Optimization of a bacterial expression system for the purification and activity analysis of a *Caenorhabditis elegans* chitin deacetylase

Molly Chirunomula
Senior Thesis, Spring 2015
Dr. Juliet A Fuhrman
Tufts Biology Department

Table of Contents:

Abstract:

Parasitic nematodes cause severely debilitating diseases such as lymphatic filariasis and onchocerciasis. One potential therapeutic target for the treatment or prevention of nematode infections is chitin deacetylase (CDA). CDA converts the biologically important carbohydrate polymer of chitin (beta-1,4-linked N-acetylglucosamine) to chitosan. The effectiveness of this process has been correlated to healthy nematode development. The aim of this study was to generate soluble, purified, and active CDA for further characterization through various activity assays. Variations on a bacterial expression system were constructed using the pET22b and pGEX-4T-2 bacterial vectors and the *Caenorhabditis elegans* CDA gene F48E3.8c. Host cell cultures carrying distinct versions of the recombinant plasmids were extracted using a variety of induction and lysis conditions. Expression of soluble CDA was accomplished using the full F48E3.8c gene, as well as a truncated version of the gene. Enzymatic activity was successfully restored following the denaturation and crude purification of CDA embedded in polyacrylamide gels. Further experiments are required to optimize the purification of active protein in solution.

Introduction:

Filarial infections and chitin deacetylase

Lymphatic filariasis and onchocerciasis caused by the nematodes *Wuchereria bancrofti, Brugia timori, Brugia malayi*, and *Onchocerca volvulus* affect over 150 million people across the globe (Lentz *et al.* 2013). Lymphatic filariasis (LF), or elephantiasis, is an extremely painful, disfiguring disease caused by filarial nematodes that infiltrate the human lymphatic system. Adult worms vary from 3-10 cm in length, and cause significant damage to the fluid balance of

the body; this can result in grotesque lymphoedema and thickening of the skin (Lymphatic filariasis 2015). LF also interferes with the immune system of the host, and often results in kidney damage, in addition to having severe social and economic impacts on symptomatic individuals (Lymphatic filariasis 2015). Onchocerciasis, or river blindness, is a similarly devastating disease; it is estimated that 500,000 people are currently blind as a result of *Onchocerca volvulus* infection transmitted by *Simulium* blackflies (African Programme for Onchocerciasis Control 2015). Fear of disease transmission has driven sub-Saharan communities away from fertile river valleys where these flies breed, decreasing agricultural yields and exacerbating poverty (African Programme for Onchocerciasis Control 2015).

Current efforts to control LF entail mass administration of albendazole with ivermectin or diethylcarbamazine in endemic countries, which effectively reduces disease transmission by killing larval stages of filaria (microfilariae) in the bloodstream, but merely paralyzes adult worms that have already developed in the body of a victim (Lentz *et al.* 2013). There is currently a great need for research to develop interventions that may allow more selective targeting and treatment of this disease. Chitin deacetylase (CDA) has been identified as a potential therapeutic target (Heustis *et al.* 2012). CDA is an enzyme that deacetylates chitin, the strong carbohydrate polymer of β-1,4-linked N-acetylglucosamine, to produce chitosan, a less resilient polysaccharide bearing multiple free amino groups at the sites where the acetyl groups were cleaved from the substrate. It has been suggested that the CDA activity in the chitinous pharyngeal lining supports normal development in the free-living model nematode *C. elegans* (Heustis *et al.* 2012). Further characterization of the role of CDA in *C. elegans* chitin metabolism can aid the future development of therapeutics to combat disease caused by filarial nematodes.

Biological functions of CDAs and PDAs vary across species

The first chitin deacetylase ever characterized was extracted from the fungus *Mucor rouxii* (Kafetzopoulos *et al.* 1993). Since then CDA has been found in numerous fungi, including *Colletotrichum lindemuthianum* (Blair *et al.* 2006; Zhao *et al.* 2014), *Aspergillus nidulans, Absidia coerulea,* and *Saccharmyces cereviseae* (Blair *et al.* 2006). Insect CDAs and bacterial peptidoglycan deacetylases (PDAs/PgdAs) have also been described as homologues of the fungal CDAs.

CDAs are classified in the CAZy (Carbohydrate-Active enZYme) database as part of carbohydrate esterase family 4 (Blair *et al.* 2005, Heustis *et al.* 2012, Psylinakis *et al.* 2005, Carbohydrate Esterase Family 4 2015). This family also includes peptidoglycan deacetylases, rhizobial NodB chitooligosaccharide deacetylases, zylanases and acetylxylan esterases. The CAZy family 4 esterases share a polysaccharide deacetylase domain that allows them hydrolyze N-linked acetyl groups on N-acetylhexosamines. As a result of their significance in the biological functioning of so many species, including a variety of fungus, bacteria, nematode, marine invertebrate, and insect species, CE4 esterases and chitin deacetylases have been the subject of significant study to characterize their patterns of developmental expression, structure, and biological function.

Studies of CDAs derived from diverse species have collectively revealed that CDAs are often critical enzymes for biological success, but the reason for this dependence varies. Despite the substantial homology seen across all CDAs, they have been found to demonstrate key differences, such as substrate specificity. For instance, Psylinakis *et al.* found that the two closely related species *Bacillus anthracis and Bacillus cereus* only deacetylate *N*-acetyl-D-glucosamine

when it is attached to *N*-acetylmuramyl-L-alanine-D-isoglutamine via a (ß- 1,4) linkage (Psylinakis *et al.* 2005). Our growing knowledge of the differing functions of various CDAs speaks to the highly distinct biological roles that CDA plays in different organisms. Numerous fungal pathogens of vertebrates and plants employ de-N-acetylation of protective chitinous structures as a modification technique that reduces the ability of host immune enzymes to target and destroy the invading organisms. The pathogenic microsporidia *Encephalitozoon cuniculi* uses a CDA to evade host immune response by making it resistant to the glycoside hydrolases native to bacteria, plants, and animals (van Aalten *et al.* 2009). CDAs can also protect nonpathogenic fungi like the yeast *Saccharomyces cerevisiae,* which relies on a CDA to make its strong, chitinous ascospore wall resistant to lytic enzymes in its environment (Blair *et al.* 2006).

Peptidoglycan deacetylases fulfill a similar function in numerous species of virulent bacteria, including *Staphylococcus epidermidis*, which requires the deacetylation of glycans on its cell surfaces to form biofilms and resist host immune mechanisms (Vuong *et al.* 2004 qtd in Deng *et al.* 2009). A similar tactic is used by *Listeria monocytogenes* (Boneca *et al.* qtd in Deng *et al.* 2009) and *Streptococcus pneumoniae* (Vollmer *et al.* 2000, 2002 qtd in Deng *et al.* 2009). *S. pneumoniae* infectivity declines significantly in a mammalian host in the absence of the *Sp*PgdA gene; this gene has even been suggested as a potential drug target for treatment of infection (Blair *et al.* 2005).

Because insects are generally not pathogenic, the significances of insect CDAs tend to be more directly related to structural development. Two CDAs, *serp* and *verm*, have been characterized in *Drosophila* to be important for shaping the tracheal tubes, important respiratory structures (Luschnig *et al.* 2006). These enzymes also regulate the structure of the epidermal

cuticle (Luschnig *et al.* 2006). Extensive characterization of nine CDAs in *Tribolium casteneum* has revealed highly differential developmental patterns of expression, suggesting that each CDA has a different function in the development of the insect (Arakane *et al.* 2009). Thus, CDAs demonstrate distinct biological functions not only between species, but also within them.

Aims of this study

This study ultimately aims to characterize nematode CDA activity by defining optimal conditions for enzymatic function. This information could inform the future use of high-throughput screening technology to identify molecules that inhibit enzymatic function, which could potentially be developed into therapeutics that combat nematode infection. Understanding the activity of CDA derived from nematodes will also lend insight into the protein's native environment, helping to describe parasitic nematode physiology, and also to delineate the biological functions of CDAs specifically in nematodes.

A promising avenue for the characterization of CDA activity is the activity assay we adapted from Blair *et al.* (2005) in a previous study (Chirunomula 2013). This assay centers on the premise that the cleavage sites at which CDA removes acetyl groups from chitin can be detected using the organic reagent fluorescamine, which binds free amino groups to produce a fluorescent product. The intensity of the fluorescence produced via reaction with fluorescamine can be correlated to the amount of deacetylated product present, which serves as a quantitative measure of CDA activity (Blair *et al.* 2005, Chirunomula 2013).

In order to proceed with our fluorescence-based activity assay, soluble and active CDA must be purified in the laboratory. Three distinct bacterial expression systems were designed in

an effort to express this protein in its soluble, active form. Cell cultures were subjected to a variety of protein extraction and purification techniques, and significant progress was made towards the development of a successful expression system.

Materials and Methods:

Bacterial cultures

3-mL overnight liquid cultures of transformed cells were transferred to larger Erlenmeyer flasks (250 - 2000 mL) and grown 2-5 h at 37˚C, with shaking at 220 rpm. Cells were induced with IPTG at varying concentrations (commonly 0.1 mM or 1 mM), and varying temperatures and durations (commonly 16˚C for 18 h, 30˚C for 2 h, or 37˚C for 2 h). Negative controls included cell cultures without CDA insert, and uninduced cultures with insert.

SDS-PAGE with Coomassie staining

10 % or 4-15% gradient polyacrylamide gels were used for standard SDS-PAGE. Samples were prepared with reducing sample buffer, heated 5 min at 85-90^{o}C, and microfuged 5 min at ~25^{o}C, 16,100*g* prior to electrophoresis. Following electrophoresis, gels were incubated with rocking in 0.1% Coomassie Blue R-250 stain (0.1% Coomassie Blue R, 10% acetic acid, 40% methanol) 45 min, ~25^{o}C; Destain A (50% methanol, 10% acetic acid) 1 h, ~25^{o}C; Destain B (5% methanol, 7% acetic acid). Gels were visualized on white light.

S*ubstrate gel analysis*

12% polyacrylamide discontinuous substrate gels were prepared with glycol chitin according to Trudel & Asselin (1990). Standard SDS-PAGE gels were poured with 0.1% glycol chitin embedded in the separating gel. Electrophoresis of protein samples was performed on duplicate gels. As a negative control, one gel was immediately stained 5 min in 40 mL of 10% Calcofluor white in 0.5 M Tris buffer (pH 9.0) and destained with rocking, 2 h, in dH_2O before photographing. Experimental gels were rocked 1-2 nights in 20 mL 1% Triton-X 100 (Calbiochem) in 50 mM Hepes (pH 7.0) to allow CDA renaturation, then stained in calcofluor/ Tris and washed 2 nights in dH_2O before photographing.

*Protein purification columns**

Novagen His-bind column was pre-wet with 6 mL 1x binding buffer. Column was loaded with cell extract (7 parts lysate : 1 part 8x binding buffer), then 20 mL 1x binding buffer, followed by 10 mL 1x wash buffer. Samples were eluted with 4 mL 1x elute buffer. Column was stripped with 4 mL 1x strip buffer. Samples of elute were collected after each step for analysis.

*Co-MAC column protein purification protocol is reported in alternate publication (McNeil 2014).

Acetone precipitation of protein samples

Protein samples were concentrated as needed by chilling at -20˚C, 1.5 h in 80% acetone and microfuging 15 min at 4^0C. Supernatant was replaced with 70% chilled acetone followed by a

second round of microfugation. Pellets were dried and resuspended in reducing sample buffer for SDS-PAGE.

