-EM), proteomics

mass spectrometry approaches, biochemical assays as well as electron paramagnetic resonance (EPR) Cryo-EM can be used to observe macromolecular supercomplexes or even basic structures. The way of stabilizing the molecule is through flash-freezing solutions of the YCF1 protein. It is then placed in an environment with an abundance of electrons and microscopic imaging can reveal detailed information about the molecule. Currently, cryo-EM methods as well as biochemical assays have been extensively used to study YCF1 and $CFTR_4$. In the approach taken in this paper, we focused on exploring the EPR technique as a complementary tool to understand YCF1 structure and conformational changes. Paramagnetic molecules undergo a unique spectroscopy signal that can be used to characterize the different conformations and structural frameworks that a protein sample can undergo. To prepare for EPR spin labels, the first step is to reintroduce pairs of cysteines, attach spin labels, and measure the EPR. A cys-less mutant protein is crucial for this type of experiment because it will allow for the reintroduction of cysteine pairs in regions of interest, for spin labeling and consequent EPR measurement. In terms of spin labels, those labeling chemicals have the capacity to bind to cysteine residues due to the formation of disulfide bonds. There is an EPR signal detection that is used to then undergo multiple power saturation experiments. The EPR spectrum could then show one of two things. It could either display non-saturating conditions in the absence of a standard baseline or in the presence of one. The g-factor and spin density could then be analyzed, along with the conformational space the EPR has noted. If the cryo-electron microscopy route was taken, it would be mainly used to analyze the structure via stabilizing the molecule. The end goal for the project presented in this paper is to have a purified cys-less mutant

product that can be used for different EPR measurements and ultimately to sample the conformational changes happening in the dynamic Ycf1 protein.

Methodology

Structure Analysis

The first step in the process entails structure analysis. It is necessary to evaluate the current Ycf1 protein models from PDB, in order to identify cysteine residues that can be probed for EPR. Cysteines are the target residues that can be used for spin labeling probes necessary for EPR. These regions also need to be solvent exposed in order for probes to interact and react by disulfide bond formation. Considering the importance of solvent accessibility for the reaction to occur, we then ignore any cysteine that is hydrophobically hidden, non-solvent exposed regions. 8 cysteine residues were identified that fulfilled these requirements in the protein structure. Then, we evaluate the corresponding plasmid DNA regions to be modified one at a time by site-directed mutagenesis. The purpose of mutagenesis is to replace the 8 cysteines with serine residues. That way, we can generate a cys-less backbone plasmid and protein that will be further used for EPR experiments. Serines are the optimal amino acids for the replacement due to its size (similar to cysteine) and chemistry (non-disulfide reactant). Ultimately, the goal is to neutralize the potential for disulfide bond formation between unwanted cysteine residues and EPR probes (Cysteine less mutant) (See fig. 2).

Plasmid

The specific YCF1-type cells used in the series of experiments have 11,124 base pairs. Several regions of important interest can be seen (fig. 2). The yeast cells can encode more than 1,800 amino acids, thus its versatility of experimentation must be well explored. One of the sections is the M13 origin, which contains 507 nucleotides. It is utilized in several recombinant DNA processes. There is also an ampicillin-resistant gene that is inserted into the plasmid. When an agar plate was used ampicillin was utilized to eliminate any unwanted bacteria that could lead to contamination and other unwanted cell types.

Primer Design for Mutagenesis

The first step in the process entails structure analysis. Selection of up to 8 cysteine regions in the plasmid DNA is chosen to be modified one at a time. The regions were locations 31, 57, 526, 951, 1086, 1157, and 1506. The conformational changes and other potential changes are then observed. The cysteines selected were on the exterior of the protein to ensure that they are solvent-exposed (See fig. 2). The higher the frequency of the new amino acid made with the DNA substitution, the better the protein expression. Successful protein expression is desired for future experiments featuring the purifying protein and EPR. In the approach taken with EPR, the DNA used in YCF1 was imputed into a sequence data sheet. Primers were then made in accorders to the most ideal places for cis-less mutations. Mutagenesis PCR will then replicate the

