Spring 5-2013

Cloning and Expression of the cbbO (0910) Gene from *Halothiobacillus neapolitanus* and Its Potential to Code for RubisCO Activase

Emily F.E. Bustin
University of Southern Mississippi

The University of Southern Mississippi

Cloning and Expression of the *cbbO* (0910) Gene from *Halothiobacillus neapolitanus* and Its Potential to Code for RubisCO Activase

by

Emily Bustin

A Thesis
Submitted to the Honors College of
The University of Southern Mississippi
in Partial Fulfillment
of the Requirements for the Degree of
Bachelor of Science
in the Department of Chemistry & Biochemistry

May 2013

Approved by

_____

Dr. Sabine Heinhorst
Professor of Chemistry & Biochemistry

_____

Dr. Sabine Heinhorst, Chair
Department of Chemistry and Biochemistry

_____

David R. Davies, Dean
Honors College

AKNOWLEDGEMENTS

I would like to thank the Heinhorst/Cannon research lab group, who have taught me much about molecular research and become great friends besides. I would like to thank Avi Biswas, for his unending patience and understanding in teaching a “12-year old” how to achieve lasting results. Lastly, I would like to thank Dr. Sabine Heinhorst for being a wonderful mentor and for always allowing me the opportunity to expand my horizons.

ABSTRACT

Ribulose-1,5-bisphosphate carboxylase/oxygenase or RubisCO is an enzyme found in autotrophic organisms that functions to fix CO_2 and therefore plays an important role in the global carbon cycle. In order for RubisCO to increase its slow enzymatic rate, it must be exposed to high amounts of its substrate, CO_2. Carboxysomes, which sequester the RubisCO and its substrate, provide this function. It was previously believed that all the components necessary for carboxysome function were encoded by genes within the traditional *cso* operon. Recently, however, a gene in an operon located downstream of the *cso* operon was found to encode the novel shell protein CsoS1D. This discovery raised the possibility that other genes located outside of the traditional *cso* operon may contribute to the structure or function of the carboxysome. One such gene is *cbbO*, which encodes a potential RubisCO activase. A RubisCO activase could play a role in maintaining RubisCO's catalytic efficiency within the carboxysome. To study the *cbbO* gene and its potential in greater detail, its gene first had to be over-expressed so that sufficient amounts of soluble recombinant CbbO protein could be purified for the generation of polyclonal antibodies. To achieve this goal, the *cbbO* gene of *Halothiobacillus neapolitanus* was PCR amplified and inserted into a plasmid vector. The genomic DNA was taken from the model organism for the study of carboxysomes, *Halothiobacillus neapolitanus*. The recombinant construct was sent for sequence determination. Then the *cbbO* gene fragment with the correct DNA sequence was ligated into a protein expression vector. Recombinant CbbO protein, which has a hexa-histidine affinity tag that is encoded on the pETDUET-1 vector, was purified by affinity chromatography and quantified and analyzed using SDS page. Finally, the recombinant CbbO protein was sent off for antibody generation in rabbits and initial bleeds were analyzed for antigen specificity using immunoblotting. The preliminary results seemed to indicate that an antibody probe for recombinant CbbO protein was obtained.

TABLE OF CONTENTS

REVIEW OF RELATED LITERATURE

For a long time, CO_2 fixation into energy-rich molecules by autotrophs, organisms that obtain their carbon from inorganic molecules, has fascinated scientists. In the infancy of the study of metabolic pathways, it was known that some organisms required CO_2 as a carbon source, and that the conversion of 1 mol of CO_2 to 2 mols of 3-phosphoglyceric acid (3-PGA) was only accomplished with the help of the enzyme, ribulose-1,5-bisphosphate carboxylase/oxygenase, or RubisCO (Shively 1998). Autotrophs being the primary producers in the global biosphere, a deeper understanding of this enzyme was required. As it turns out, RubisCO is the most abundant protein on Earth. It is also an extremely slow enzyme and only produces three to ten 3-PGA molecules per second per active site (Feller 2008). Therefore, at normal intracellular concentrations of CO_2 and ribulose-1,5-bisphosphate (RubP), the other substrate of RubisCO, the activity of RubisCO remains relatively low. However, as with all enzymes, once the concentration of the limiting substrate, in this case CO_2, is increased, the catalytic rate of RubisCO also increases, and allows it to fix more CO_2 per unit of time. For some autotrophic bacteria to successfully increase the concentration of CO_2 and increase the enzymatic rate of RubisCO, it seemed likely that some kind of CO_2 concentrating mechanism must occur (reviewed in Shively 1998 & Heinhorst 2010). In 1956 a particular type of polyhedral inclusion bodies or bacterial micro-compartments were discovered in a cyanobacterium (Figure 1). Their cellular role was at first unknown. Since then the inclusions have been determined to be made of an outer thin (3-4 nm) protein shell and to function by sequestering large amounts of the enzyme RubisCO within their core (Figure 2). Once this property was discovered, they were called carboxysomes (Shively 1998). Bacterial micro-compartments have been thought by Yeates et. al. to be the highest possible level of sub-cellular organization for bacteria (2008). It has also been determined that regardless of the presence of RubisCO, the carboxysome shell is still assembled and functions normally (Heinhorst 2010).

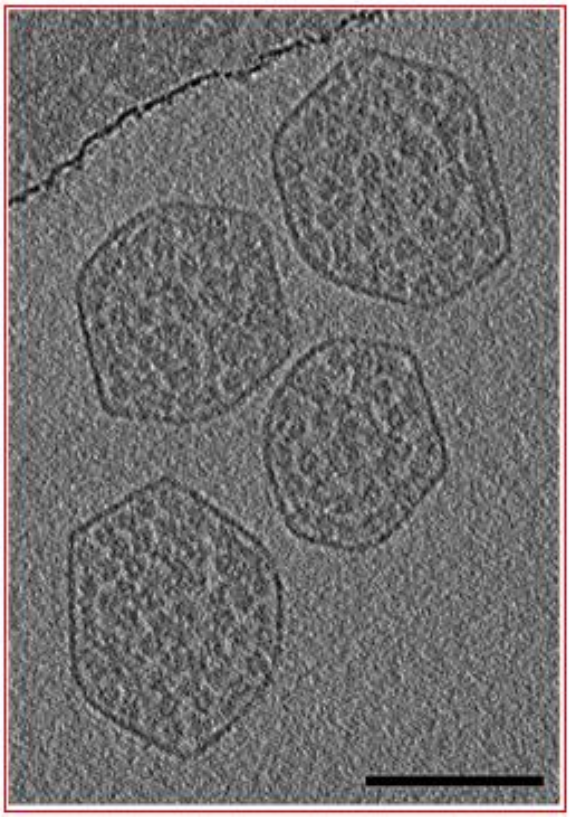


Figure 1: A cryo-electron tomogram of isolated *H. neapolitanus* carboxysomes filled with RubisCO.*Source: C. Iancu & G. Jensen, California Institute of Technology*

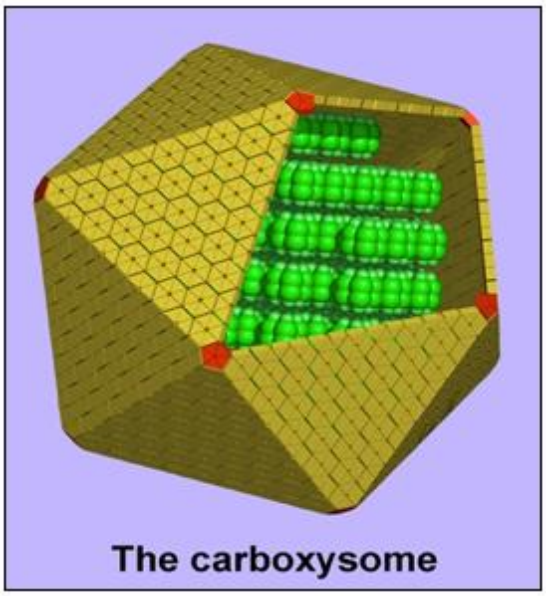


Figure 2: An artist rendering of a carboxysome shell filled with RubisCO holoenzyme molecules. *Source: T.O. Yeates, University of California, Los Angeles*

The gene region in autotrophs that codes for the carboxysome proteins is called the *cso o*peron. Traditionally, this has been considered the only operon responsible for encoding the structural and functional carboxysome components. Downstream of this operon, however, resides another operon that has recently been studied and is now known to encode at least one additional necessary carboxysome component, the novel shell protein CsoS1D (Roberts 2012). In between the traditional *cso* operon and the gene for CsoS1D is the *cbbO* gene, which is annotated as encoding a Von Willebrand Factor-like protein that may function as a RubisCO activase or perhaps a chaperone that guides RubisCO to the carboxysome interior (1997 Hayashi). According to Hayashi, the CbbO protein has been shown to enhance RubisCO activity and stability in *Pseudomonas hydrogenothermophila* (1997 Hayashi).