Western blotting

Nitrocellulose was rocked 5 min in ppH_2O; nitrocellulose, filter paper, and gel were rocked 15 min in Towbin Transblot Buffer (0.025 M Tris-base, 0.19 M glycine; 20% methanol, 80% ppH_2O). Unstained SDS-PAGE protein gel transfers to nitrocellulose were assembled and run 30 min at 15 V in transblotter (BioRad Semi-Dry Transfer Cell). Nitrocellulose was blocked 1 h in 1% BSA/PBS, incubated 3 h at ~25^{o}C or overnight at 4^{o}C in primary antibody and washed in 0.05% Tween/PBS, then incubated in secondary antibody by the same protocol. Control blots were incubated in secondary antibody only. Blots were developed ~20 min with 0.03% 4-chloronaphthol in PBS, with 2 uL 30% H_2O_2 per 10 mL. Blots were stored in ppH_2O.

Bacterial vector preparation

Novagen pET 22b and GE Healthcare Life Sciences pGEX-4T-2 vectors were digested with NE BioLabs Restriction Endonucleases, and dephosphorylated with NE BioLabs Shrimp Alkaline Phosphatase (rSAP). Protocols were performed as directed by the manufacturer.

CDA insert preparation

CDA inserts were digested to generate sticky ends using NE BioLabs Restriction Endonucleases as directed by the manufacturer. NE BioLabs Mung Bean Nuclease was briefly used as a polishing agent, also as directed. Constructs II and III inserts were PCR-amplified from an

F48E3.8c cDNA template stock ("yk template") using Life Technologies Elongase Enzyme Mix as directed by manufacturer.

Vector-insert ligation

CDA insert ligations into Promega pGEM-T Easy vector and prepared bacterial vectors were performed in a 3:1 insert:vector ratio by mass (pGEX-4T-2 vector 4970 bp; pET22b vector 5493 bp; nontruncated CDA insert ~1.1 kb; truncated CDA insert ~800 kb). Protocol performed as directed in Promega *pGEM-T and pGEM-T Easy Vector Systems: Technical Manual*. Negative control ligations were performed without insert.

Transformation into JM109 cells

Vector-insert ligation reactions were transformed into Promega JM109 *E. coli* competent cells as directed in *Promega* E. coli *Competent Cells: Technical Bulletin*. Transformation reactions were performed in Falcon 17 x 100 mm, Polypropylene Round-Bottom Tubes. Reactions were plated on LB agar with ampicillin (pET22b and pGEX confer ampicillin resistance). Supercoiled plasmid (Promega Control DNA) was transformed into JM109 cells as a positive control. Negative controls included the transformation of ligation reactions without insert, and transformation of water instead of ligation reaction.

Plasmid Extraction from JM109 Cells

Following transformation, recombinant plasmids were extracted from JM109 cells using Promega Wizard Plus SV Minipreps DNA Purification System as described by manufacturer. 3-

mL LB/ampicillin cultures of bacterial clones were incubated overnight at 37˚C in test tube rotator in preparation for plasmid extraction.

Agarose gel electrophoresis of DNA

Analytical and preparative agarose gel electrophoresis of DNA (including PCR products, digested vectors with and without insert, and uncut/cut recombinant plasmids) was performed in 0.8% agarose/TBE buffer gels with 0.01% ethidium bromide, by standard protocol.

Transformation into E. coli *cells*

Recombinant plasmids confirmed to contain CDA insert were transformed into Novagen Rosetta and Rosetta2 *E. coli* strains as directed in Novagen User Protocol TB009 (LB agar was prepared with carbenicillin and chloramphenicol, as Rosetta and Rosetta2 *E. coli* strains are chloramphenicol-resistant). Permanent cultures of CDA expression hosts were stocked in 15% glycerol/LB with antibiotics at -80˚C.

DNA extraction from agarose gel

CDA insert was purified prior to ligation through gel electrophoresis of digested, recombinant T-easy.CDA. Portion of agarose gel containing insert was excised (visualized under low levels of UV exposure) and extracted using MP Bio Geneclean II kit, as directed by manufacturer.

CDA extraction from expression hosts

Cells carrying recombinant plasmids were lysed in 0.4% lysonase in Novagen BugBuster Protein Extraction Reagent. Sarkosyl lysates were prepared by adding 1% or 10% sarkosyl to standard lysis preparation. Lysates with protease inhibitors were prepared by adding 10 ug/mL aprotinin and 10 ug/mL leupeptin to the lysis preparation. Inclusion bodies extracted from lysates were soaked overnight in 10% sarkosyl/ST Buffer (10 mM beta-mercaptoethanol, 50 mM Tris, 300 mM NaCl, 5 mM ZnCl2). Sonication of cells suspended in sterile PBS was performed using Branson Sonifer 450; one standard "round" of sonication entailed 6, 15-second, constant pulses at intervals of 1 minute.

GSH sepharose 4B purification

Cellular extracts containing 10% sarkosyl were prepared for GSH sepharose 4B purification by preparing a 5-fold dilution in ST buffer with 4% Triton X-100 and 40 mM CHAPS. 50 uL of Pharmacia GSH Sepharose 4B was pelleted at 500*g* and washed 3 times in PBS. Beads were then incubated 30 minutes with rotating and unbound fraction was removed. Beads were washed 5 times in PBS. To extract bound fraction for analysis, beads were heated at 85˚C in 25 uL reducing sample buffer for 5 minutes and microfuged at room temperature, 1000*g*.

Results:

***Expression system I: pET22b.CDA in Rosetta(DE3)* E. coli**

IA) Preliminary expression

Protein gel electrophoresis of previously constructed pET22b.CDA/Rosetta(DE3) lysates revealed specific enhancement of CDA in extracts of induced cells compared to uninduced cells. However, this construct did not drive export of CDA into the periplasmic fraction of cellular extracts as intended (data not shown; experiment documented by McNeil 2014). Specific enhancement of CDA occurred only in the whole lysate, and not in the soluble fraction of the lysate, indicating that the CDA protein was largely insoluble (Figure 1). Manipulations of induction temperature and inducer concentration failed to increase CDA solubility (Figure 2). Furthermore, reducing the temperature of the induction caused a significant decrease in the net yield of CDA present in cellular extracts, regardless of the concentration of inducer (Figure 3).

IB) Column purification

Purification of Expression System I cellular extracts was first attempted under native conditions using the Novagen His-Bind column, which contains a resin charged with Ni^{2+} ions. The C-terminal His_6 tag on the recombinant CDA protein was expected to chelate around the cation, causing the protein to adhere to the column matrix. Purified CDA could then be eluted by flooding the column with imidazole, outcompeting the protein for binding sites on the resin and causing the release of purified CDA into the eluate. Protein composition of the eluted fraction of construct I cellular extracts was found to be highly mixed. Although the eluate contained a

detectable amount of a protein matching the mobility of CDA, equivalent or greater amounts of other proteins were present as well; the column failed to specifically isolate the CDA (Figure 4).

A second column purification attempt was made using the Novagen Co-MAC cartridge, which is charged with Co^{2+} instead of Ni^{2+}ions. To increase the yield of CDA, inclusion bodies from construct I cells were denatured and solubilized with urea prior to passage through the column. Purification resulted in the specific isolation of CDA and two lower molecular weight proteins, which were likely degradation products (Figure 5). The concentrated Co-MAC eluate was shown to renature and demonstrate activity via a substrate gel assay adapted from Trudel and Asselin (1990) (Figure 6). Eluate was analyzed via electrophoresis on a polyacrylamide gel embedded with glycol chitin. The SDS in the gel was then outcompeted by Triton X-100 to allow protein renaturation and substrate deacetylation. Calcofluor-white staining, which hyperfluoresces in the presence of chitosan, enabled visualization of regions of the gel containing active protein. Despite the harsh denaturation, purification, and renaturation conditions entailed by this method, CDA generated from construct I cells demonstrated the potential to catalyze substrate conversion (Figure 6). The hyperfluorescent band that appeared on the gel at a mobility of ~37 kDa also served to provide conclusive evidence for the identification of this protein as CDA.

IC) Final evaluation of the construct

Further experimentation revealed that denatured CDA eluted from the Co-MAC column was particularly resistant to renaturation in solution (McNeil 2014). Enzyme activity could not be restored in solution, and it was ultimately concluded that the pET22b.CDA/Rosetta(DE3)

expression system could not be utilized to purify active and soluble CDA. The failure of this system to generate soluble CDA could potentially be due to the presence of a hydrophobic signal peptide encoded upstream of the multiple cloning site (MCS) in the pET22b vector (Figure 7). This signal peptide did not seem to be directing CDA export; no secreted protein was detected through analysis of the periplasmic fraction of cellular lysates (McNeil 2014). Additionally, the peptide was not cleaved from the CDA following protein translation. Thus, it served no functional purpose in the construct, and may have been reducing CDA solubility. Upon re-evaluation of the recombinant plasmid design, a new bacterial vector was selected to express *C. elegans* F48E3.8c.

***Expression system II: pGEX-4T-2.CDA in Rosetta* E. coli**

II-A) *Selection of the pGEX-4T-2 bacterial vector in expression system design*

pGEX-4T-2 (pGEX) was selected among many commercially available bacterial plasmid vectors to carry the CDA gene in the new expression construct. Two features of the pGEX plasmid made it a strong candidate to replace pET22b as the CDA vector in the previously tested expression system. pGEX lacks the hydrophobic signal peptide encoded by pET22b, which increases its likelihood of producing soluble protein. Additionally, pGEX encodes a Glutathione-S-Transferase (GST) tag upstream of the MCS. This tag has been reported to aid protein solubility by improving the accuracy of protein folding (EMD Millipore 2011). The GST tag also enables fusion protein purification via GSH sepharose. GST has a high affinity for the ligand glutathione (GSH); GST fusion proteins are therefore selectively bound by GSH sepharose beads. These proteins can be eluted by flooding the beads with excess GSH. In the case of

enzymatic proteins, GST tags also indicate the likelihood of activity. GST must be properly folded in order for binding and subsequent elution of the purified fusion protein to occur; thus, fusion proteins that are effectively purified are likely to be present in their proper conformation.

II-B) Subcloning of the recombinant plasmid

Multiple subcloning schemes were attempted while constructing the recombinant pGEX.CDA plasmid used in Expression System II. The CDA insert was initially cleaved out of the construct I recombinant plasmid, pET22b.CDA, via a sequential digest with NcoI, Mung Bean Nuclease (MBN), and Xho I. MBN was used to convert the NcoI sticky end of the insert to a blunt end, allowing for its later ligation to the SmaI-cut blunt end of the vector. The purified insert was ligated directly into SmaI/XhoI cut pGEX. Transformation of the resulting plasmids into competent cells produced very few colonies. Upon DNA extraction and analysis, it was found that the plasmids taken up by these cells did not contain the insert (data not shown). It was likely that polishing the NcoI sticky end of the insert with MBN had failed to produce a blunt end that could be ligated to the SmaI blunt end. A new tactic was adopted to avoid using SmaI by introducing a new restriction site to the 5' end of the insert via PCR (Figure 8).

PCR primers were designed to amplify the F48E3.8c transcript using cloned full-length cDNA (yk1130903) as template. This purified F48E3.8c template was a previously generated isolate from a bulk *C. elegans* cDNA library. The reverse primer for the gene had been previously designed by Heustis (2012) to encode an XhoI site, and the forward primer was newly designed to encode a BamHI site; these restriction sites were thus incorporated into the final

PCR amplimer. All 8 PCR amplification reactions successfully generated amplimers of approximately 1.1 kb in length, equal to the size of the desired cDNA fragment (Figure 9A).