new and old strands of DNA. The KLD reactions will ensure that the wild type is digested, leaving the DNA of interest. It is ideal only to substitute one base pair per modification site in mutagenesis procedures, and also to select the highest frequency codon for the desired amino acid to be changed. The sequence of the protein is problematic for base changes because it thermodynamically poses a challenge, due to the size of the protein. Designing the primers in the DNA template is the next step. Forward and reverse primers were used in a non-traditional format. NEBaseChanger is an online website that has an algorithm to create the best primers to use for each cysteine modification, via the substitution method. This allows for no frameshift in the day, thus changing only one amino acid. Only the forward primer carrier the change to mutate the DNA. The primer design is for substitutions that occur in the middle of the forward primer. A minimum of 10 complementary nucleotides have to be present on the 3’ end of the forward primer. The 5’ end reverse primer anneals to the 5’ end of the forward primer. This allows for a more efficient primer to be created, as there is typically only a forward primer used in this process.

Q5 mini-prep

The Q5 Site-Directed Mutagenesis Kit was utilized to miniprep the Neb 5 cells. The Neb5 cells were selected over the top 10 cells as a higher rate of success has been shown with Neb5 cells. There are three main phases to the kit. First is the amplification of exponential amplification (PCR) and then there is kinase, ligase, and DPN1 (KLD). The last phase is transformation. Exponential amplification entails a mixture to be made.

The mixture contained a Q5 Master mix at 12.5 uL, 1.25uL of forward primer (10uM), 1.25uL of reverse primer (10uM), 9.0 nuclease-free water, and 1uL of template DNA. The cycling conditions for the PCR were 25 cycles with a minimal extension time of 2 minutes. The KLD reaction is the second phase of the mini-prep. It combines 1uL of PCR product, 2X KLD Reaction Buffer 5uL, 10X KLD Enzyme mix at 1uL, and 3uL of H_2O. Due to the nature of the reaction, adding the enzyme last is crucial. Otherwise, the reaction will not progress in a cohesive manner. The mix was left to incubate at room temperature before the third phase took place. Transformation first occurred when 5uL of KLD mix is 50uL of the Neb5 Cells. The cells then underwent several incubation periods on ice and had a heat shock. The last steps allowed for 950uL of SOC to be added to the cells and incubated for an hour before being spread onto a plate for growth overnight.

A plasmid first needed to be introduced to bacterial cells. In order to ensure that only the correct plasmid will be replicated throughout the process, an ampicillin resistance gene was added. Propagation would then occur in the bacteria cells to then allow the cells to be prepared for transformation. Various miniprep techniques were used to then create DNA stock. PCR mutagenesis would amplify the cysteine mutated plasmid. The PCR tubes along with the YCF1 cells were then introduced together so the yeast cells could grow. The top 10 cells were used in QIA miniprep and then sent off to sequencing to confirm that only the plasmid DNA was left in the solution. The sequencing process is Sanger Sequencing. The DNA is then purified furthered and plated in a YNB-HIS plate that allows for the plasmid to be translated. Transformation of cells could then accrue and be later verified in accuracy.

QIAprep Spin miniprep Kit

5mL of bacterial overnight culture was centrifuged and resuspended the cells in 250uL in a buffer (p1). 250uL was then added and inversion occurred to allow for the lysis reaction to take place. After 350uL was added, the mixture was centrifuged for 10 minutes. All the centrifuging in this miniprep stage was at 13,000rpm. The 800 uL gathered from previous steps, was centrifuged for 1 minute, and a vacuum was formed to collect the solution through the QIAprep Spin miniprep column. The column was then rinsed with 0.75 mL buffer PE. The flow-through was discarded after the mixture was centrifuged. The vacuum method was applied to the solution. The solution was then centrifuged for 2 minutes and the column opening was left open to air dry for 5 minutes. The column was placed in a 1.5mL centrifuge tube, so the elution of the DNA can begin. A total of 40uL of TrisBuffer (10mM) was added to the column and the column was left standing for 5 minutes before being centrifuged for 2 minutes. The DNA can now be collected and the concentration can be measured via a microplate spectrophotometer (BioTek Epoch 2). There were 8 sections being measured at 2uL per measurement. The sections were first calibrated with distilled water, before having the sample concentrations measured. Epoch 2 was alerted that the DNA being measured was double-stranded, to ensure a more accurate reading.