Shively et. al (1998) expressed the need for a greater understanding of genomic function, with regards to the carboxysome, outside of the *cso* operon. Thus, the necessity to study the *cbbO* gene downstream of the traditional operon is apparent. The hypothesis for this research project is that the CbbO protein can be expressed as a soluble recombinant protein in amounts large enough for the generation of polyclonal antiserum in rabbits. The recombinant CbbO protein can be used to study protein-protein interactions in conjunction with the large and small subunits of RubisCO and other partner proteins, including CbbQ, another potential RubisCO activase. Once the antibody is generated and confirmed via Western Blot analysis, it can be used to identify if the CbbO protein is present in carboxysomes or if CbbO is involved in carboxysomal structure or function.

As seen in Figure 3 and as reported by Shively et al. (1998) the carboxysome gene regions are very similar in several different chemoautotrophic species.

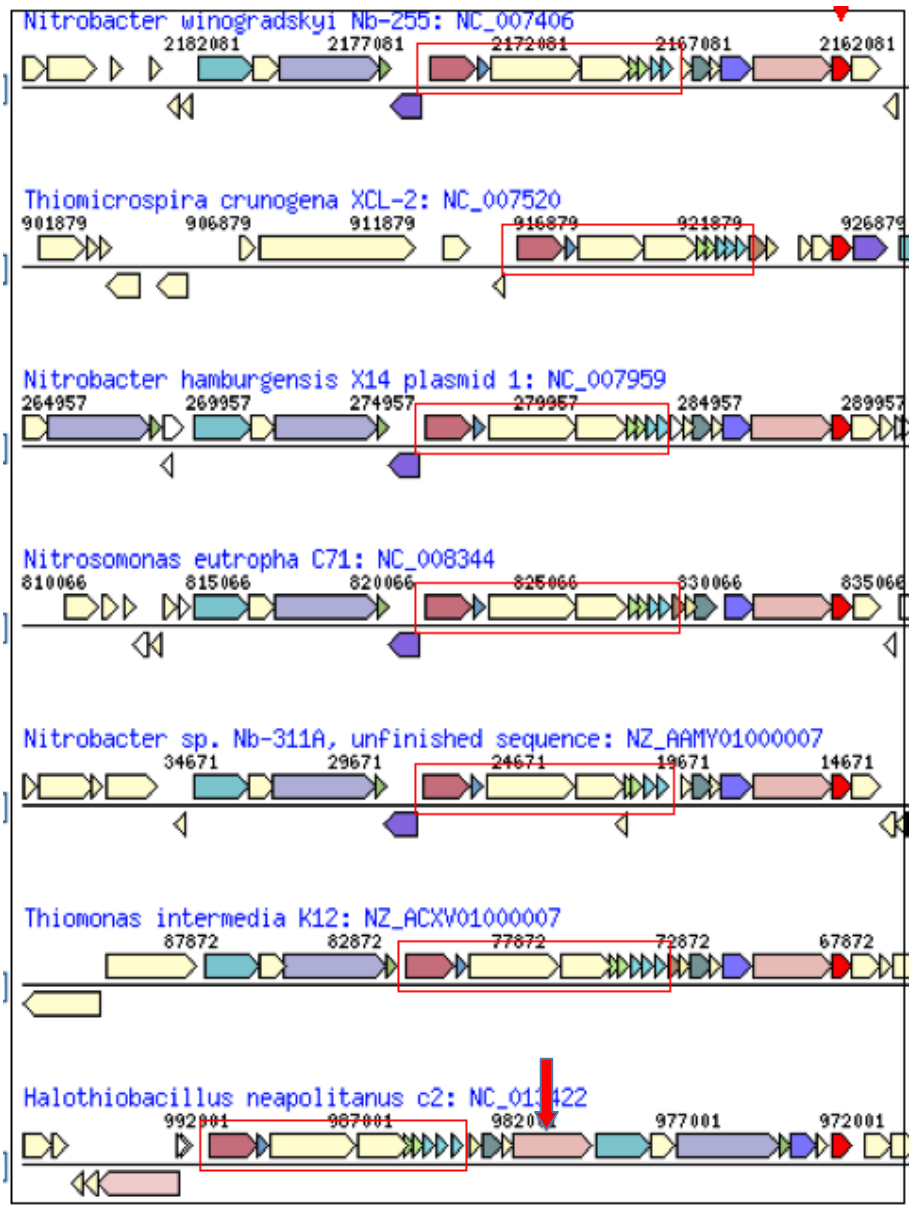
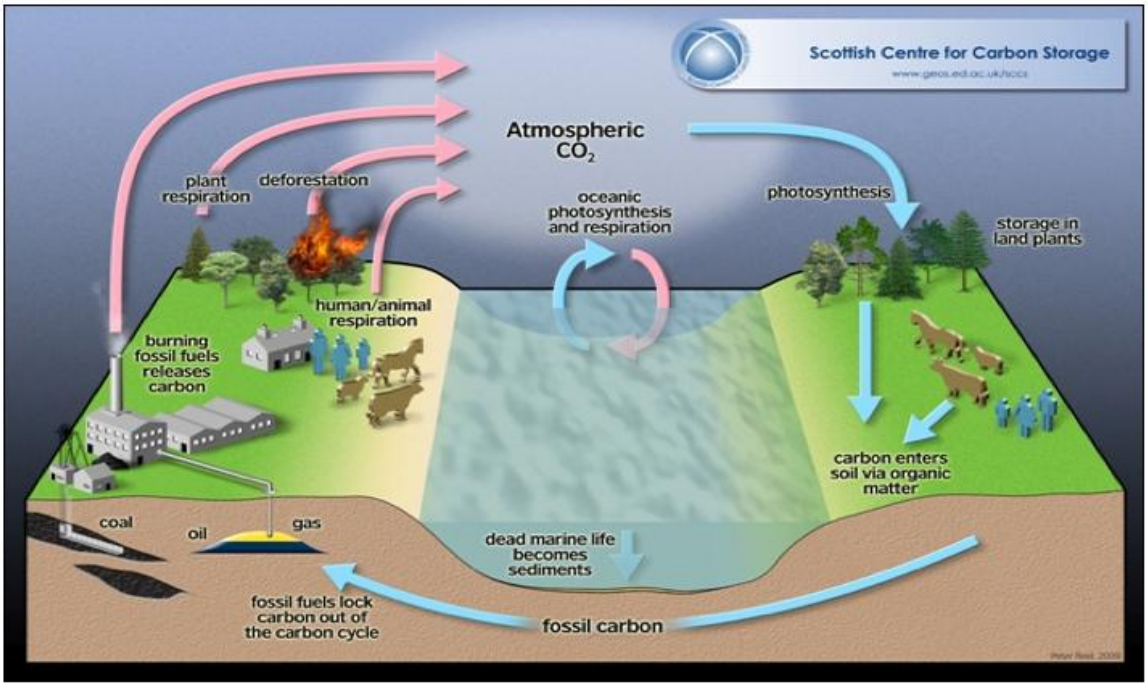


Figure 3: Gene annotation denoting the traditional *cso* carboxysome operon (in red box) for various chemoautotrophs, and the *cbbO* gene (pink, marked by a downward red arrow) thought to code for a potential RubisCO activase.

The findings of this study have important implications for biotechnological applications. One would be the manufacture of synthetic carboxysomes. These nano-cages could be used as an advanced biochemical approach to a more selective drug delivery system. Another would be the use of recombinant DNA technology to engineer more efficient means of CO_2 fixation, thereby possibly decreasing the quantities of

CO_2, a contributing green-house gas, from the atmosphere by bacteria like *H. neapolitanus*. This effect is possible due to the CO_2 fixation function carried out by carboxysomes in autotrophs found in Earth's vast oceans and water ways as seen in Figure 4 (Shively 1998).

For this study the sulfur oxidizing bacterium *Halothiobacillus neapolitanus*, an obligate autotroph will be used as the genomic model organism. This bacterium has been widely studied and its genome has been sequenced. Additionally, carboxysomes from this organism can be purified and their composition and function studied *in vitro* (Heinhorst 2010).



Source: Scottish Centre for Carbon Capture and Storage, University of Edinburgh

Figure 4: The global carbon cycle, showing sources of atmospheric CO_2 and processes that remove the greenhouse gas from the atmosphere.

If a better understanding of carboxysome structure and function can be obtained, a greater understanding of the *in vivo* micro-compartment assembly pathway may also be accomplished (Yeates 2008). The findings could also give further evidence to the mechanism by which RubisCO obtains a kinetic advantage through being sequestered within the carboxysome.

It has been hypothesized that the *cbbO* gene is a component of another operon that helps regulate CO_2 fixation in conjunction with the traditional *cso* operon (Hayashi, 1999). The overall aim of this project will be to express the *cbbO* gene as soluble, recombinant CbbO protein that will enable future researchers to develop an antibody probe for the presence of CbbO in the cell and in carboxysomes. In order to study the *cbbO* gene, the gene will be amplified from the model organism *H.neapolitanus*, a gram-negative bacterium that was the first chemoautotroph from which purified carboxysomes were obtained (Cannon 2001). Carboxysomes within *H. neapolitanus* can be seen in Figure 5. Other autotrophic organisms have been found to contain the *cbbO* gene. However, *H. neapolitanus* will be used since its carboxysomes have already been thoroughly studied (Cannon, et. al. 1983). The gene of interest inserted in the protein expression vector can be found in Figure 8.