All PCR reactions were completed using the Life Technologies DNA polymerase enzyme mix *Elongase*, which consists of a mixture of thermostable DNA polymerases derived from *Taq* as well as *Pyrococcus* species. One of the enzymes incorporated into this cocktail adds an overhanging adenine nucleotide to each 3' end of the generated PCR amplimer. Promega's pGEM T-easy vector is constructed with an overhanging thymine on each 3' end. Thus, the amplified, truncated CDA gene (with the added restriction sites) could be directly ligated into Promega's pGEM T-easy vector and cleaved out with BamHI and XhoI to reliably generate sticky ends. This tactic was adopted because simply digesting the vector alone appeared to require the removal of too few nucleotides from each end of the insert for accurate cutting to occur (data not shown). The regeneration of the 1.1 kb insert from the digest of the recombinant T-easy.CDA plasmid indicated that the ligation was successful (Figure 9B).

The pGEX vector was prepared for ligation through a simultaneous BamHI/XhoI double digest. The slight difference in size observed between the uncut and cut plasmid verified that the vector had been cleaved (Figure 9C). Regeneration of the 1.1 kb insert through digestion of the recombinant pGEX.CDA plasmid later indicated that ligation of the CDA insert into pGEX had been successful, and that construction of the pGEX.CDA vector was complete (Figure 9D).

II-C) Preliminary expression

Preliminary induction and lysis of pGEX.CDA in Rosetta under standard conditions resulted in extremely high expression of insoluble CDA (Figure 10). Gel electrophoresis of

cellular lysates gave no indication that any soluble CDA had been generated (Figure 10). Manipulating the concentration of inducer, induction temperature, and duration of induction did not appear to affect fusion protein expression or improve solubility (Figure 11). Substrate gel experimentation confirmed the identity of the specifically enhanced electrophoresis bands in the insoluble fractions of the gel as CDA, and also proved that the expressed protein had the potential to catalyze chitin deacetylation (Figure 12).

II-D) Test for aggregation due to interstrand disulfide bridges

A possible cause of CDA insolubility was the inappropriate formation of interstrand disulfide bridges. This would have caused the protein to aggregate and precipitate out of solution following host cell lysis. If this hypothesis were to be verified, the recombinant pGEX.CDA plasmid could be transformed into a new strain of host cells expressing thioredoxin, which would aid the regulation of disulfide bridge formation and reduce the occurrence of unnatural disulfide bridges.

The positive signal at the expected mobility of the ~62 kDa fusion protein seen through SDS-PAGE could not be interpreted as evidence against the disulfide bridge hypothesis, because protein samples were prepared for electrophoresis under reducing conditions; these conditions would have allowed aggregates to disassociate and produce electrophoresis bands at the expected mobility.

In order to test the hypothesis, cellular lysates were analyzed on a substrate gel under two sets of conditions: reducing and nonreducing (Figure 13). If the hypothesis were true, the samples analyzed under nonreducing conditions would have contained multiple selectively

expressed protein aggregates of different sizes. It is also likely that these aggregates of varying weights would not have demonstrated activity. Results indicated that the disulfide bridge hypothesis was false. Under both reducing and nonreducing conditions, a positive signal for active CDA was observed at the mobility of the fusion protein (Figure 14A). Additionally, as evidenced by a comparison of CDA band intensities on the substrate gel stained for active protein (Figure 14A) versus total protein content (Figure 14B), protein activity was proportional to protein quantity under both reducing and nonreducing conditions. It was ultimately concluded that interstrand disulfide bridge formation was not the cause of CDA aggregation and insolubility.

II-E) Attempted mechanical solubilization

Sonication was attempted in place of chemical lysis to extract CDA from bacterial host cell culture. Sonication of the cells did not appear to improve the solubility of CDA; no soluble GST.CDA was detected in sonicates after either one or two rounds of sonication. Little to no soluble CDA was extracted into the soluble fraction of sonicates subjected to methods standard cell lysis. Throughout sonication experiments, a large amount of CDA was expressed, but it remained entirely insoluble (Figure 15).

II-F) Solubilization using sarkosyl

Variations on a solubilization protocol for GST fusion proteins using sarkosyl, adapted from Tao *et al.* 2010 and Park *et al.* 2011, were tested (Figure 16). Sarkosyl was added in varying concentrations to cell lysis buffer. Alternatively, insoluble fractions of lysates (produced with or

without sarkosyl) were soaked in sarkosyl-containing buffer (specific experimental conditions described in Figure 16). Soluble CDA was consistently extracted from lysates produced with 10% sarkosyl in the lysis buffer. Solubilization was most effective when sarkosyl was added to the lysis buffer, and the insoluble fraction of the resulting lysate was then treated with 10% sarkosyl soak (Figure 16). The soluble fraction obtained from this soaking step was used in subsequent attempts to renature and purify CDA.

II-G) Attempted GSH-sepharose purification

GST tags must be properly folded to enable GSH sepharose purification of GST fusion proteins. Because sarkosyl is a denaturant, steps were taken to restore GST to its proper conformation following lysis in the presence of sarkosyl. These renaturation techniques, adapted from Tao *et al.* 2010 and Park *et al.* 2011, involved the dilution of the sarkosyl, along with the addition of Triton and CHAPS to cellular lysates.

The first attempt at GST.CDA renaturation resulted in the extreme reduction in solubility of all proteins except for two, which weighed roughly 62 kDa and 37 kDa (Figure 17). These proteins were identified as GST.CDA and CDA, respectively. Thus, GST.CDA and CDA were more abundant in the soluble fraction of the renatured cellular lysate than any other protein (Figure 17). The reason for this unexpected, partial purification following renaturation is unknown, and results could not be replicated subsequently.

As expected, CDA lacking the GST fusion tag did not bind GSH sepharose, and was consequently removed from the beads in the unbound fraction (Figure 17). However, a detectable amount of CDA.GST remained bound to the beads during all wash steps and was specifically

pulled down from the beads in purified form. Because CDA requires the GST fusion tag in order to bind GSH, the identity of the ~62 kDa protein was verified to be CDA.GST, and the ~37 kDa protein was verified to be CDA.

Although this initial purification attempt was partly successful, purified GST.CDA was eluted in relatively low yield. To reduce to the cleavage of the GST tag from CDA, a fresh cell culture was lysed in the presence of two protease inhibitors: aprotinin and leupeptin. GSH sepharose purification of the lysates was inferior (Figure 18). Protease inhibitors did not prevent cleavage of GST from CDA, and the lane of the electrophoresis gel containing the bound fraction from the pull down displayed an even weaker and far less specific signal at the mobility of GST.CDA (Figure 18).

The decreased yield and specificity of the fusion protein purification of cell lysates prepared with protease inhibitors could be explained by the composition of the starting material, which contained less total protein than that of the previous purification. For unknown reasons, the specific enhancement of CDA in the soluble fraction of renatured lysates was also reduced. Over all, adding protease inhibitors to the lysis buffer of cultured cells failed to improve the yield of purified GST.CDA.

GSH sepharose purification of fresh lysates without protease inhibitors was repeated in an attempt to improve the yield and specificity of CDA purification (Figure 19). Although a somewhat stronger signal was generated at the mobility of GST.CDA in the bound fraction of lysate eluted from the beads, this signal was highly nonspecific; the fusion protein was not effectively isolated from other proteins in the cellular extract (Figure 19). As noted, previous results could not be replicated. It was ultimately concluded that the pGEX.CDA/Rosetta

expression system needed further revision in order to allow the production of soluble CDA that could be effectively purified.

Expression system III: pGEX-4T-2.CDAt in Rosetta2 E. coli

III-A) Rationale for limits of truncated CDA gene

The optimal portion of the splice variant F48E3.8 to include in the truncated gene fragment was determined using Phyre2 software along with Heustis's cross-species alignment of the Nod B (catalytic polysaccharide deacetylase) homology domain (Heustis 2012). Phyre2 uses a database of resolved protein structures to predict the three-dimensional structure formed by an input amino acid sequence. Entry of the full-length *C. elegans* F48E3.8c gene into the database generated a structure that demonstrated high homology to the *Helicobacter pylori* hydrolase hp0310, a putative polysaccharide deacetylase. The primary sequence of this protein was truncated, stretching from Met 57 to Asp 319 of the full-length splice variant F48E3.8c. Further investigation revealed that this truncated sequence conserved the 5 amino acid sequence motifs highlighted in Heustis's polysaccharide deacetylase alignment (Heustis 2012). It was concluded that the truncated CDA gene generated using Phyre2 was more likely to be soluble than the full-length protein, while still retaining the ability to fold into a functional protein. The DNA sequence encoding the truncated protein was shorter than the original by roughly 300 nucleotides, and it contained 75% fewer cysteine residues than the original. It was hypothesized that these qualities would improve the solubility of the CDA.

III-B) Subcloning of the recombinant plasmid

The subcloning scheme utilized to construct the recombinant plasmid pGEX.CDAt was nearly identical to that used for the subcloning of pGEX.CDA (Figure 20). This schematic was carried out successfully.

The designated portion of the F48E3.8c gene was amplified from the previously used yk template, again using primers designed to add a BamHI cut site to the 5' end of the insert, and an XhoI cut site to the 3' end. All 6 PCR reactions containing template DNA generated amplimers that were roughly 800 bp in length, matching the expected size of the 804-nucleotide amplified region of the cDNA (Figure 21A). As before, PCR reactions were completed using the enzyme *Elongase*, allowing for the direct ligation of the amplified, truncated CDA gene (with the added restriction sites) into Promega's pGEM T-easy vector. Following the transformation and extraction of this recombinant plasmid from JM109 competent cells, an analytical digest was completed. The insert was successfully regenerated through BamHI/XhoI digestion, indicating that the ligation had been completed effectively and that the sequences of the cut sites added to the amplimer were accurate (Figure 21B).

Following preparation of the pGEX plasmid and CDAt insert, ligation was completed, and the recombinant plasmid was transformed into JM109 cells and re-extracted from multiple clones for analysis by XhoI/BamHI double digest. The regeneration of the ~800 kb insert once again indicated that the ligation was successful and that the plasmid could be transformed into its final host, Novagen's Rosetta2, for protein expression (Figure 21D).

III-C) Preliminary expression

SDS page analysis of cellular lysates revealed that under standard induction and lysis conditions, cells carrying the new recombinant plasmid expressed large quantities of CDA relative to other proteins (Figure 22). However, this protein was completely insoluble. Manipulating various parameters of the cellular induction (including temperature, duration, and concentration of the inducer) did not appear to affect protein solubility. However, adding 10% sarkosyl to the buffer in which cells were lysed did effectively solubilize the CDAt, regardless of induction conditions (Figure 23).

III-D) Attempted GSH-sepharose purification

Because all soluble, crude protein extracts containing CDAt also contained the denaturant sarkosyl, solution conditions needed to be adjusted so that the CDAt could renature prior to protein purification, without losing its ability to solubilize. This was attempted by diluting the sarkosyl present in the extracts and outcompeting it with a combination of two milder detergents (Triton and CHAPS). The quantity of soluble CDAt present in the protein sample soluble fraction of diluted lysate containing Triton and CHAPS was not successfully purified using GSH-sepharose beads (Figure 24). Regardless of the conditions in which cells were induced, minimal protein was eluted from the beads. Further, the protein composition of this eluted fraction was very nonspecific, and did not appear to contain a specifically enhanced quantity of CDAt (Figure 13). The GSH purification of construct III lysates was ultimately unsuccessful.