Yeast Growth

Once we have our plasmid, the plasmid needed to be introduced to component yeast cells for expression. It was chemically transformed (the cells) with a yeast transformation kit, and then plate the transformants into a selection media plate. The colonies are then grown in the plate and are inoculated into a starter culture in 70 mL YMB liquid media. Approximately 20mL is transferred into the 1 mL culture. Everything should be at 4℃ at the start of this procedure. To first harvest the cells, the yeast culture is added to 1L centrifuge tubes. The centrifuge is used for 25 mins. The supernatant is decanted except for the last 50mL in the tube. The pellet was then resuspended using the leftover supernatant and transferred to a 50mL conical tube. It was spun down for 20 minutes. The solution is then decanted and the mass of the pellet is noted and stored at -80℃.

The buffer for YCF1 should be prepared using 50 mM Tris and 300 mM NaCl. Approximately 3.1mL of the buffer will be mixed in for every 1 gram of cells. The mixture should then be on the ice to thaw. A protease is then added to the buffer solution and 1 tablet is used for every 50mL of the buffer. The cell pellet is then scraped from the inner linings of the tube and added to a chilled mixer. Buffer is then added to the pellet. Fungus disruption beads were added to fill up the mixer. The bead beater is then assembled and the plastic exterior is filled with ice to keep the cells cold at all times. Methanol was used to lower the melting temperature of the ice. The mixer was used in 40-second intervals, with 5 minutes to chill (repeated 7 times). A vacuum filter removed the glass beads as they were rinsed with a buffer. Filtrate was added to ultracentrifuge

tubes and spun down. The supernatant was decanted and the crude membrane's mass was recorded and stored at -80℃.

Solubilization of the membrane was the next step. Three different buffers were created with a mixture of 50 mM Tris-Cl, 300 mM NaCl, imidazole, and HCl. The first buffer is made for 15mL for every gram of Chorismate Mutase. 1x APL is added from the 100x stock. The buffer is then added to 2mM AEBPS to be dissolved. The Chorismate Mutase is then added to a buffer bottle and buffer was used to rinse tubes to get as much of the Chorismate Mutase submerged into the buffer. A stir bar was then added and left to stir for 5 hours. The solubilized membrane was spun down in a 4℃ ultra-centrifuge for 30 minutes. The supernatants were then extracted via a vacuum filter and the final volume of the lysate was recorded. In total, 30 mM imidazole was added to lysate before the loading process.

Nickel Purification

5mL IMAC cartridge was used for the column. Pressure for the column was set at 0.30 and the delta MPa is around 0.02-0.03 MPa. The pressure for the instrument was set to 3MPa. A washing line was pumped with CV buffer, then degassed water. The lysate was then loaded onto the column using a 2mL per minute flow rate. The test tube rack was then loaded in a specified order (Fig.6). The filter is then made wet with a 50 mL 100X Amicon concentrator. The mixture was centrifuged for 5 minutes and the fractions are combined in a 50mL conical tube with buffer. The excess water was

drained and the remaining solution was centrifuged for 20 minutes. A 200uL pipette was then used to transfer the protein to the Costar Spin-X column and spun for 2 minutes.

Size Exclusion

The SEC Superose column was prepped with 30 mL buffer at a 0.5 mL/minute rate. The test tube rack was then loaded with 45 test tubes which will allow for YCF1 to elute from the column. A 1 mL syringe and the needle were then required to inject the sample. This part proceeded with extra caution as if air bubbles arose, they had to be quickly eliminated by pulling back the plunger until the bubbles floated to the top of the column. After all the bubbles were removed, the data was collected. An SDS page can be used to confirm where the protein would elute (Fig.7). The protein was then stored in the buffer at 4 ℃. The protein is then ready to perform further confirmatory experiments, such as EPR or Cryo-EM and biochemical assays, to view in-depth details of the new structures formed.