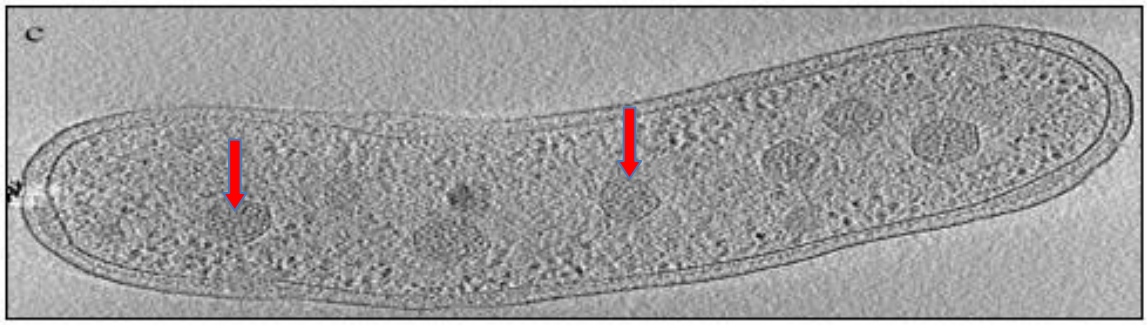


Figure 5: Cryo-electron tomogram showing an *H. neapolitanus* cell containing carboxysomes (indicated by the downward facing red arrows) that are filled with RubisCO.

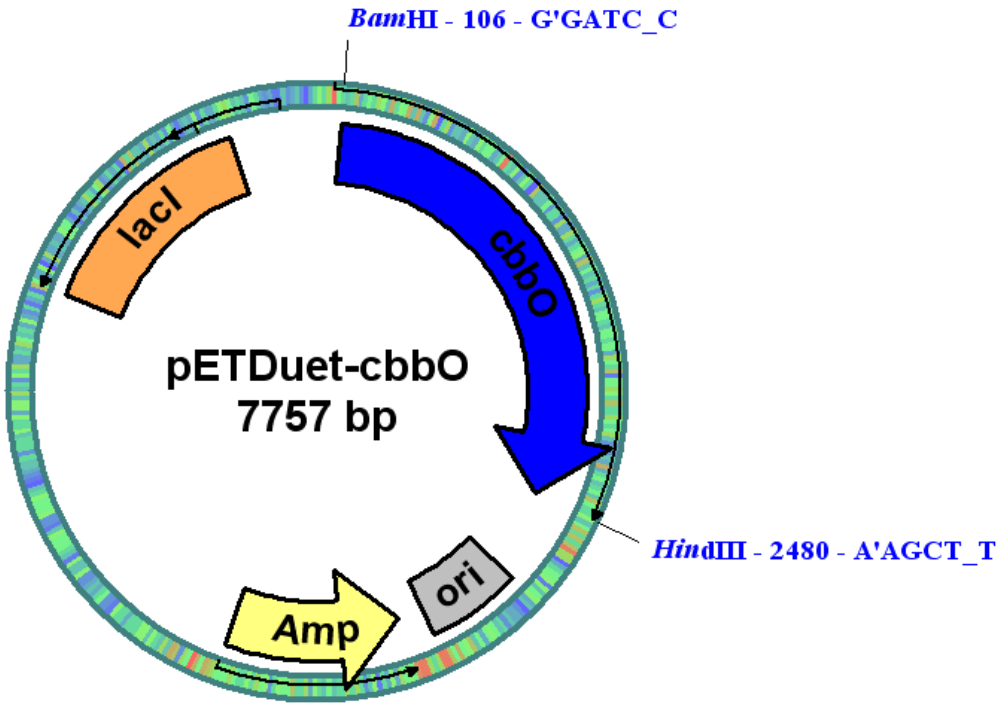


Figure 6: Gene map of protein expression vector with inserted *cbbO* gene of interest and restriction enzyme sites.

EXPERIMENTAL PROCEDURES

Materials

Media

Luria Bertani (LB) broth

10 g/L Bacto tryptone

5g/L Bacto yeast extract

10 g/L NaCl

Autoclave for 20 minutes. at 121°C and 15 psi

For solid LB medium, add 1.5% w/v agar before autoclaving

Super Optimal with Catabolite Repression (SOC) medium

20 g/L Bacto tryptone

5 g/L Bacto yeast extract

0.58 g/L NaCl

0.19 g/L KCl

0.95 g/L $MgCl_2$

3.6 g/L glucose

Dyes

Ethidium Bromide stock solution

0.02 g/mL Ethidium Bromide in deionized water

Agarose gel tracking dye (10 X)

100 mM EDTA (pH 8.0)

50% (v/v) glycerol

1% (w/v) SDS

0.1% (w/v) bromophenol blue

Buffers

TAE buffer pH 8.0

40 mM Tris-HCl (pH 7.8)

20 mM Na-acetate

2 mM EDTA

PMSF/PTSF stock solution

100 mM phenylmethylsulfonylfluoride (PMSF)

100 mM p-toluenesulfonylfluoride (PTSF)

This solution was made with 100% ethanol

Lysis buffer, pH 8.0

50 mM Tris-HCl (pH 8.0)

5 mM β-mercaptoethanol (BME)

1 mM PMSF/PTSF

Phosphate buffered saline (PBS), pH 7.4

137 mM NaCl

2.7 mM KCl

4.3 mM Na_2HPO_4

1.4 mM KH_2PO_4

Colony Lysis Buffer

50 mM Tris-HCl, pH 8.0

10 mM EDTA

100 ug/mL RNase A

*Storage condition: 4 °C after adding RNase A

Affinity Chromatography Buffers

pProEx wash buffer A/ column equilibration buffer

20 mM Tris-HCl (pH 8.0)

100 mM KCl

10 % (v/v) glycerol

20 mM imidazole

pProEX wash buffer B

20 mM Tris-HCl (pH 8.0)

20 mM KCl

10% (v/v) glycerol

Elution buffer

20 mM Tris-HCl (pH 8.0)

100 mM KCl

10% (v/v) glycerol

400 mM imidazole

Protein dialysis buffer for recombinant proteins, pH 8.0

10 mM Tris HCl (pH 8.0)

400 μM PMSF/PTSF

Protein Gel Electrophoresis

SDS-PAGE running buffer (Laemmli buffer), pH 8.5

25 mM Tris

192 mM glycine

1% w/v SDS

4X SDS-PAGE loading buffer

200 mM Tris HCl (pH 6.8)

40% v/v glycerol

8% w/v SDS

10 % v/v 2-mercaptoethanol

Sterile deionized water to adjust final volume

Western blot transfer buffer, pH 7.4

25 mM Tris

192 mM glycine

20% v/v methanol

Store at 4°C

Immunoblot blocking buffer

5% non-fat dry milk dissolved in PBS buffer containing 0.1% v/v Triton X-100

Buffer always prepared fresh

Antibiotic Solutions

Stock: 100 mg/mL ampicillin prepared in deionized water and filter sterilized

Final working concentration: 100 μg/mL

Stock: 50 mg/mL kanamycin prepared in deionized water and filter sterilized

Final working concentration: 50 μg/mL

Plasmids and E. coli Strains

New England Biolabs (NEB) 5-alpha Competent *Escherichia coli (E. coli*)– for cloning work

NEB BL21(DE3) Competent *Escherichia coli (E. coli)* –for protein expression

pCR 4Blunt-TOPO vector (Invitrogen)- for confirming the sequence of PCR product

pET DUET-1 vector (Novagen) – for protein expression

Methods

Polymerase chain reaction (PCR)

Typically, 20 uL PCR reactions were set up in 0.2 ml thin-walled PCR tubes. Each reaction contained a high-fidelity DNA polymerase and PCR buffer both from Invitrogen, 250 uM of each dNTP, 200 nM each of forward primer [Von FacA: BamHI:F, 5'-GGATCCGATGAACCCAGCGACTGAA, Tm: 62.8 °C] and reverse primer [Von FacA: HindIII:R, 5'-AAGCTTCTATCGCGTCATCGACAAAT, Tm: 58.8 °C], 100 ng of genomic DNA purified from the model organism as template, and 2 units of DNA polymerase. The Bio-Rad MyCycler thermal cycler was pre-heated to 95 °C before inserting the PCR tubes. A typical PCR protocol consisted of an initial denaturation step at 95 °C for 3 minutes, an initial 5 repeat cycles of denaturation (95 °C for 30 s), annealing (47.2 °C for 30 s) and extension (72 °C for 2.5 minutes), and 25 repeat cycles of denaturation (95 °C for 30 s), annealing (52.2 °C for 30 s) and extension (72 °C for 2.5

minutes). These cycles were followed by a final extension step at 72 °C for 10 minutes, and a hold step at 4 °C.