III-E) Evaluation of the construct

Construct III must be subjected to further laboratory testing before evaluation of its potential to produce active CDAt in solution is complete. Thus far, this construct has been found to specifically express large quantities of soluble CDA when cells are lysed in the presence of sarkosyl.

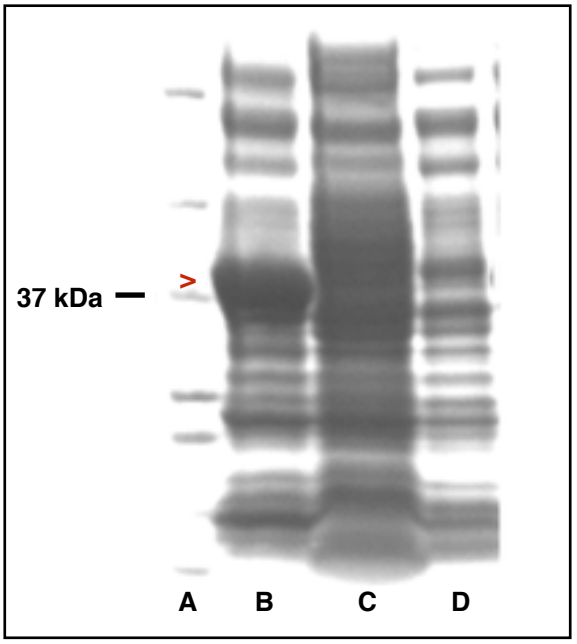

Figure 1. Induction specifically enhanced CDA in whole lysate. Gel electrophoresis of pET22b with CDA insert transformed into Rosetta(DE3) *E. coli*. Induction was performed at 37^{0}C with 1 mM IPTG. A) Molecular weight markers; B) Induced whole lysate; C) Uninduced whole lysate; D) Soluble fraction of induced lysate. Arrow indicates putative CDA protein band.

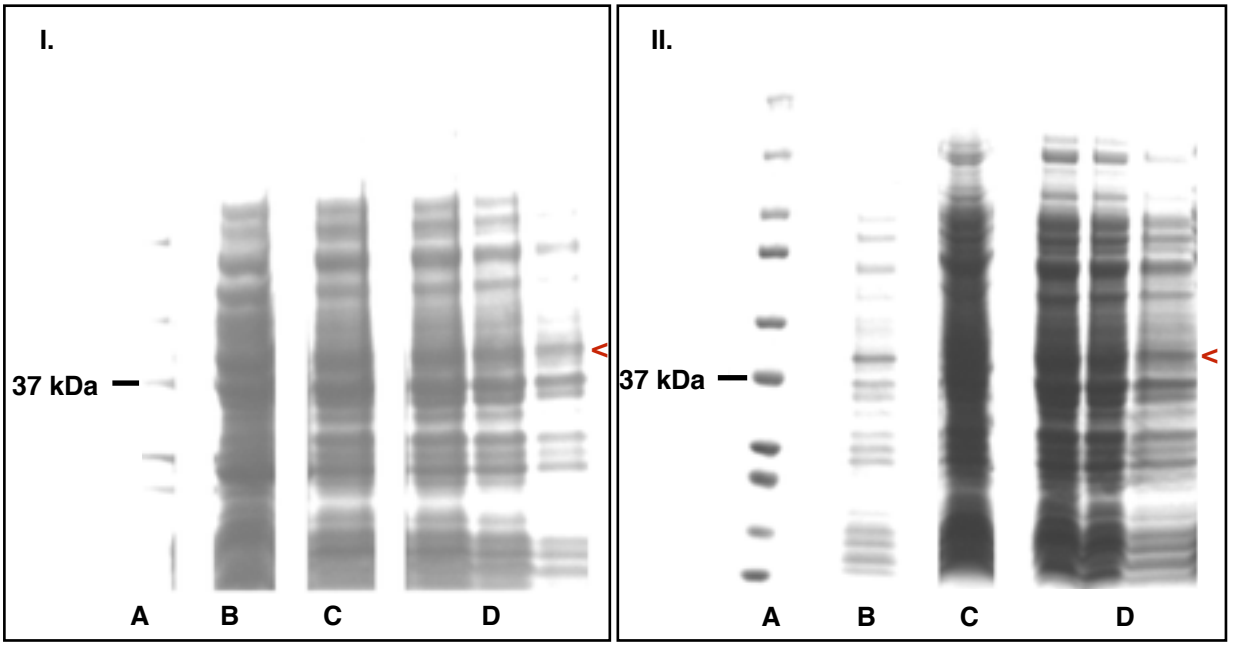

Figure 2. Cold inductions did not specifically enhance CDA in whole lysate or soluble fraction. Cell extracts of Rosetta(DE3) *E. coli* carrying pET22b plasmid with CDA gene insert. Panel I represents cells induced with 0.4 mM IPTG; Panel II represents cells induced with 1.0 mM IPTG. All inductions were performed at 18^0 C. A) Molecular weight markers; B) Uninduced pooled lysate; C) Induced pooled lysate; D) Induced soluble fraction of lysate loaded in decreasing concentrations from left to right. Red arrows indicate putative CDA protein bands.

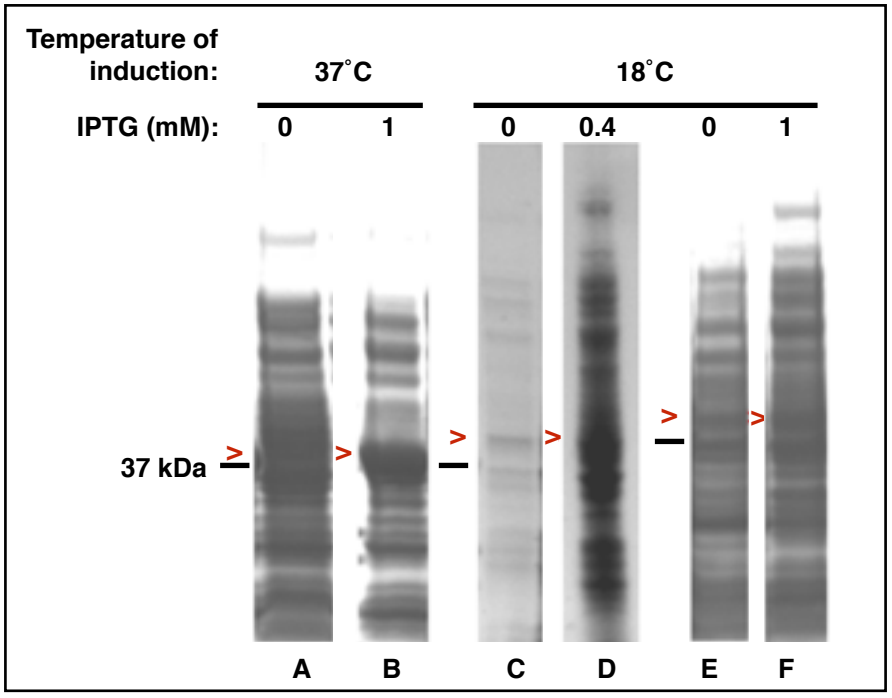

Figure 3. 37^{0}C induction specifically enhanced CDA whereas 18^{0}C did not. Coomassie Blue stained whole cell lysates of Rosetta DE3 *E. coli* carrying pET22b plasmid with eukaryotic CDA gene. Samples A, C, and E were uninduced. Samples B, D, and F were induced at temperatures and IPTG concentrations as indicated. Black bars indicate position of 37 kDa weight marker. Red arrows indicate position of putative CDA protein bands. 37^{0}C induction significantly enhanced CDA protein band whereas 18^{0}C induction did not appear to enhance CDA at either inducer concentration.

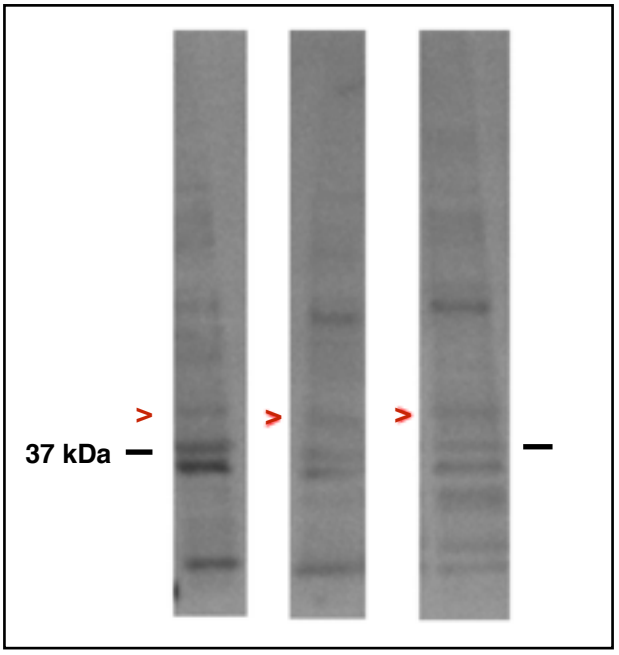

Figure 4. CDA protein was only partially retained by the His-bind column. Western blotting with anti His-6 primary antibody (Sigma H1029) and secondary antibody (Ram-Px). All fractions were concentrated by acetone precipitation. From left to right: original lysate before preparation with binding buffer for passage through column; column wash with 1x wash buffer (unbound fraction); column elute with 1x elute buffer (bound fraction). Red arrows indicate putative CDA protein bands.

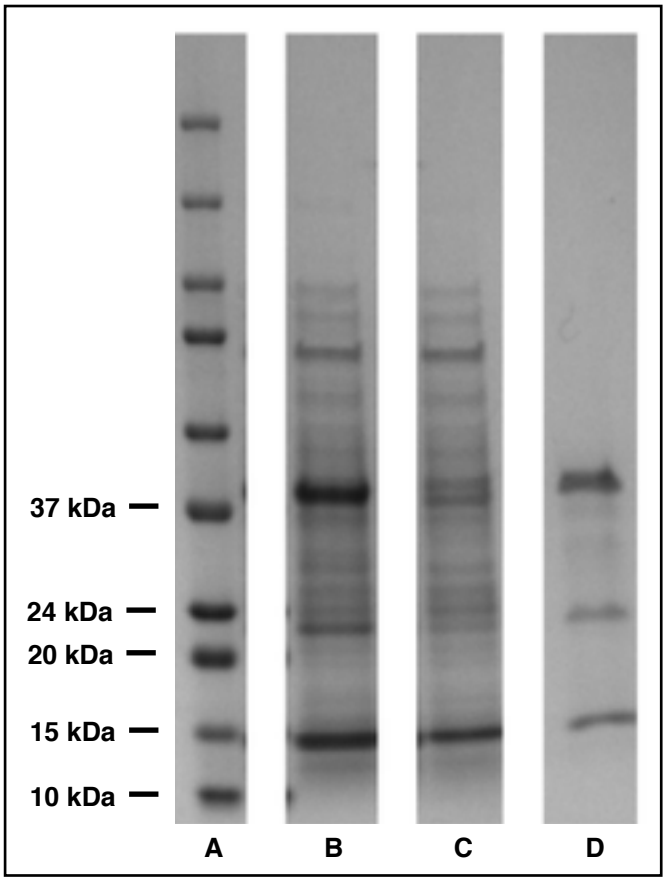

Figure 5. Co-MAC purification cartridge successfully isolated CDA from the original lysate. Recombinant protein in *E. coli* lysate was denatured with urea before purification. A) Molecular weight markers; B) Starting material; C) Flow-through (unbound fraction) of prepared starting material through column; D) Protein eluted with imidazole (bound fraction). Gel was stained with Coomassie Blue following electrophoresis. (Column purification and gel electrophoresis completed by William McNeil.)