Figures

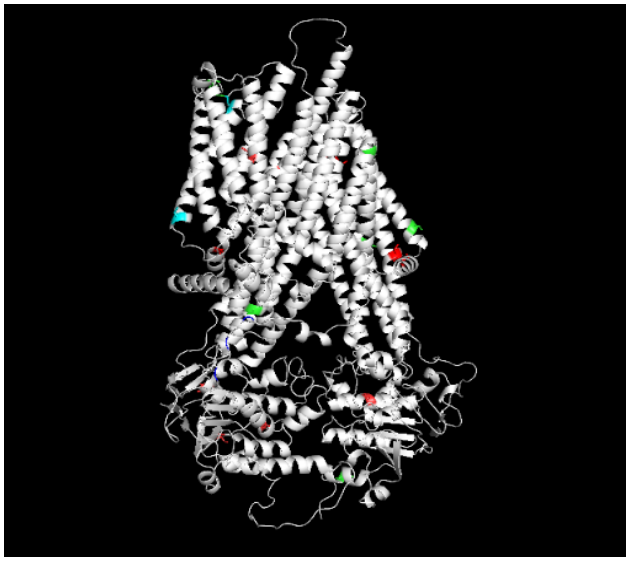


Fig.1. YCF1_83nu has several locations that could be used for cis-less mutations. Solvent-exposed regions are often best to select as they have the range to make the most notable confirmation changes on the exterior of the protein. The red regions are not solvent-exposed. The blue regions were considered to possibly be solvent-exposed. Whereas, the green regions are the selected locations of the cysteines that were substituted.

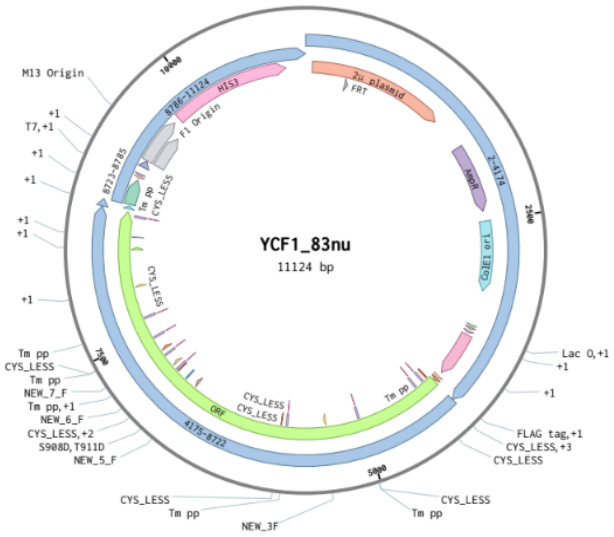


Fig.2. There are 11,124 base pairs in the YCF1_83nu mutant. Multiple crucial parts of the cell allow for the proceeding experiments to be successful. One is the ampicillin resistance gene. Throughout several steps, ampicillin can be used to kill unwanted contamination by inserting this gene. The HIS3 gene codes for an enzyme that accelerates the biosynthesis of histidine. M13 has the role of the production of viruses, especially in replication processes.

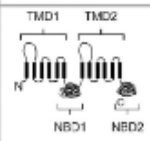
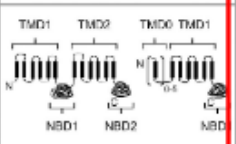
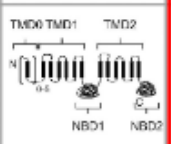
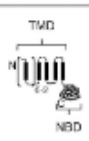
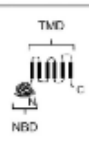
Human ABC-Family	A	B	C	D	G
Transporter Topology					
Lipophilic Substrates	(Phospho)lipids (A1, A3, A4, A7, A12) Sphingomyelin (A1, A3) Cholesterol (A1, A2, A5)	Phospholipids (B1, B4) Sphingolipids (B1) Bile salts (B11) Drugs (B1)	Phospholipids (C1) Bile salts (C1, C2, C3) Steroids (C1, C10) Drugs (C1, C2)	(VLC)FA (D1-D4)	Lipids (G4) Cholesterol (G1, G4, G5/G8) Steroids (G2, G5/G8) Drugs (G2)
Subcellular Localization	Plasma membrane (A1, A4, A7) Lysosome (A2, A5) Lamellar bodies (A3, A12)	Plasma membrane (B1, B4, B11)	Plasma membrane (C1, C2, C3, C10)	Peroxisome (D1-D4)	Plasma membrane (G5/G8) Endosomes (G1, G4)

Fig.3. ABC transporters have subfamilies (A, B, C, D, and G). The topology, substrates, and localization are shown. YCF1 is a part of the ABCC family where the gene associated with the mucus membrane thickening of cystic fibrosis is also found. Adapted from: Hellmich et al. (2016). BBA Biomembranes 1859 (2017) [14].