TOPO Cloning & Transformation

The DNA amplification product was inserted into a Zero Blunt TOPO Vector using the Invitrogen Zero Blunt TOPO PCR Cloning Kit. All protocols were taken from the Invitrogen user manual. In a PCR tube, 10 μL of PCR product was combined with 1 μL of salt solution and 1 μL of the TOPO Vector, both taken from the Invitrogen kit. The solution was gently mixed with a pipette, then incubated undisturbed for 30 minutes at room temperature. After incubation, 100 uL of chemically competent NEB α (NEB) *E. coli* were gently added to the TOPO reaction and incubated for 30 minutes on ice. The *E. coli* cells were then subjected to heat shock at 42° C for one minute, incubated on ice for one minute, then incubated for one hour with shaking at 37 °C in 250 μL of SOC medium to allow for recovery. The *E. coli* containing the TOPO vector were able to grow on the media plates containing kanamycin since the gene encoding a kanamycin resistance protein was present in the TOPO vector. Individual colonies were streaked onto similar LB plates containing the appropriate amount of kanamycin for short-term storage at 4 °C. Liquid overnight culture of 5 mL of LB media and 50 μg/mL of kanamycin were inoculated with samples from selected streak plates. The plasmids were then isolated from the bacterial cultures using the Qiagen Plasmid Miniprep Kit. The purified DNA was quantified using NanoDrop to determine the exact concentration along with approximate purity of the purified DNA sample based on the ratio of the double-stranded DNA to the protein contamination: 260/280 nm (Fisher Scientific).

Agarose Gel Electrophoresis

To separate and analyze the nucleic acids, the technique of agarose gel electrophoresis was used. The 0.8% agarose gels were prepared in TBE buffer. A 1 kb dsDNA marker (NEB) and 6X loading buffer (2.5% Ficoll, 11 mM EDTA, 3.3 mM Tris-HCl pH 8.0, 0.017% SDS, 0.015% Bromophenol Blue) were purchased from New England Biolabs Inc. Loading buffer was added to samples and the gels were subjected to 100 volts until the bromophenol blue tracking dye migrated approximately ¾ down the length of the gel. The GELS were stained by using 0.5 μg/ mL ethidium bromide for ten minutes, and then destained in distilled

water for an additional ten minutes. The nucleic acids were visualized under UV light and analyzed using gel analysis software, VersaDoc (BioRad Laboratories, Hercules, CA).

Restriction Enzyme Digestion of DNA

Restriction enzyme digestion of plasmid DNA was carried out in a 40 µL reaction. Each reaction contained 20 µL of pET DUET-1 vector (26 ng/µL), 2 µL each of 10x NEB Buffer 2 and NEB Buffer 3, 12 µL of sterile water, 1 µL of HindIII and 1 µL of BamHI (both from New England Biolabs). The two restriction enzymes were the last components added to the reaction mixture. The 40 µL mixture was placed in a 37 °C water bath overnight before gel electrophoresis. Agarose gel electrophoresis was then used to confirm the presence of the *cbbO* gene. After analysis with restriction digestion, the TOPO plasmid containing the *cbbO* gene was sent to Eurofins MWG Operon Inc. for sequence determination of the inserted DNA.

DNA recovery from agarose gels

To recover the DNA from a gel, the band of interest was excised using a clean razor blade, and the excised fragment was transferred to a 1.5 mL microcentrifuge tube. The GeneClean Gel DNA Recovery Kit Protocol was used to purify the DNA from the gel fragment. The gel piece was dissolved in a volume of NaI solution corresponding to three times its weight. The gel slice and NaI solution were incubated to melt the agarose. The reaction tube was placed in a 45 °C -55 °C water bath for 1 minute. The contents of the tube were mixed by gently pipetting up and down. The incubation was continued until all the agarose was dissolved. The amount of GLASSMILK© needed was calculated and added to the sample and mixed gently using a pipette. The sample was then incubated at room temperature for 5 minutes. Then the GLASSMILK© with the bound DNA was pelleted by centrifugation at 14,000 x g for 5 s. The supernatant was discarded. Five hundred µL prepared NEW Wash was added and the pellet resuspended. The sample was centrifuged again at 14,000 x g for 5 s and the supernatant once again discarded. This wash step was repeated once. The pellet was allowed to dry and a volume of water equal to that of the amount of GLASSMILK© initially added was used to resuspend the pellet. The sample was centrifuged at 14,000 x g for 30 s. The supernatant containing the DNA was carefully separated by pipetting into a fresh tube.

DNA Ligation reaction

For ligation of the *cbbO* gene into the pET DUET-1 vector (Novagen), a 10 µL reaction volume was used. The reaction mixture included 1 µL of pET DUET-1 expression vector (26 ng/µL) linearized with the restriction enzymes Bam HI and Hind III, 2 µL of 10x T4 DNA Ligase Reaction Buffer (NEB), and 1 µL of T4 DNA ligase (NEB). Ligation reactions were incubated at 16 °C overnight and used to transform chemically competent NEB α *E. coli* cells as described previously.

Colony Lysis of Transformants and Restriction Enzyme Digestion

Following transformation, 30 well-separated *E. coli* colonies that had grown on selective, ampicillin-containing medium were re-streaked onto a patch plate and allowed to grow overnight. Using a small, sterile disposable pipette tip, a portion of each patch was transferred to 3 µL of water in a 0.5 mL microfuge tube, and the bacteria were re-suspended. Then 8 µL of Colony Lysis Buffer (pH 8.0) were added to each microfuge tube and mixed by vortexing the sample briefly. Each sample was heated at 100 °C for 30 s and allowed to cool to room temperature. Then a 4 µL reaction mixture containing 1.4 µL of 10X NEB Buffers 2 and 3 (1:1 ratio; pH 7.9), 0.2 µL of 1 mg/mL Bovine Serum Albumin (BSA), 0.5 µL of BamHI, and 0.5 µL of HindIII was added to each sample. The samples were placed in a water bath and incubated at 37 °C for 30 minutes. The cellular debris of each sample was pelleted by centrifugation at 14,000 rpm for 1 minute, and 2.5 µL of agarose gel tracking dye (pH 8.0) were mixed with each sample supernatant before loading onto an agarose gel for electrophoresis. Following electrophoresis, the gel was allowed to stain with ethidium bromide for 10 minutes. The DNA was visualized using the VersaDoc System (BIO-RAD). The colonies that showed evidence of an insert were targeted for plasmid DNA preparation and restriction digestion.

Small Scale Protein Expression

Two-milliliter overnight cultures of *E. coli* transformants that contained the pET DUET-1 plasmid with the appropriate insert were grown in Luria-Bertani (LB) broth with 100 µg/mL of ampicillin. Aliquots (0.5 mL) of these cultures were used as an inoculum for 50 mL working cultures, which were incubated at 37 °C

with agitation for 2.5 hours. Once the cultures reached an optical density of 0.6 to 0.8 at a wavelength of 600 nm, 1-mL aliquots were removed and centrifuged at 10,000 X g for 3 minutes. The supernatants were discarded and the pellets were placed on ice (un-induced samples). To the remaining cultures isopropyl-β-D-thiogalactoside (IPTG) was added (0.3 mL) to a final concentration of 0.6 mM. The cultures were incubated at 37 °C for 5 hours or at 25 °C overnight, before 1 mL samples were removed and centrifuged as before (induced samples). The induced and un-induced samples were re-suspended in 100 µL of PBS buffer (pH 7.4). For sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis, 5 µL of the suspensions were mixed with 5 µL of 4X Laemmli Loading Buffer, and 10 µL of distilled water. The mixture was boiled for 10 minutes and loaded onto a vertical 12% Laemmli SDS-polyacrylamide gel. The samples were subjected to electrophoresis at 100 V for 1.5 hours. The gels were washed in distilled water for 30 minutes, stained with Gel-Code Blue (Fisher Scientific) overnight and destained in sterile water to visualize the polypeptide bands. Induced and un-induced samples were analyzed using SDS-PAGE.

Large Scale Protein Expression

Protocols for induction and purification of the HT-CbbO protein were taken from Life Technologies pProEX HT Prokaryotic Expression System Protocol. Based on the small-scale expressions, one of the *E. coli* cultures found to express recombinant CbbO was prepared for protein production by adding 15 mL of the overnight starter culture to 1 L of LB media along with the appropriate amount of ampicillin antibiotic. The culture was shaken at 37 °C for four hours until the optical density at 600 nm (OD_{600}) was approximately 0.6. Production of T7 RNA polymerase and transcription of the *cbbO* gene were induced by addition of 0.5 mM IPTG. Cells were incubated with shaking at 225 rpm at 30 °C for another 4 hours before they were harvested by centrifugation at 10,000 x g for 10 minutes. Cell pellets were stored at -20 °C until they were used. The bacterial pellet was then re-suspended in 40 mL of Lysis buffer and lysed by three passages through a chilled French pressure cell at 138 MPa. The lysed cell suspension was centrifuged for 10 minutes at 10,000 x g to remove the cell debris, and unbroken cells. The crude extract, supernatant, and pellet were analyzed by SDS-PAGE, using 5 uL of each suspension.

Affinity Chromatography

Using the protocol taken from Life Technologies, the resulting 40 mL of supernatant was loaded onto 4 mL of Ni-NTA resin equilibrated with wash buffer A and 0.5 mM PMSP/PTSF and agitated on a rotary shaker for 24 hours at 4 °C to maximize binding. The supernatant was then allowed to flow through the Ni-NTA column and the resulting resin was washed with 40 mL of wash buffer A containing 20 mM imidazole, followed by 20 mL of wash buffer B without imidazole, and finally 20 mL of wash buffer A containing 20 mM imidazole to remove contaminating proteins with non-specific binding. The CbbO protein was then eluted from the column using elution buffer containing 300 mM imidazole. The purified protein samples were dialyzed against 10 mM Tris HCl at pH 8.0 to remove the imidazole and stored at -20 °C for further use in producing antibodies. Then, SDS-PAGE was used to verify the presence of the protein.