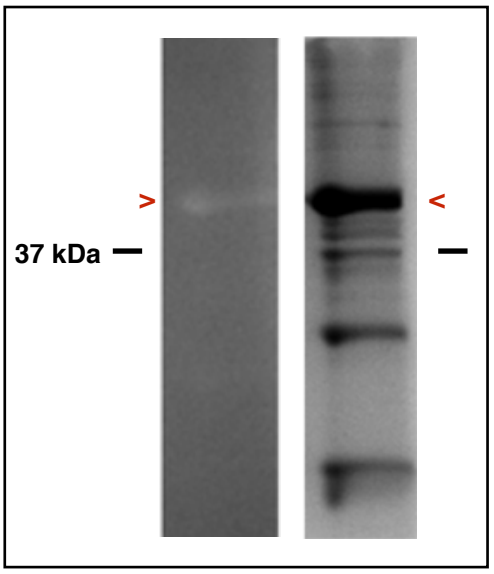

Figure 6. Concentrated, purified CDA was renatured and regained enzymatic activity. Red arrows indicate CDA protein bands. A) Substrate gel containing glycol chitin; stained with calcofluor after renaturation in 1% Triton-X; visualized on UV light. B) Same substrate gel stained in Coomassie blue; visualized on white light.

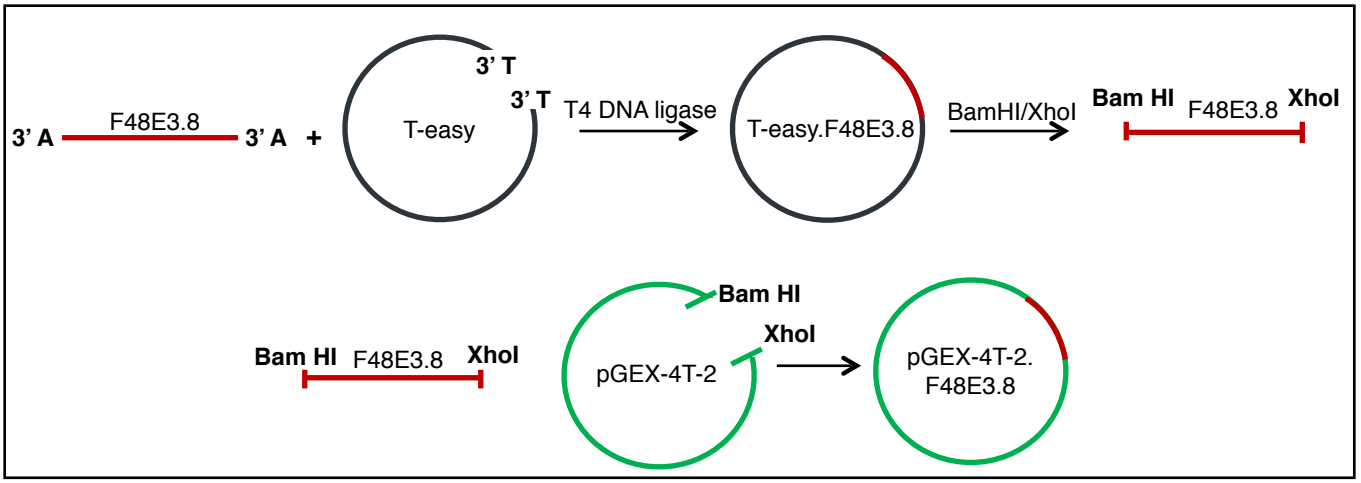

Figure 7. Subcloning scheme utilized to construct recombinant plasmids for expression systems II and III. F48E3.8 represents isoform c of one *C. elegans* CDA gene. System II utilized the full splice variant whereas system III utilized a truncated version of the splice variant. F48E3.8 with BamHI and XhoI restriction sites was created through a PCR amplification, ligated into T-easy, and cleaved out of T-easy to generate sticky ends. The gene was then ligated into the BamHI/XhoI-digested, dephosphorylated pGEX-4T-2 vector.

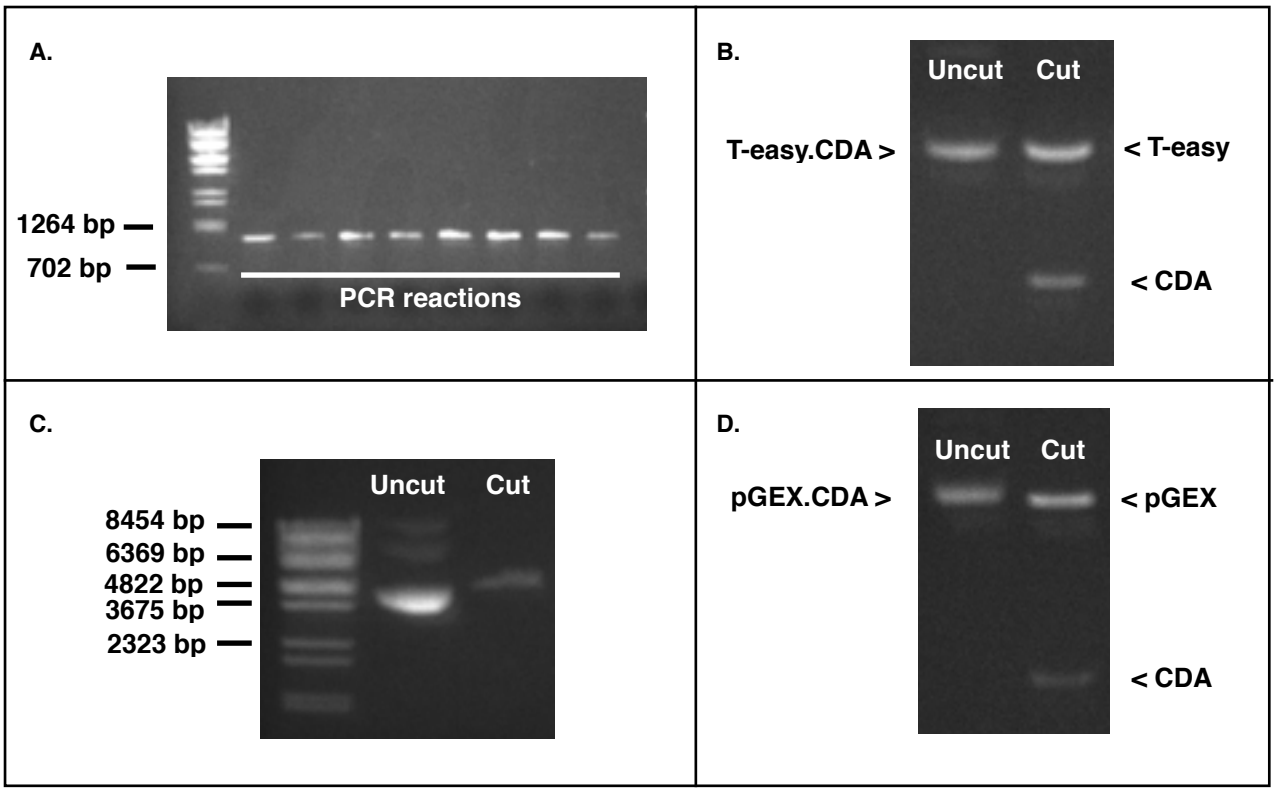

Figure 8. Agarose DNA electrophoresis illustrated progress of expression system II subcloning process. A) PCR amplimers of CDA insert all near expected mobility (~1100 bp); control reaction without template yielded no amplimer (data not shown). B) T-easy.CDA extracted from JM109 cells. Insert was regenerated through BamHI/XhoI double digest ("cut"), indicating effective ligation. C) Shifted mobility between uncut, supercoiled pGEX plasmid (~4900 bp) and XhoI/BamHI double digest indicated effective cleavage of vector. D) pGEX.CDA extracted from JM109 cells; ligation was effective.

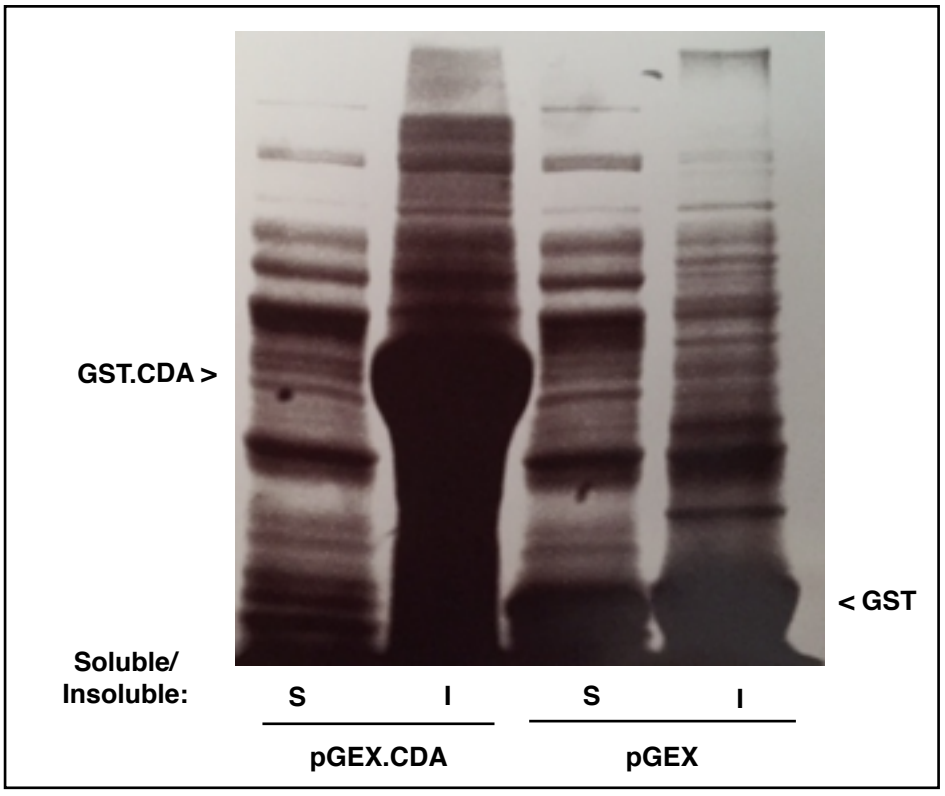

Figure 9. pGEX.CDA/Rosetta lysates induced and lysed in standard conditions. S indicates soluble fraction of lysate; I indicates insoluble fraction. 10% polyacrylamide gel stained in Coomassie Blue-R250.