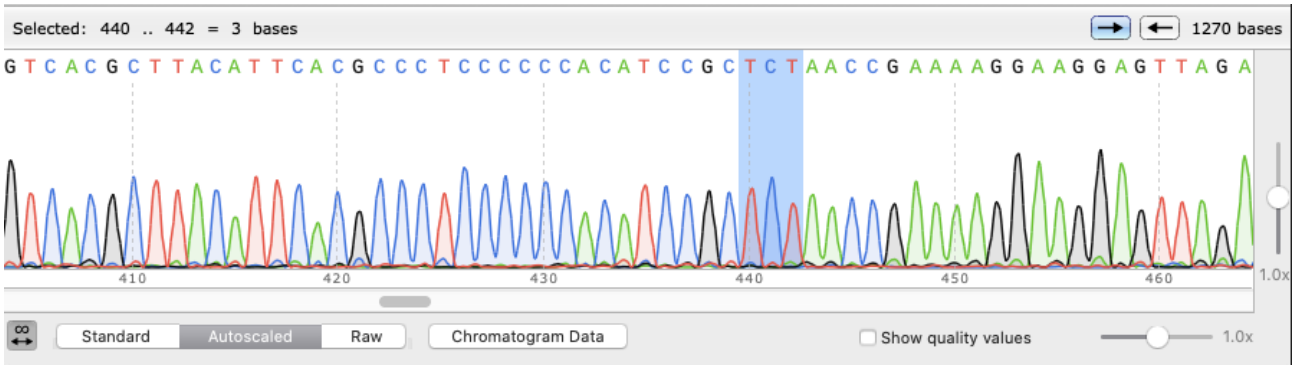


Fig. 4. The figure depicted demonstrates a successful sequencing data set where the plasmid DNA was purified without contaminants, thus able to proceed to future experimentation. The triplet codon TGT that codes for cysteine was successfully replaced by a TCT codon that codes for serine.