Estimation of Protein Concentration

A BCA protein assay (Pierce, Rockford, IL) was performed using a Beckman Coulter DU 800 Spectrophotometer (Beckman Coulter, Fullerton, CA). Bovine Serum Albumin (BSA) standards were prepared using 1-40 μg total protein per sample. Samples contained protein and distilled water to a 100 μL total volume and 900 μL of BCA reagent was added. Samples were incubated for 30 minutes at 37 °C. Using the standard curve generated by the BSA standards, protein content of each sample was determined.

Western Blotting

To verify the presence of the recombinant CbbO protein, a western blot was performed using antibodies raised against the six histidine residue tag. The protein was separated using an SDS-PAGE following the protocol described earlier. The protein was transferred to a nitrocellulose membrane by lateral electro-transfer at 250 mA for 30 minutes at 4 °C using Western Transfer Buffer (as described in the buffer protocol). To develop the blot, the membrane was soaked in 5% milk solution (PBS with 0.1% Triton) to prevent non-specific binding to the membrane. For antibody recognition a 1:4000 dilution of Anti *CbbO* Rabbit Polyclonal from Cocalico Biologicals was used as the primary antibody. The secondary antibody

was a 1:10,000 dilution of a commercial Goat Anti-Rabbit IgG antiserum, with 5% milk and PBS-0.1% Triton washes in between each antibody. Thermo Scientific SuperSignal West Pico Chemiluminescence Substrate was used to develop the blot for imaging. The chemiluminescence agent was prepared by adding 750 uL of Peroxide solution and 750 uL of Luminol/Enhancer solution. The blot was incubated for 2 minutes before being imaged using the VersaDoc Imager from BioRad.

RESULTS

PCR amplification and cloning of the *cbbO* gene into the TOPO vector

Primers were designed to amplify the chosen region of DNA from *H. neapolitanus* by the PCR technique. The amplified DNA was then visualized using gel electrophoresis (Figure 9). The "S" lane represents the standard 1 kb DNA ladder used for size comparison of the amplified gene of interest. Lanes 1 and 2 represent the same amplified *cbbO* gene. A strong band corresponding to the gene of interest was found between standard bands 3 kb and 2 kb. This result was consistent with the expected gene size of 2.367 kb.

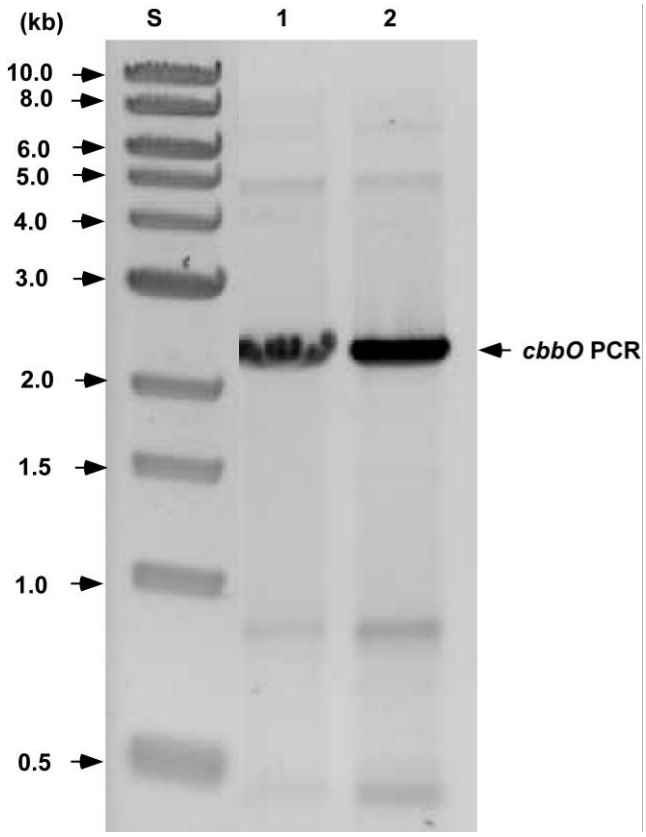


Figure 7: PCR amplification of the *cbbO* gene. "S" lane represents the standard with the respective size of the gene indicated with an arrow. Lanes 1 and 2 represent the amplified *cbbO* gene.

The gene of interest was cloned into the pCR 4Blunt-TOPO vector and the resulting *cbbO* (TOPO) plasmid used to transform the NEB alpha strain of *E. coli*. Overnight cultures of 3 mL were grown using ten colonies along with 50 ug/uL of kanamycin. The plasmid DNA was purified from the overnight cultures using Qiagen mini-prep kit following manufacture protocols. Restriction digest was performed on the purified plasmid DNA using Bam HI and HindIII to verify the presence of the *cbbO* insert. Gel electrophoresis was preformed to confirm the gene of interest (Figure 10). The lane "S" represents the 1 kb DNA ladder standard. The linearized TOPO vector was found to be contained in all 10 samples between standard bands 4 kb and 3 kb, as expected. This was consistent with the known size of the TOPO vector, 3.956 kb. Lanes 3, 4, 7, 8, 9, and 10 showed the excised gene of interest between standard bands at 3 kb and 2 kb. This result was consistent with the expected gene size of 2.367 kb. Small fragments could be seen in lanes 1,2, 5, and 6 but none of these samples also contained a band large enough to represent the gene of

interest.

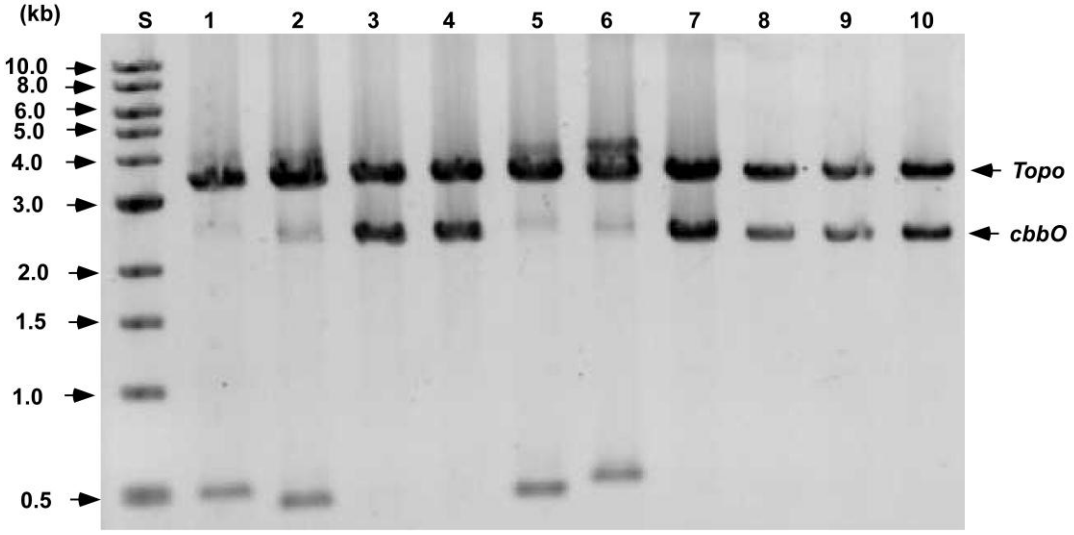


Figure 8: Restriction digest of plasmids containing the amplified *cbbO* gene cloned in the TOPO vector. Lane "S" is the standard and lanes 1-10 represent the enzyme-digested *cbbO* excised from the TOPO vector. The vector band was found between standard bands of 4 and 3 kb, and the gene of interest (2.1 kb) was found between standard bands of 3 and 2 kb.

Cloning of the *cbbO* gene into the protein expression vector pET Duet-1

The sample of the TOPO cloning reaction (Figure 10, lane 4) contained a band that likely represented the *cbbO* gene was used for the purification of the insert for insertion into the pET DUET-1 vector. Digestion with Bam HI and HindIII was designed to insert the *cbbO* gene in the same reading frame as the 6-His protein tag. This measure ensured that the CbbO protein would be translated with a N-terminal 6-His tag, which aids in its purification from the cellular supernatant by affinity chromatography. Selected transformants of NEB alpha *E . coli* were chosen for recombinant plasmid DNA purification and restriction digestion. Gel electrophoresis was performed to confirm the presence of the *cbbO* gene fragment (Figure 11). The lane "S" represents the 1 kb DNA ladder standard. Lanes 1-5 represent restriction enzyme-digested plasmids of individual clones. The pET DUET-1 vector band was observed in all five samples. The band representing the *cbbO* gene was expected to migrate between standard bands of 3 and 2 kb and was present in all five samples. However, samples 3 and 5 seem to also contain a significant quantity of another nucleic acid fragment larger than the *cbbO* gene and the protein expression vector. This band may

represent the *cbbO* gene attached to the linearized pET DUET-1 vector. This vector along with the insert would correspond to ~8.1 kb in size. Samples 2 and 4 were chosen for sequencing. The sequencing results confirmed the sequence of the *cbbO* gene without mutations. Figure 8 represents the final vector map of the *cbbO* gene cloned into the pET DUET-1 plasmid. This construct was eventually used for protein expression.