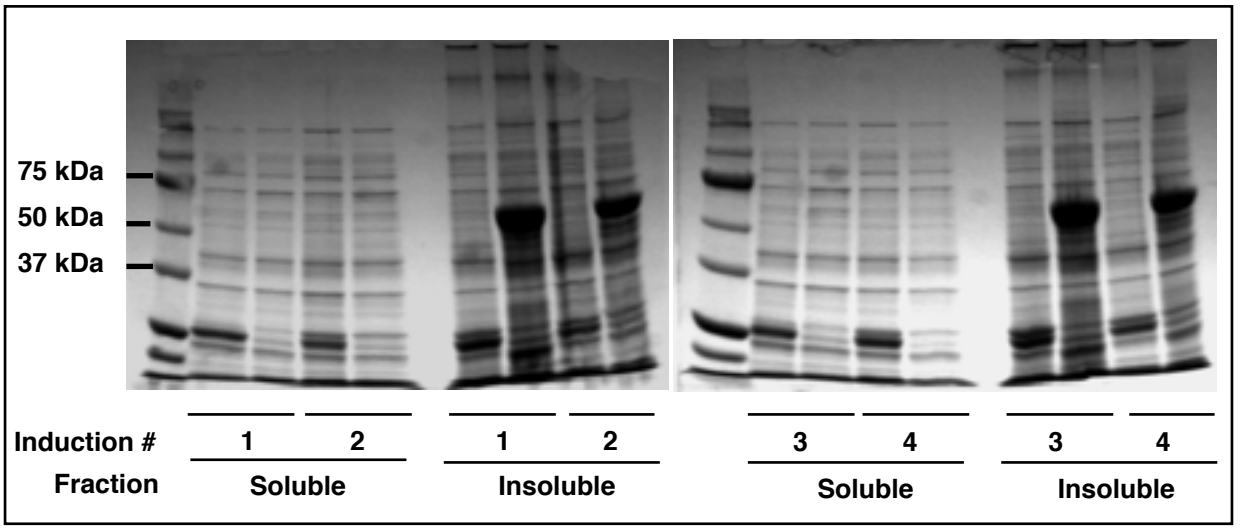

Figure 10. Protein gel electrophoresis of pGEX.CDA/Rosetta lysates (construct II). Induction conditions: 1) 1 mM IPTG, 37˚C, 2 h; 2) 1 mM IPTG, 16˚C, 21 h; 3) 0.1 mM IPTG, 37˚C, 2h; 4) 0.1 mM IPTG, 16˚C, 21h. The lefthand lane in each pair contains lysates of cells carrying pGEX without CDA; the righthand lane contains induced lysates. Regardless of variations in induction temperature, duration, and inducer concentration, CDA consistently localized to the insoluble fraction of cellular extracts.

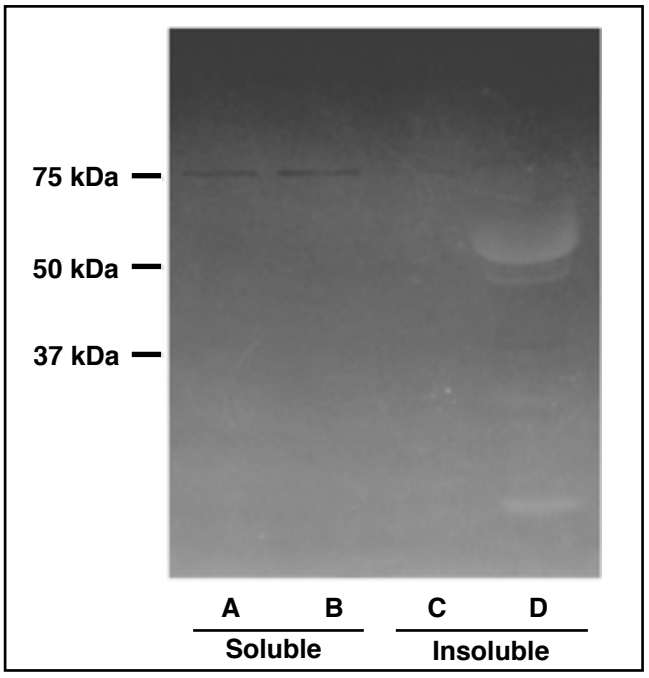

Figure 11. Electrophoresis of Rosetta cell lysate transformed with empty pGEX vector (A; C) and recombinant pGEX.CDA (B; D). Lanes (A) and (B) contain soluble fractions of lysate; (C) and (D) contain insoluble fractions.10% polyacrylamide gel embedded with glycol chitin. Calcofluor white fluorescent staining specific for chitosan. Bright band appeared at expected mobility of GST.CDA fusion protein (~62 kDa).

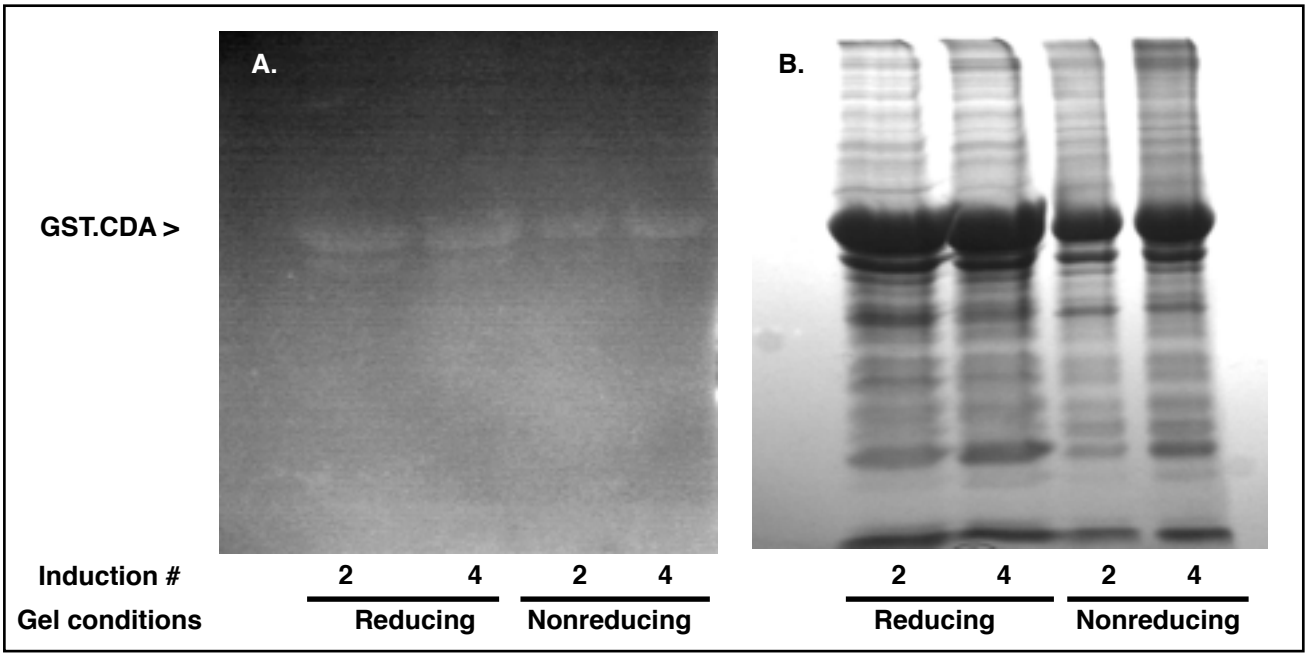

Figure 12. Inappropriate disulfide bridge formation was not responsible for CDA insolubility. Protein gel electrophoresis of pGEX.CDA/Rosetta (construct II) lysates under reducing and nonreducing conditions. Induction conditions: 2) 1 mM IPTG, 16˚C, 21 h; 4) 0.1 mM IPTG, 16˚C, 21h. A) Substrate gel embedded with glycol chitin, incubated in Triton X-100 after electrophoresis, and stained with Calcofluor white. Visualized on UV light. Bright bands indicate soluble CDA. B) Same gel stained in Coomassie Blue R-250, visualized on white light. Aggregation did not occur in the absence of beta-mercaptoethanol, and enzymatic activity was proportional to quantity of CDA under both reducing and nonreducing conditions.

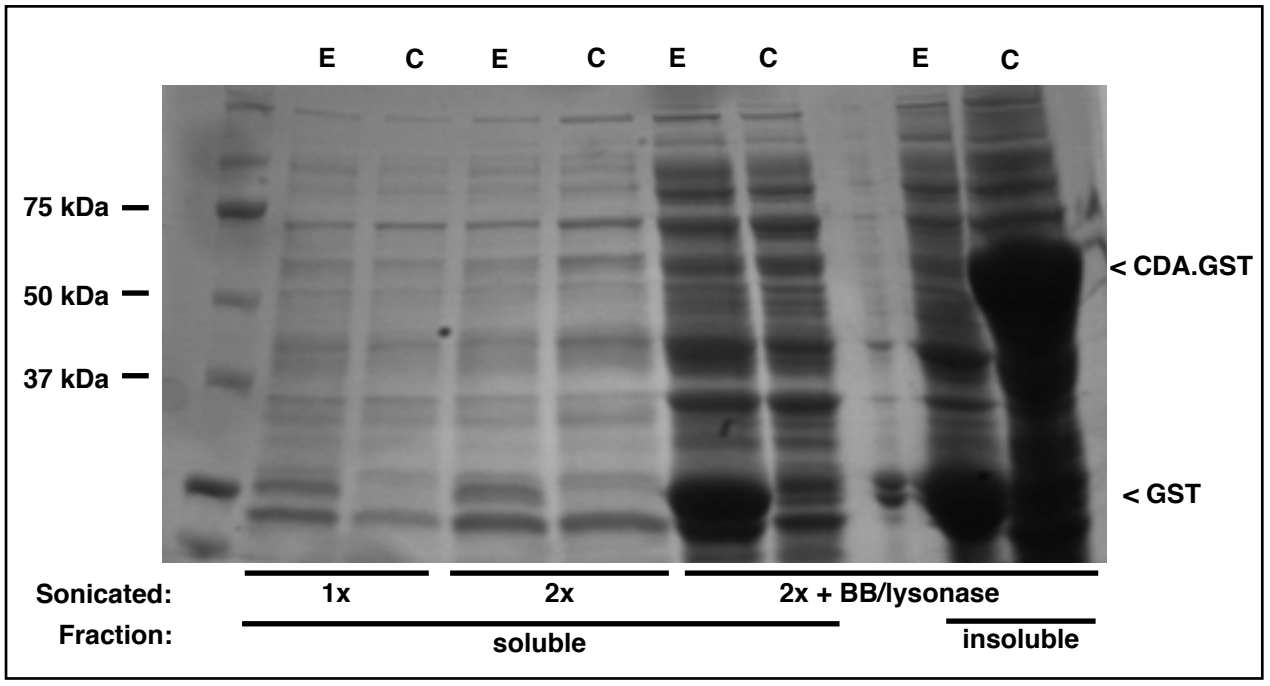

Figure 13. SDS-PAGE analysis of sonicated and sonicated/chemically lysed pGEX.CDA/ Rosetta extracts. "E" indicates cells carrying vector without insert; "C" indicates cells carrying vector with CDA insert. Number of rounds (sets of 6, 15-second pulses at 60-second intervals) of sonication indicated. Sonication failed to solubilize CDA.GST (~62 kDa). 10% polyacrylamide gel stained in Coomassie Blue-R250.