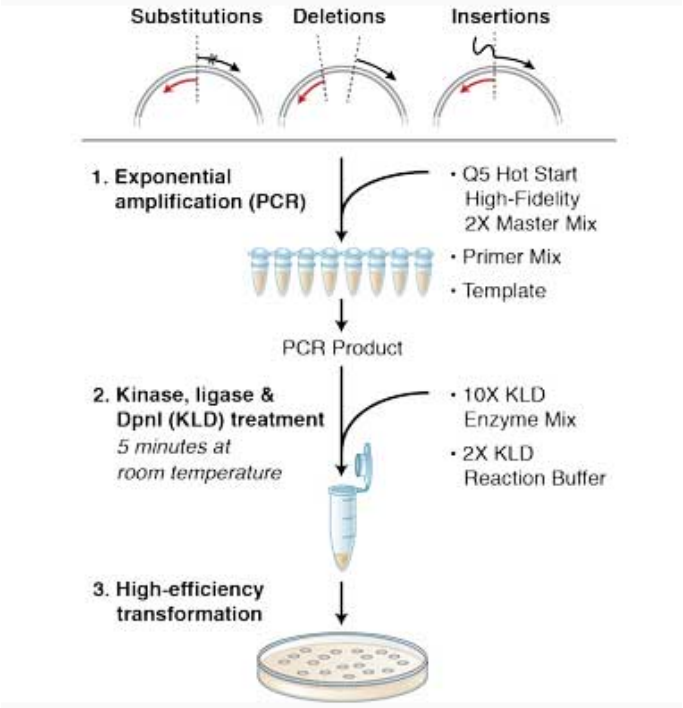


Diagram from: Nebchanger.com

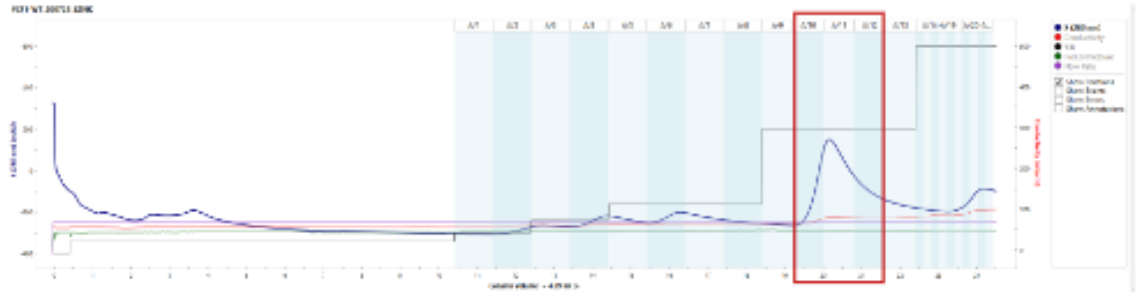
Fig.5. A demonstration of the qrt-PCR with the KLD reaction notes how transformation is performed in its most optimal way. The substitution process allows for more efficient insertion of new material.

10 CV (50 mL)	6% B
2 CV (10 mL)	10% B
2 CV (10 mL)	16% B
4 CV (20 mL)	24% B
4 CV (20 mL)	60% B **elution**
6 CV (30 mL)	100% B

* if eluting manually, make sure to wash pump B before starting run (30 mL, 10 mL/min, bypass)

Fig.6. Columns were used in conjunction with an Affinity IMAC to allow for nickel purification. The percentages on the right-hand side of the table represent the percentage of a buffer that is placed in each lane.

IMAC:



SEC:

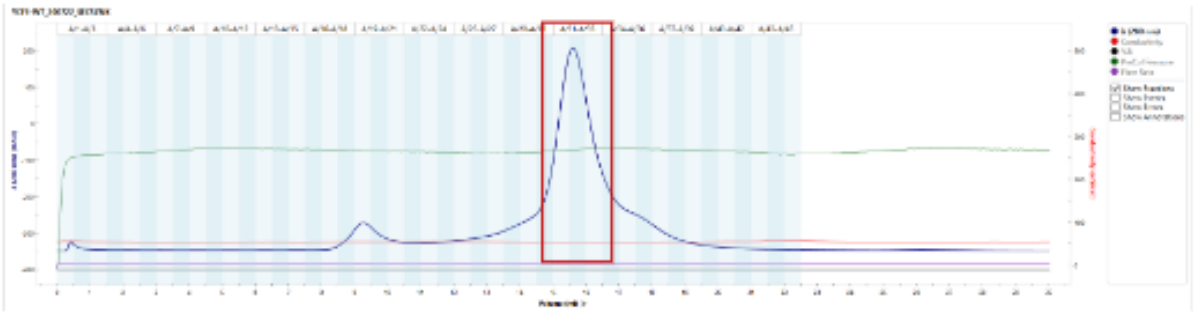


Fig.7.a. Wild-type YCF1 protein Purification Scheme. Immobilized metal affinity chromatography (IMAC), provided a way to separate the various proteins in the solution based on their behavior with transition metals and other proteins found in the solution. The protein of interest eluted in the boxed-in regions in both the IMAC and SEC. The SEC allowed for the proteins to elute in accordance with their size. In which the protein was eluted at the 170 kB region.