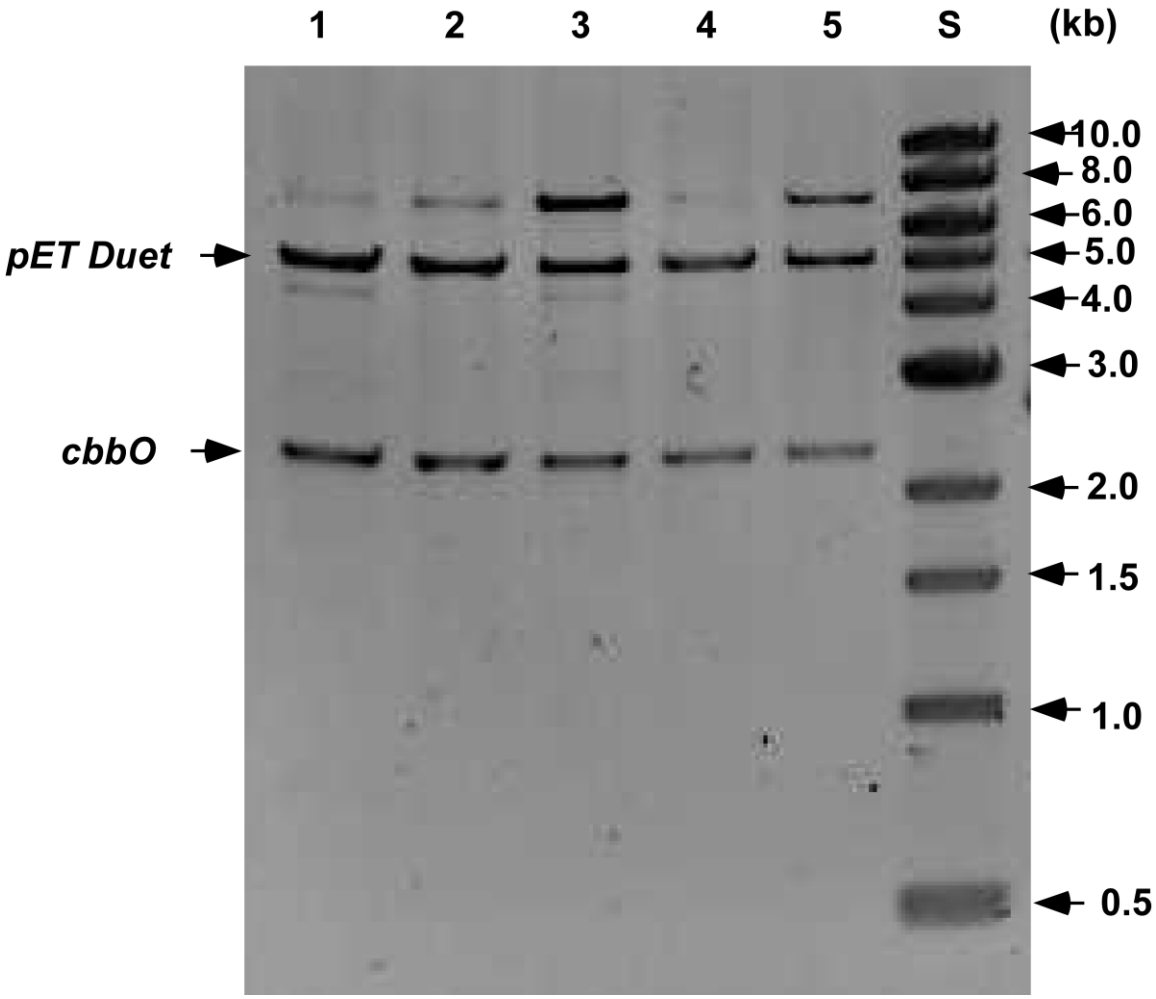


Figure 9: Plasmid DNA digests containing *cbbO* in the pET Duet-1 Protein Expression Vector. Lanes 1-5 represent DNA samples from individual clones. The band representing the gene of interest migrated between standard bands of 3 and 2 kb. Samples 2 and 4 were sent for sequence confirmation. Lane "S" represents the location of the size standard bands.

Small-scale protein expression of Histidine Tagged-CbbO (HT-CbbO)

The pET DUET-1 construct containing the *cbbO* gene was transformed into the BL21DE3 chemically competent strain of *E. coli*, which is more suitable to protein expression than NEB5alpha. Overnight

cultures of 5 mL were then used for small-scale protein expression to test whether the selected clones expressed the CbbO protein. The small-scale protein expression was carried out following the protocol described in the materials and methods section. Figure 12 shows the protein samples after separation on an SDS-PAGE gel and after being stained with Coomassie stain. The even-numbered lanes represent un-induced samples and the odd-numbered lanes represent induced samples. The inducer chemical IPTG was used to overexpress the gene of interest (*cbbO*) in the samples. The lane marked "H" represents purified carboxysomes from *H. neapolitanus* and the lane marked "S" represents the protein standard. The expected molecular weight of the His-tagged recombinant CbbO protein was 90.14 kDa. As expected in Figure 12, lanes 2, 4, and 6 containing the induced samples showed the distinct protein band between 118.6 kDa and 78.9 kDa representing the CbbO protein. The CbbO protein migrated between standard bands 118.6 kDa and 78.9 kDa.

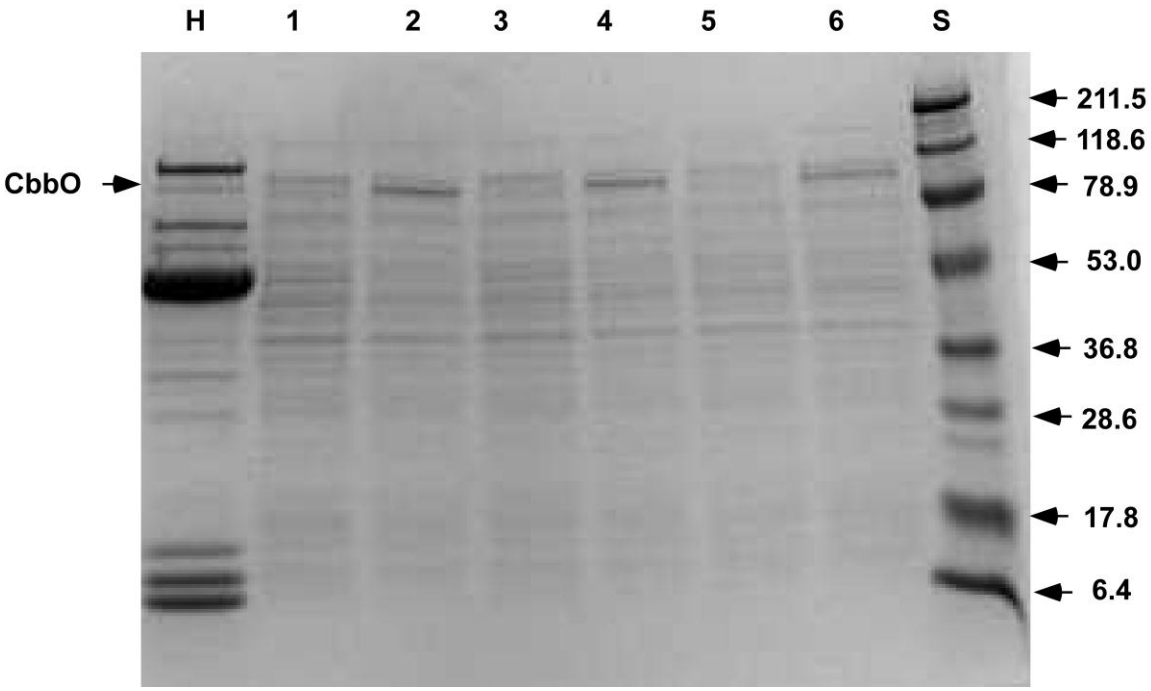


Figure 10: Coomassie stain denaturing Polyacrylamide Gel of the samples from the small scale protein expression. The odd-numbered lanes represent un-induced samples and even-numbered lanes represent induced samples. The lane marked "H" represents purified carboxysomes from wild-type *Halothiobacillus neapolitanus*. Molecular Weight of the His-tag protein: 90.14 kDa. The arrow indicates the presence of the protein.

Purification of HT-CbbO using affinity column chromatography

Larger cultures (1 L) were then grown from the 5 mL overnight cultures of the pET DUET-1 vector containing the *cbbO* gene transformed into the BL21DE3 *E. coli* strain from samples 2 and 4. Protein expression was performed by inducing with IPTG. The HT-CbbO protein was purified from the induced cultures using Ni-NTA affinity chromatography (see Materials and Methods for description). Samples of each fraction were analyzed by separating them on an SDS-PAGE (Figure 13). Lane "S" represented the Bio-Rad protein standard. An arrow indicates the presence of the HT-CbbO protein. Significant quantities of the protein were seen in: [the lysed bacterial supernatant (Lane 1), bacterial pellet (Lane 2), initial flow-through (Lane 3), and the first three eluates (lanes 4, 5, and 6) containing 100, 200, and 300 mM imidazole respectively]. Minimal amounts of the purified protein were observed in the three eluates. It may also be seen that more protein was present in the lysed bacterial supernatant and initial flow through fractions than the pelleted bacterial fraction. The eluates, lanes 4, 5, and 6 all show the protein in decreasing concentrations and possibly increasing purity due to increasing lack of background bands. All three eluates were combined for each sample and a BCA protein assay was used to determine the quantity of protein present. For the protein sample from the column elutions run from sample 2 and 4, approximately 0.7 mg/mL and 0.9 mg/mL of protein were obtained for samples 2 and 4 respectively.