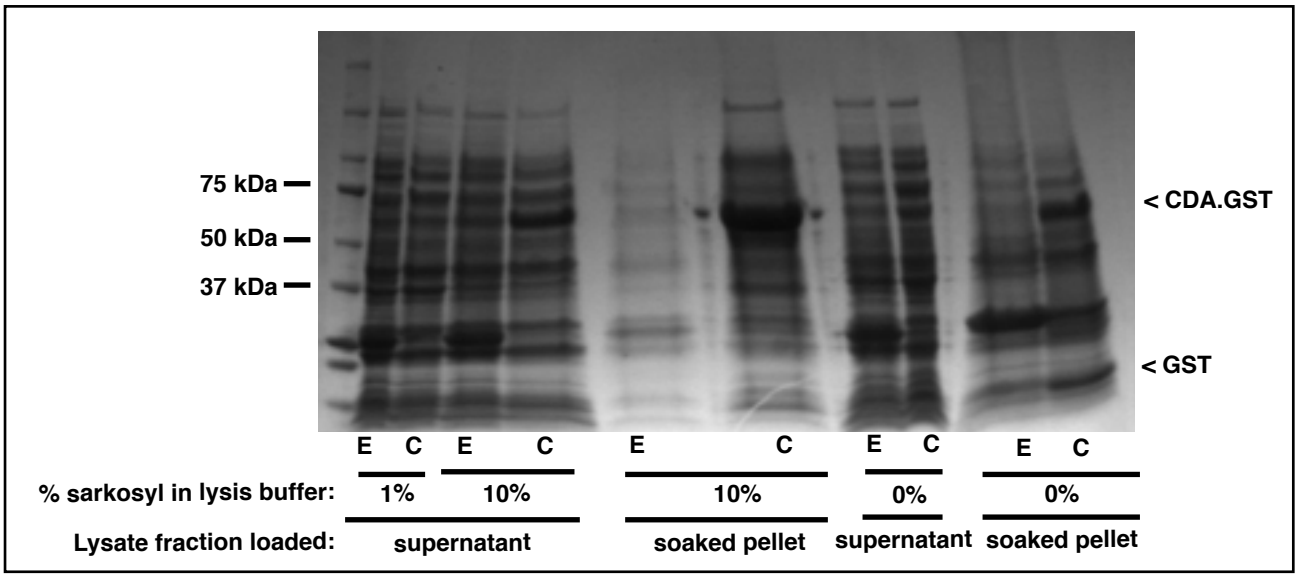

Figure 14. SDS-PAGE analysis of pGEX.CDA/Rosetta extracts . "E" indicates cells carrying vector without CDA insert; "C" indicates cells carrying recombinant vector. Lysis buffer contains varying concentrations of sarkosyl (% weight by volume) as indicated. "Supernatant" indicates soluble fraction of lysate directly following lysis; "soaked pellet" indicates soluble fraction extracted from inclusion bodies from lysis, following overnight soak in 10% sarkosyl/ ST buffer. 4-15% polyacrylamide gel stained in Coomassie Blue-R250.

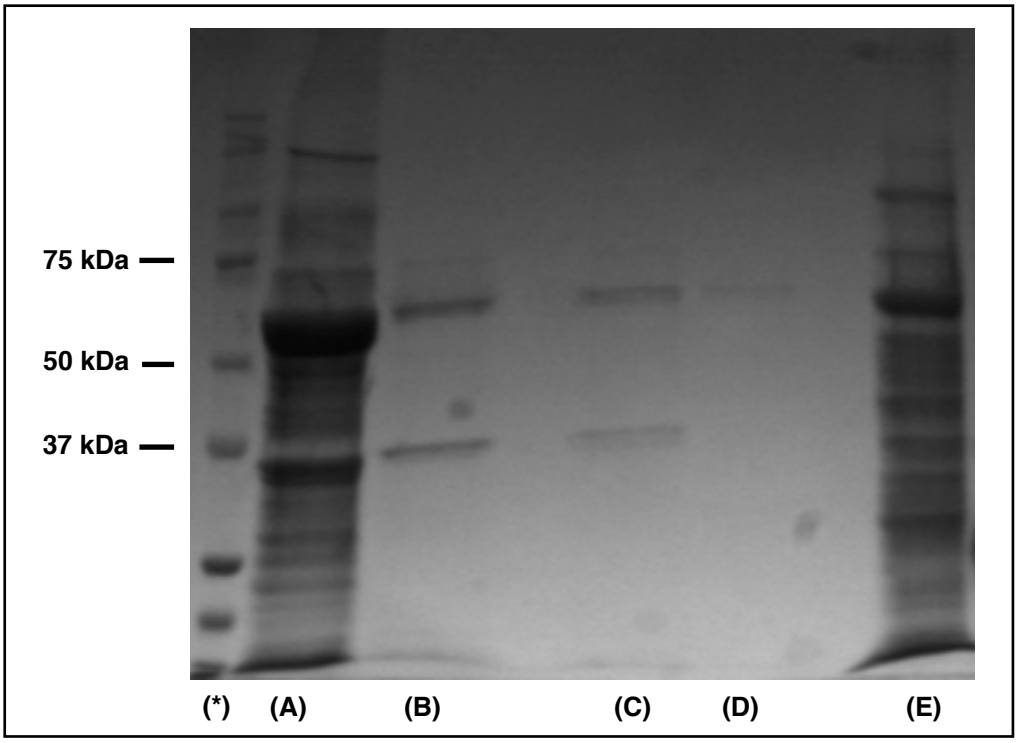

Figure 15. GSH sepharose purification fractions of cellular lysate derived from Rosetta cells, transformed with pGEX.F48E3.8c recombinant plasmid. Cells were lysed in BugBuster with lysonase and 10% sarkosyl; insoluble lysate was soaked overnight in 10% sarkosyl/ST buffer. Resulting soluble (A) and insoluble (E) fractions were analyzed via electrophoresis. Supernatants were diluted to 2% sarkosyl in 4% triton and 40 mM chaps (B). Remaining lanes were loaded as follows: (C) GSH-Sepharose unbound fraction; (D) GSH-Sepharose bound fraction; (*) indicates molecular weight markers. 10% polyacrylamide gel stained in Coomassie Blue R-250.

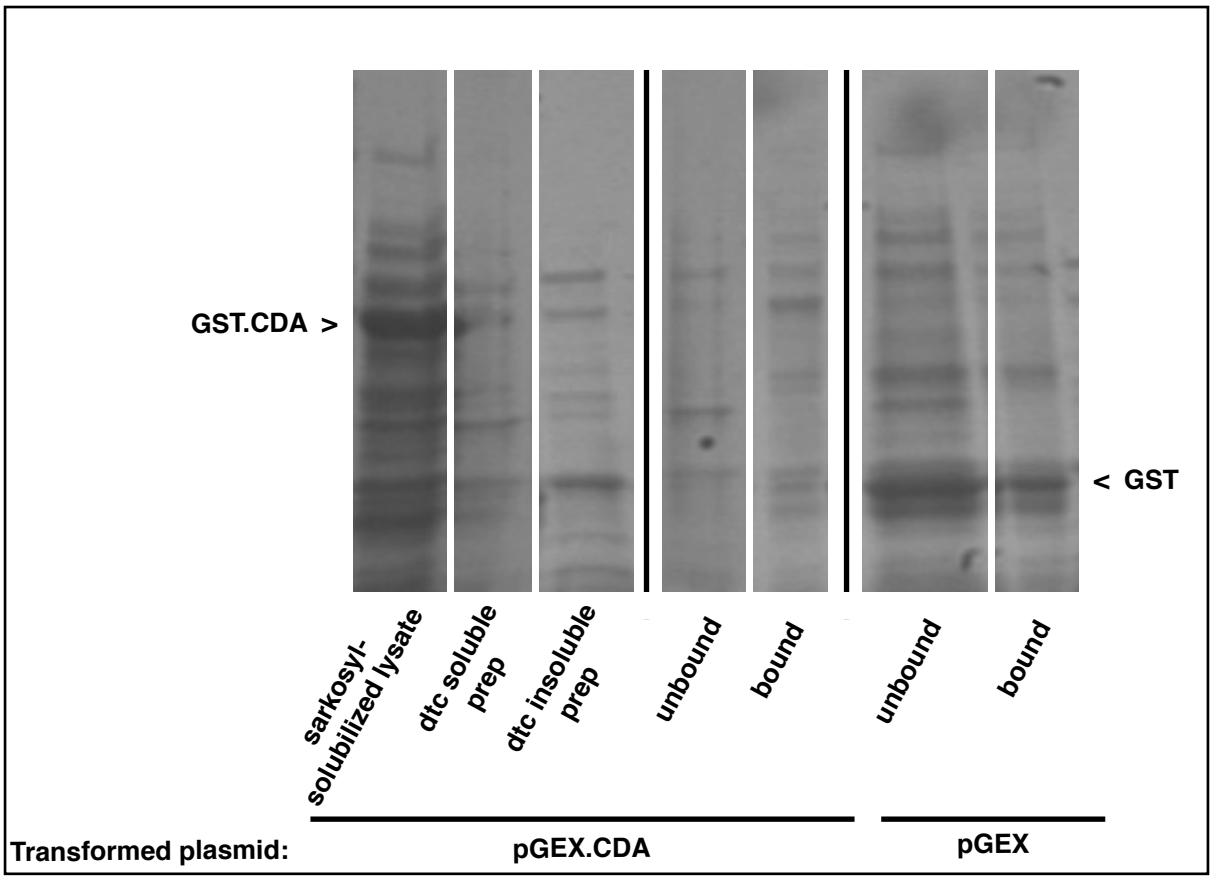

Figure 16. GSH sepharose purification of pGEX.CDA/Rosetta cell extracts. Soluble lysate containing 10% sarkosyl was diluted to 2% sarkosyl in 4% triton and 40 mM CHAPS ("dtc") prior to incubation with beads; soluble and insoluble fractions were analyzed separately here. Empty pGEX vector in Rosetta was also purified as a positive control. Somewhat nonspecific binding occurred in both experimental and control trials. SDS-PAGE on 4-15% polyacrylamide gel, stained in Coomassie Blue R-250.

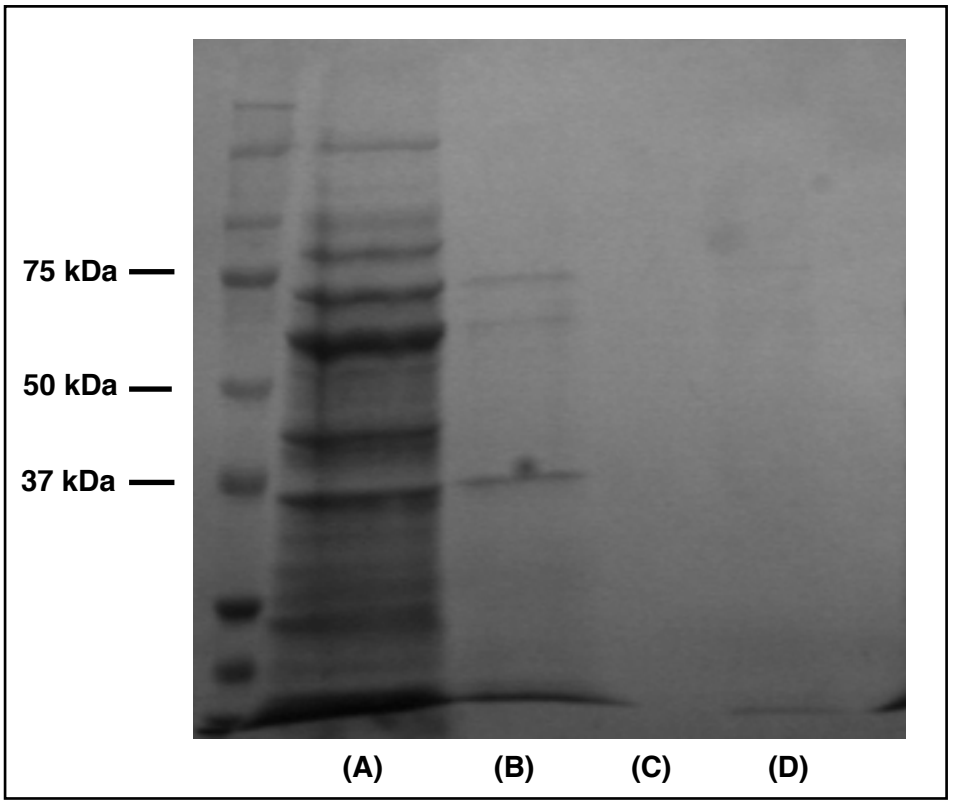

Figure 17. GSH sepharose purification of cellular lysate derived from Rosetta cells, transformed with pGEX.F48E3.8c recombinant plasmid. Cells were lysed in BugBuster with lysonase ,10% sarkosyl, 10 ug/mL aprotinin and 10 ug/mL leupeptin; insoluble lysate was soaked overnight in 10% sarkosyl/ST buffer. Supernatants (A) were diluted (B) to 2% sarkosyl in 4% triton and 40 mM chaps. Remaining lanes were loaded as follows: (C) unbound fraction; (D) bound fraction. SDS-PAGE on 10% polyacrylamide gel stained in Coomassie Blue R-250.