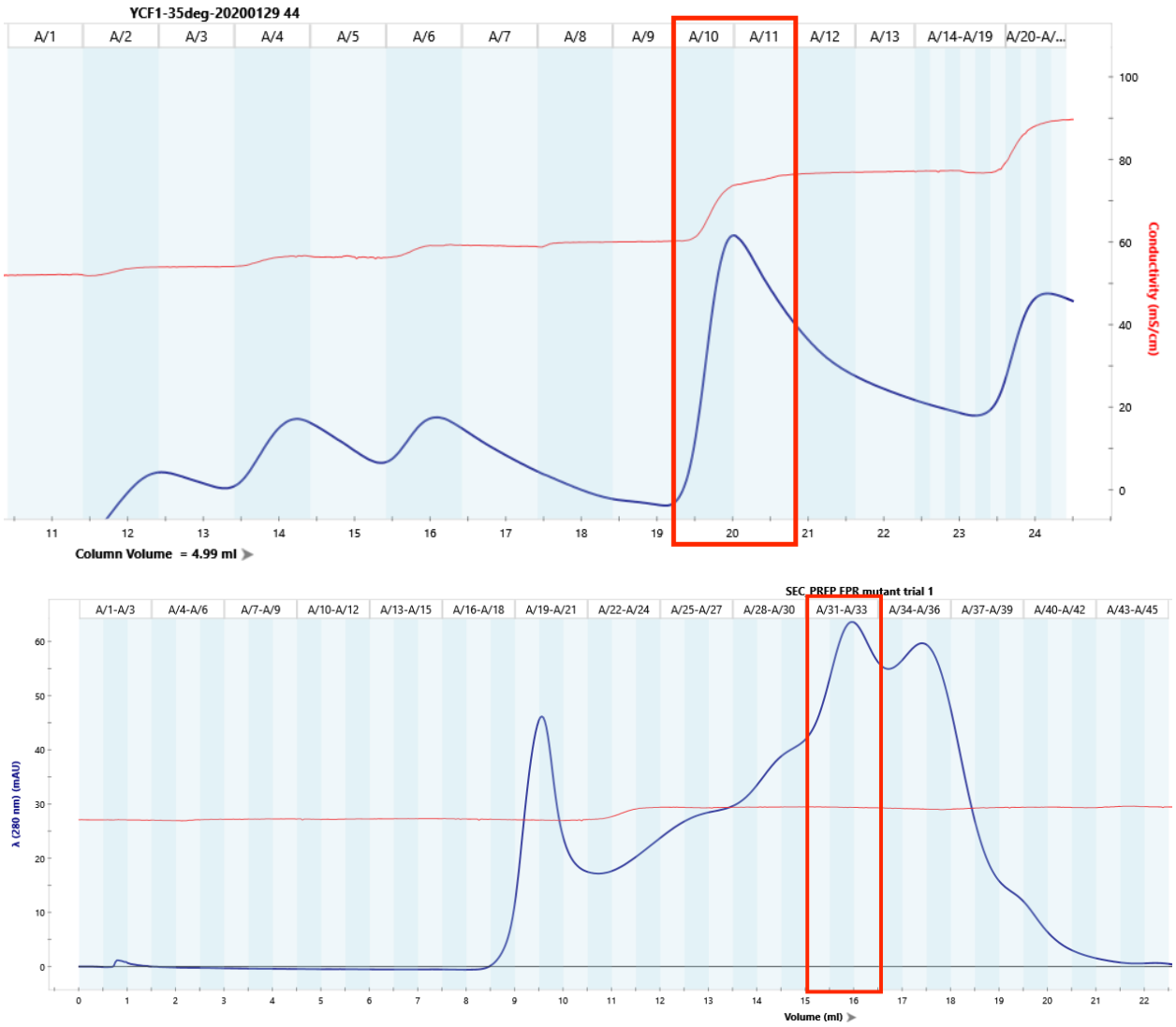


Fig.7.b. Cys-less YCF1 protein Purification Scheme. IMAC purification of the Cys-less YCF1 protein followed by a size exclusion chromatography purification to assure maximum purity of the samples. As observed the YCF1 protein peak came with other protein aggregates and degradation species. This could be due to the mutations that disrupted part of the homogenous population of the protein. Nonetheless, the eluted peaks of around 15-16 mLs represent the well-folded protein that can be further used for validation experiments.

Discussion

The importance of modifying the cysteine is so we can track the cysteine modification sites and to sample the overall EPR probing abilities. At the moment, we were able to find the areas of interest for EPR. In regards to successful outcomes, the structure analysis and plasmid preparations proved to be accurate. The selection of 8 cysteine regions on YCF1 that were modified, was successful. The ampicillin resistance element when first transferring the yeast cells to the agar plate eliminated the wild-type and significantly lowered the probability of contamination. Only 2 rounds of experiments were classified as possibly contaminated, thus requiring over 80% of trials to be purified. As for the primer designs, all solvent-exposed regions were able to have an effective forward and reverse primer annealing that allowed for mutagenesis to be set up correctly. Our sequencing results yielded that yielded, that it is possible to modify specific locations of Ycf1 plasmid and further modified the cysteines in the protein structure

The Q5 and QIA spin mini preps allowed the cells to be prepared for transformation and no contamination issues were found during these steps. Yet, nearly a third of all trials sent off to sequencing revealed inconclusive data and needed to be repeated. The yeast growth was successful in the solubilization of the membrane. And, for the IMAC purification, it was successful in every run. In the end, the DNA was successfully mutated and the protein was expressed and purified.