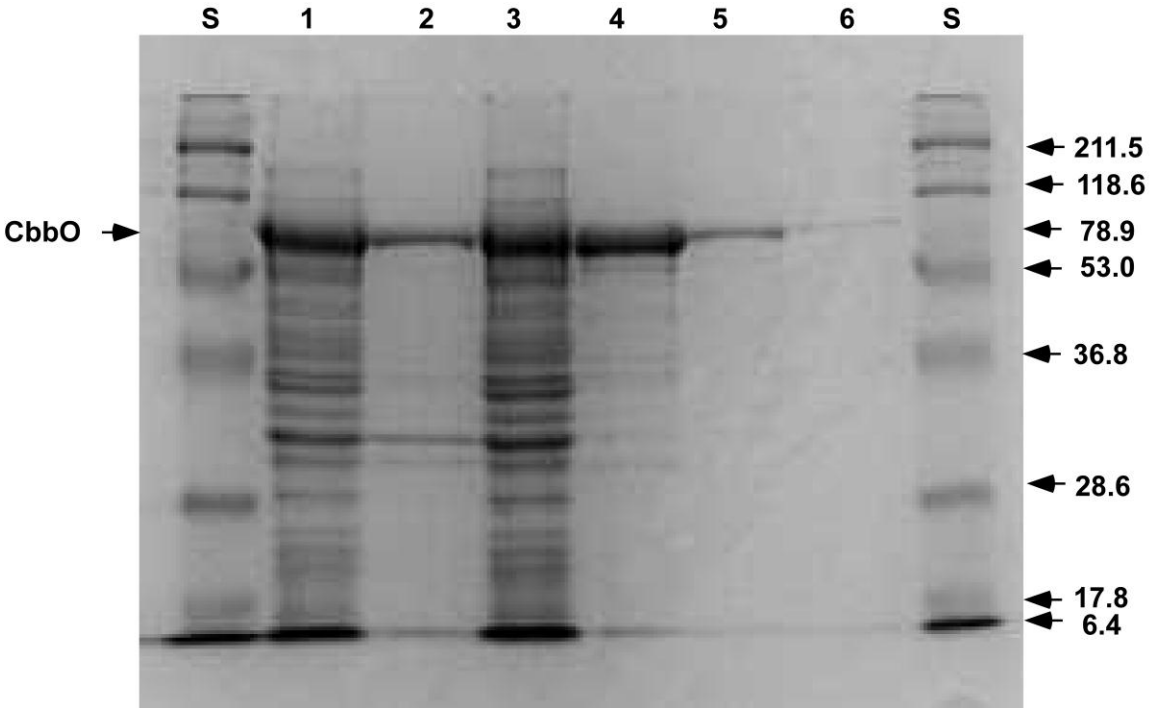


Figure 11: SDS PAGE samples from large scale protein expression. Lane 1 represents the lysed bacterial supernatant sample. Lane 2 represents the lysed bacterial pellet. Lane 3 represents the column flow-through. Lanes 4,5, and 6 show eluates 1, 2, and 3 respectively. Lane "S" represents the molecular weight standards. An arrow indicates the presence of the CbbO protein.

Western blot analysis of HT-CbbO

Protein samples 2 and 4 were then sent to Cocalico Biologicals Inc. for the generation of polyclonal antibodies in rabbits. A Western was completed for the 1st and 2nd test bleed (Figure 14). The figure, 14 a represents the Coomassie stain and 14 b represents the Immunoblot. For the 2nd bleed, a 1:3,000 dilution was used for the primary antibody and a 1:10,000 dilution was used for the secondary antibody. Very little or no target antigen CbbO seemed to be present in the crude extracts. Lane "S" represents the protein standards. Lane "H' represents isolated purified carboxysomes from *H. neapolitanus*. Lane 1 represents the semi-purified CbbO protein. The protein of interest was perhaps visible between standard bands 118.6 kDa and 78.9 kDa, and is represented by the arrow. However, the CbbO protein band was very faint. No target antigen CbbO was visible in the isolated purified carboxysomes.

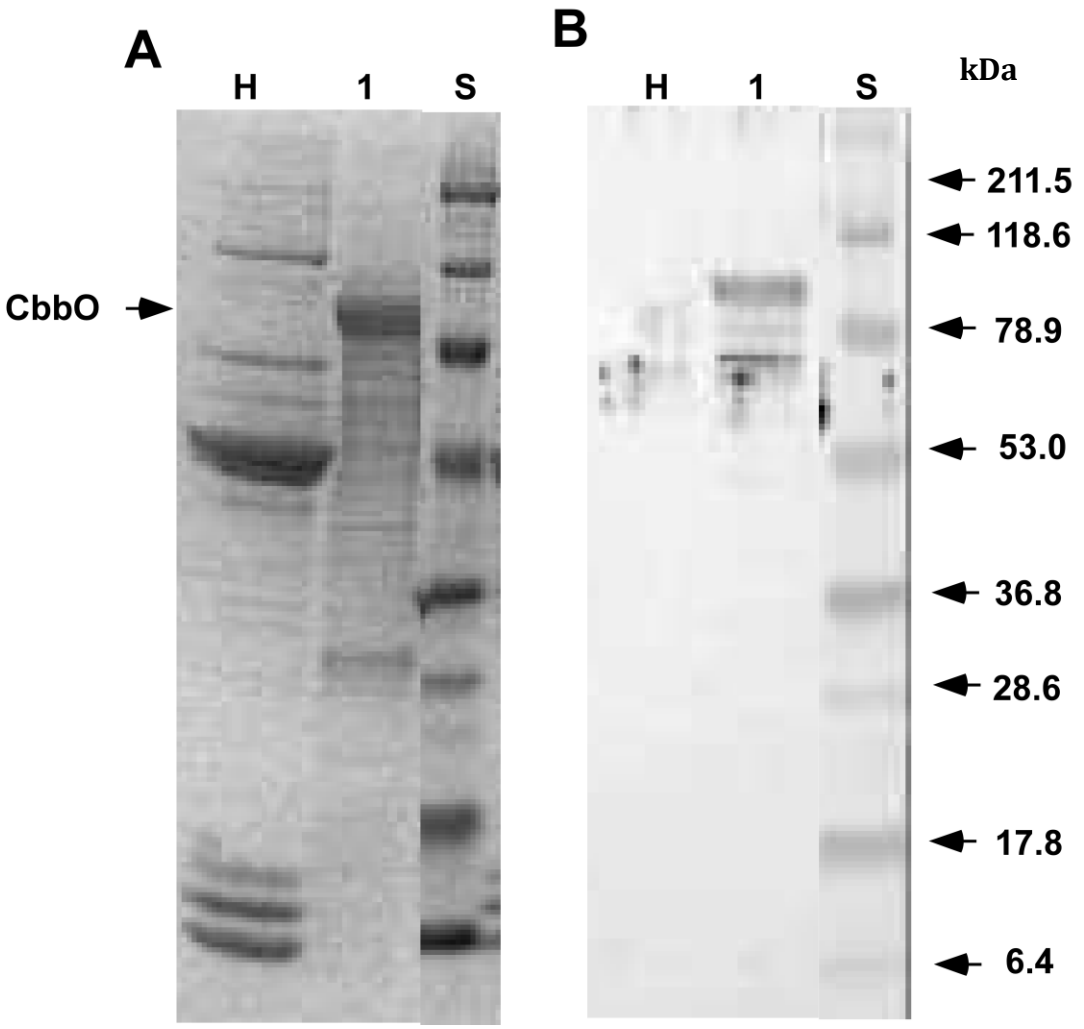


Figure 12: Results of a Western Blot using the 2nd test Bleed of a polyclonal antibody generation. A. Represents the Coomassie stain and B represents the immunoblot. A 1:3,000 dilution was used for the primary antibody and a 1: 10,000 dilution was used for the secondary antibody. The protein of interest is perhaps visible between standard bands 118.6 kDa and 78.9 kDa, and is represented by the arrow. The lanes labeled "S" represent the standards. Lanes "H" represent the isolated purified carboxysomes from *H. neapolitanus*. The 1st lane represents the purified CbbO protein samples.