In terms of purification of this newly designed Cys-less Ycf1 protein, we observed that the purification was not optimal since there were aggregated and degraded species present. However, we are still able to obtain a percentage of the well-folded protein.

This sample can then be further verified by mass spectrometry and thermostability assays to validate for proper folding, size, and biochemical properties.

Once this biochemical validation is further verified we can proceed with the EPR assays by 1) Re-introduction specific cysteine pairs in regions of interest for EPR measurements; 2) Expression and Purification; 3) Measurement of the EPR signal; 4) Data Quantification and Visualization. Ultimately, we should be able to quantify different population species and measure the changes happening during conformational changes in this dynamic protein.

Acknowledgments

I would like to thank Dr. Tomasiak and Rodolpho for guiding me throughout the project and exposing me to various new lab techniques that will be utilized in the future. I also would like to thank the Chemistry and Biochemistry department for funding many of the necessary solvents and equipment required for experiments in the lab.

Citations

1. Kalsotra T, Khullar S, Agnihotri R, Reddy MS. Metal induction of two metallothionein genes in the ectomycorrhizal fungus Suillus himalayensis and their role in metal tolerance. Microbiology (Reading). 2018 Jun;164(6):868-876. doi: 10.1099/mic.0.000666. Epub 2018 May 15. PMID: 29762106.
2. Gueldry, O. et al. Ycf1p-dependent Hg(II) detoxification in Saccharomyces cerevisiae. Eur. J. Biochem. 270, 2486–2496 (2003).
3. Li, Z. S. et al. A new pathway for vacuolar cadmium sequestration in Saccharomyces cerevisiae: YCF1-catalyzed transport of bis(glutathionato) cadmium. Proc. Natl Acad. Sci. USA 94, 42–47 (1997).
4. Szczypka, M. S., Wemmie, J. A., Moye-Rowley, W. S. & Thiele, D. J. A yeast metal resistance protein similar to human cystic fibrosis transmembrane conductance regulator (CFTR) and multidrug resistance-associated protein. J. Biol. Chem. 269, 22853–22857 (1994).
5. Liu, F. et al. Molecular structure of the human CFTR ion channel. Cell 169, 85–95 (2017).
6. Bickers, S. C., Benlekbir, S., Rubinstein, J. L. & Kanelis, V. Structure of Ycf1p reveals the transmembrane domain TMD0 and the regulatory region of ABCC transporters. Proc. Natl Acad. Sci. USA 118, e2025853118 (2021).
7. Marks, D. S., Hopf, T. A. & Sander, C. Protein structure prediction from sequence variation. Nat. Biotechnol. 30, 1072–1080 (2012).
8. Naren, A. P. et al. CFTR chloride channel regulation by an interdomain interaction. Science 286, 544–548 (1999).

9. Winter, M. C. & Welsh, M. J. Stimulation of CFTR activity by its phosphorylated R domain. Nature 389, 294–296 (1997)

10. Dong, W., Xu, C., Li, C. *et al. ycf1*, the most promising plastid DNA barcode of land plants. *Sci Rep* 5, 8348 (2015).

11. Amar, M. H. (2020). Ycf1-ndhF genes, the most promising plastid genomic barcode, sheds light on phylogeny at low taxonomic levels in Prunus persica. *Journal of Genetic Engineering & Biotechnology*, *18*.

12. Khandelwal, N.K., Millan, C.R., Zangari, S.I. *et al.* The structural basis for regulation of the glutathione transporter Ycf1 by regulatory domain phosphorylation. *Nat Commun* 13, 1278 (2022).

13. Tusnády GE, Bakos E, Váradi A, Sarkadi B. Membrane topology distinguishes a subfamily of the ATP-binding cassette (ABC) transporters. FEBS Lett. 1997 Jan 27;402(1):1-3. doi: 10.1016/s0014-5793(96)01478-0. PMID: 9013845.

14. Jennifer Neumann, Dania Rose-Sperling, Ute A. Hellmich. Diverse relations between ABC transporters and lipids: An overview. Biochimica et Biophysica Acta (BBA) - Biomembranes,Volume 1859, Issue 4, 2017,605-618.