DISCUSSION & CONCLUSION

The *cbbO* gene was successfully amplified using the Polymerase Chain Reaction technique (Figure 9). Minimal background was seen in both sample lanes, suggesting selected proper amplification of the selected genomic sequence only. It was then cloned into the TOPO vector. In lane 6, a nucleic acid band

can be seen above the standard kb that are expected for the size of the gene of interest and the empty TOPO vector (see Figure 10). The gene of interest was then cloned into the pET DUET-1 protein expression vector. Once again, lanes 2, 3, and 5 seem to indicate a nucleic acid band above the standard kb size expected for the *cbbO* gene and the empty pET DUET-1 vector (see Figure 11). The results of the TOPO and pET DUET-1 cloning reactions could indicate an incomplete digestion by the restriction enzymes, BamHI and HindIII. Small cultures of the pET DUET-1 construct were then grown and three samples were induced with IPTG and three were not. The samples were then analyzed using SDS-PAGE and the resulting gel clearly showed little or no protein expression for the un-induced samples and protein expression for the induced samples (see Figure 12). This result was likely due to the fact that the original pET DUET-1 vector used to create the recombinant pET DUET-1 vector with the inserted *cbbO* gene, required an inducer to activate transcription at the ATG start region (see Figure 8). Once again small cultures were grown for purposes of protein expression. The culture was spun down to remove the media supernatant and the bacterial pellet re-suspended. A sample of this lysed bacterial supernatant was set aside. Then, the re-suspended bacteria was lysed and spun in order to separate the cellular components. The supernatant containing the lighter density protein was purified using a Ni-NTA column. The pellet of lysed cellular components was also re-suspended and set aside. Once the lysed supernatant had passed through the column, additional eluates were performed with wash buffer to elute protein remaining in the column. After lysis, most of the protein was found in the expected purified supernatant fraction (lane 3), though some protein could be seen in the lysed pellet fraction as well (lane 2). It has been hypothesized that protein was seen in the denser pellet fraction because of protein mis-foldings causing the protein to reside in inclusion bodies rather than the cytosol. For sample 2, we obtained 0.70 mg/mL of protein and for sample 4, we obtained 0.93 mg/mL of protein. As determined by a BCA assay, this was a significant amount of protein.

In order to further study this protein, were sent to an outside firm for the generation of polyclonal antibodies. Once the antibodies were received, their affinity for the protein was tested using a Western and Immuno Blot (see Figure 14 a and 14 b). In the lanes “H” the anti-body was used to probe for the presence of the protein in wild-type, isolated pure carboxysomes. No protein band of the appropriate size was seen, however. Therefore, it was hypothesized that the CbbO protein is not sequestered into carboxysomes itself, though it still may function to help carboxysomes sequester RubisCO. Additionally, the antibody was able

to be successfully used as a probe for the CbbO protein as seen by the lanes, 1, in which the CbbO protein was present as indicated by its position relative to the protein standard. The antibody displayed cross-reactivity as evidenced the presence of the immunoblot signal.

The *cbbO* is a gene encoded downstream of the traditional carboxysome operon in *H. neapolitanus* which can be expressed as a significant amount of soluble recombinant protein as determined by a BCA assay. The protein could also be used for generation of polyclonal antibodies. The efficacy of these antibodies remains to be seen.

Future researchers should be able to complete a Western Blot for the final bleed in order to determine the efficacy of the antibody probe for the presence of CbbO protein. A radiometric RubisCO assay could also be performed with and without CbbO to elucidate the activase property of CbbO. Yet another avenue for future work would be to determine the structure of the CbbO protein. Determining the structure of the CbbO protein could prove advantageous in determining its biochemical role in the fixation of CO_2 by RubisCO as well as contribute to a greater understanding of the mechanism for creating the intact carboxysome *in vivo*. Finally, future researchers may be able to confirm the involvement of CbbO in RubisCO packaging. The future work should be conducted in order to elucidate the importance of the CbbO protein in the structure and function of carboxysomes. This elucidation could be completed by knocking out the cbbO gene and by measuring the RubisCO enzyme activity in the carboxysome granules and observing the carboxysome structure using transformation electron microscopy.

REFERENCES

Bobik, T. (2005, November 6). Polyhedral organelles compartmenting bacterial metabolic processes. *Applied Microbiology and Biotechnology, 70*, 517-525. doi:10.1007/s00253-005-0295-0

Bonacci, W., Teng, P., Afonso, B., Niederholtmeyer, H., Grob, P., Silver, P., & Savage, D. (2012, January 10). Modularity of a carbon-fixing protein organelle. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 478-483. doi:10.1073/pnas.1108557109

Burey, S., Poroyko, V., Ergen, Z., Fathi-Nejad, S., Schuller, C., Ohnishi, N., & Fukuzawa, H. B. (2007). Acclimation to low [CO_2] by an inorganic carbon-concentrating mechanism in Cyanophora paradoxa. *Plant, Cell & Environment, 30*, 1422-1435. doi:10.1111/j.1365-3040.2007.01715.x

Cannon, G.C. and J.M. Shively, (1983) Characterization of a homogenous preparation of carboxysomes from *Thiobacillus neapolitanus*. Arc. Microbiol. 134 (1): p. 52-59.

Esparza, M., Cardenas, J.,Bowien, B., & Jedlicki, E. H. (2010). Genes and pathways for CO_2 fixation in the obligate, chemolithotrophic acidophile, Acidithiobacillus ferroxidans, Carbon fixation in A. ferroxidans, BMC Microbiology, 10 (229), 1-15. Retrieved March 2012

Feller U, Anders I, Mae T (2008). "Rubiscolytics: fate of RubisCO after its enzymatic function in a cell is terminated". *J. Exp. Bot.* **59** (7): 1615–24. doi:10.1093/jxb/erm242. PMID 17975207.

Gonzales, A., Light, Y., Zhang, Z., Iqbal, T., Lane, T., & Martino, A. (2005, August 13). Proteomic analysis of the CO_2-concentrating mechanism in the open-ocean cyanobacterium synechococcus WH8102. *Canadian Journal of Botany, 83*, 735-745. doi:10.1139/B05-056

Heinhorst, S., & Cannon, G. (2010, April 27). Addressing microbial organelles: A short peptide directs enzymes to the interior. *Proceedings of the National Academy of Sciences, 107*(17), 7627-7628. doi:10.1073/pnas.1003433107

Heinhorst, S., & Cannon, G. C. (2008, September). A new, leaner and meaner bacterial organelle. *Nature Structural & Molecular Biology, 15*(9), 897-898. Retrieved 2012

Hayashi N.R, Arai H, Kodama T, Igarashi Y. (1997) The novel genes, *cbbQ* and *cbbO*, located downstream from the RubisCO genes of Pseudomonas hydrogenothermophila, affect the conformational states and activity of RubisCO. *Biochem Biophys Res Commun, 241,* 565-569.

Kereld, C., Heinhorst, S., & Cannon, G. (2010). Bacterial Microcompartments. *Annual Review of Microbiology, 64*, 391-408. Retrieved 2012

Marin, B., Nowack, E., Glockner, G., & Melkonian, M. (2007, June 5). The ancestor of the Paulinella chromatophore obtained a carboxysomal operon by horizontal gene transfer from a Nitrococcus-like gamma-proteobacterium. *BMC Evolutionary Biology, 7*(85), 1-14. doi:10.1186/1471-2148/7/85

Menon, B., Heinhorst, S. S., & Cannon, G. (n.d.). Proton Permeability of a Bacterial Microcompartment. 1-23.

Miyakoshi, M., Nishida, H., Shintani, M., & Yamane, H. &. (2008, July 23). High-resolution mapping of plasmid transcriptomes in different host bacteria. *BMC Genomics*, *10*(12), 1-15. doi:10.1186/147-2164-10-12

Roberts, E., Cai, F., Kerfeld, C., Cannon, G., & Heinhorst, S. (2011, December 9). Isolation and Characterization of the Prochlorococcus Carboxysome. *Journal of Bacteriology*, *194*(4), 787-795. doi:10.1128/JB.06444-11

Shively, J., Keulen, G., & Meijer, W. (1998, 1998). Something from Almost Nothing: Carbon Dioxide Fixation in Chemoautotrophs. *Annual Review of Microbiology, 52*, pp. 191-230. Retrieved 2012

Tsai, Y., Sawaya, M., Cannon, G., Cai, F., Williams, E., Heinhorst, S., & Kerfeld, C. Y. (2007, June). Structural Analysis of CsoS1A and the Protein Shell of the Halothiobacillus neapolitanus Carboxysome. *Public Library of Science: Biology, 5*(6). Retrieved 2012

Valle, E., Kobayashi, H., & Akazawa, T. (1987, December 7). Transcriptional regulation of genes for plant-type ribulose-1,5-bisphosphate carboxylase/oxygenase in the photosynthetic bacterium, Chromatium vinosum. *European Journal of Biochemistry*, 483-489. Retrieved 2012

Yagi, J., Sims, D., Brettin, T., Bruce, D., & Madsen, E. (2009). The genome of Polaromonas naphthalenivorans strain CJ2, isolated from coal tar-contaminated sediment, reveals physiological and metabolic versatility and evolution through extensive horizontal gene transfer. *Envirnomental Microbiology, 11*(9), 2253-2270. doi:10.1111/j.1462-2920.2009.01947

Yeates, T., Kerfeld, C., Heinhorst, S., Cannon, G., & Shively, J. (2008, August 4). Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nature Reveiws Microbiology*. doi:10.1038/nrmicro1913

Yooseph, S., Nealson, K. R., Dupont, C., Kim, M., Johnson, J. M., Ferriera, S. B., . . . al., e. (2010, November 4). Genomic and Functional adaptation in surface ocean planktonic prokaryotes. *Nature, 468*. doi:10.1038/nature09530

Zimmerman, S., & Ferry, J. (2008). The Beta and Gamma Glasses of Carbonic Anhydrase. *Current Pharmaceutical Design, 14*, 716-721. Retrieved 2012