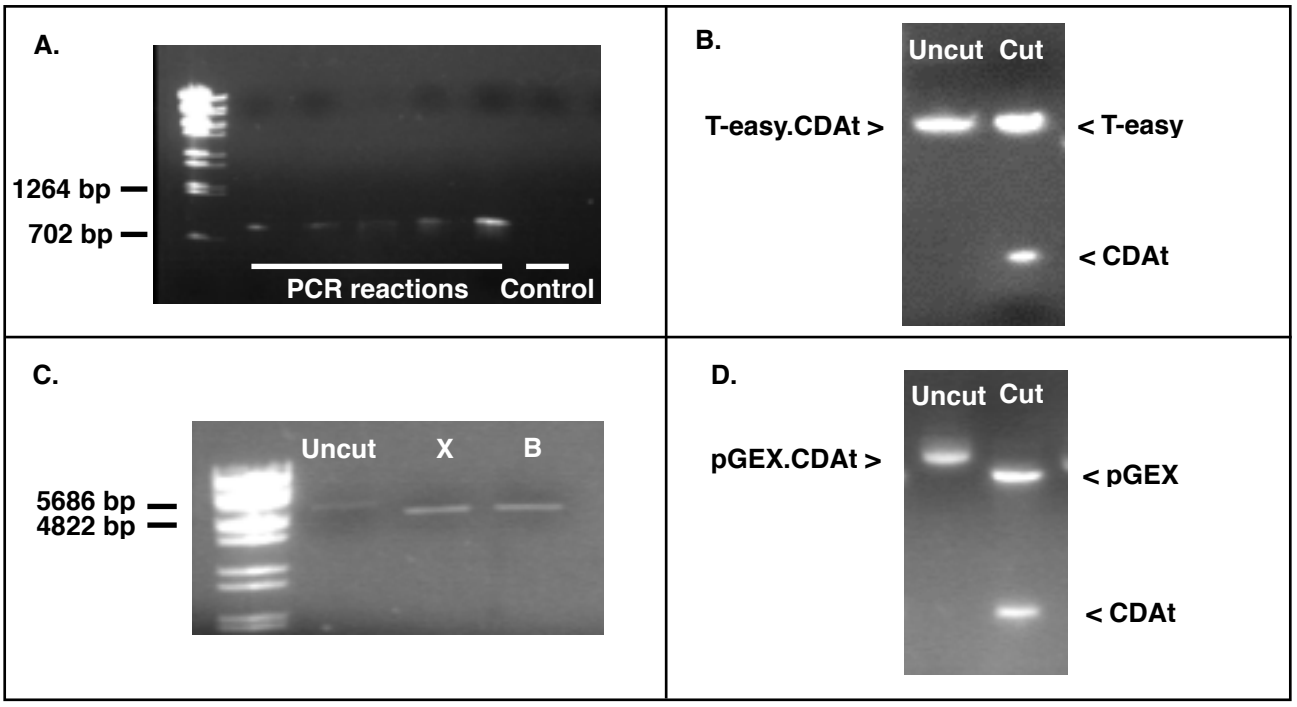

Figure 18. Subcloning truncated F48E3.8c cDNA for expression system III. A) PCR amplimers of truncated CDA insert all generated signal near expected mobility (~804 bp); control reaction contained no template. B) T-easy.CDAt extracted from JM109 cells. Lower molecular weight band appeared in lane containing BamHI/XhoI double digest ("cut"), not in lane containing whole plasmid ("uncut"), indicating effective ligation. C) Agarose gel electrophoresis of uncut pGEX plasmid (~4900 bp) and fractions of two XhoI/BamHI sequential double digests prior to addition of second enzyme (lane X digest contains only XhoI; lane B digest contains only BamHI). Shifted mobility between uncut and cut plasmids indicated effective cleavage of vector. D) pGEX.CDAt extracted from JM109 cells; ligation was effective.

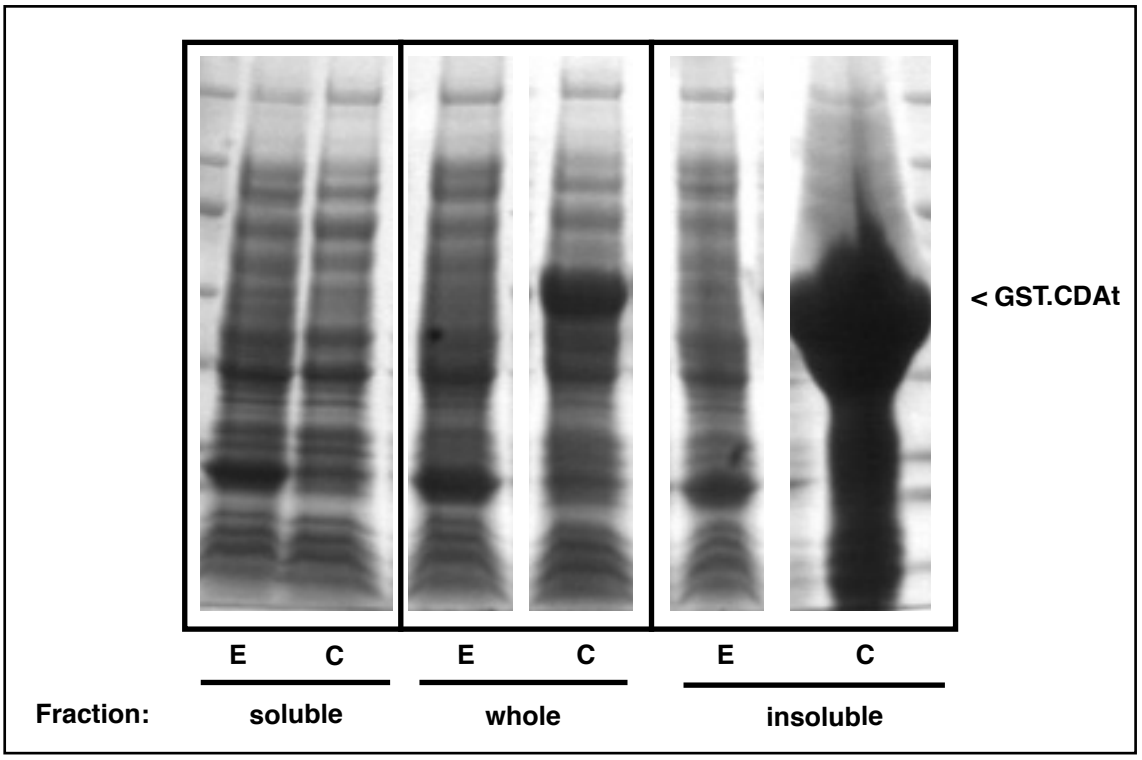

Figure 19. Fractions of induced and uninduced extracts of Rosetta2 *E. coli* cells carrying recombinant pGEX.CDAt plasmid. E represents cells carrying pGEX without CDAt; C represents cells carrying recombinant plasmid. Heavy specific enhancement of CDA visualized entirely in whole lysate and insoluble fraction. SDS-PAGE on 4-15% polyacrylamide gradient gel stained in Coomassie Blue R-250.

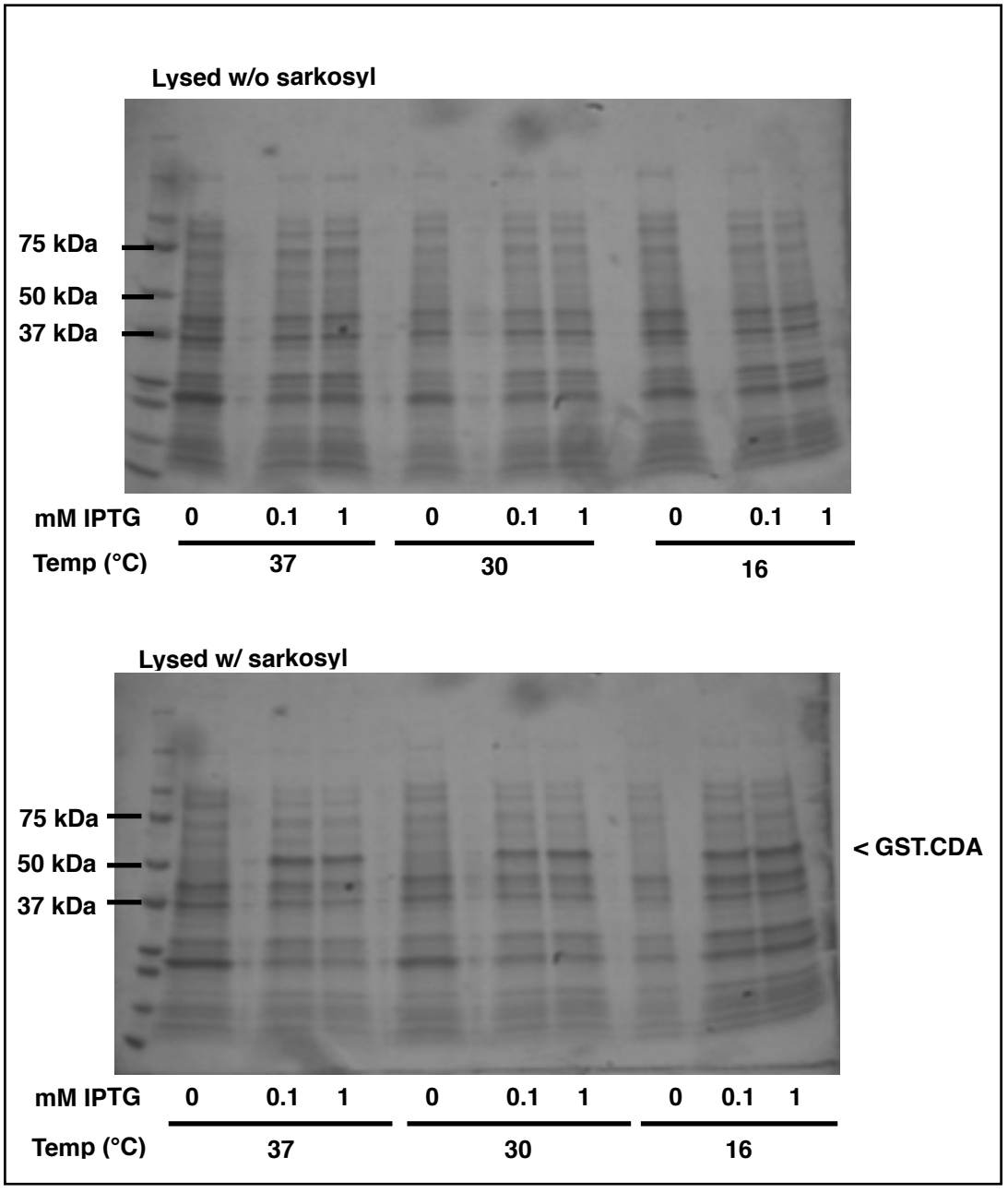

Figure 20. Soluble fractions of pGEX.CDAt/Rosetta2 cell extracts induced at varying temperatures and concentrations of IPTG. Upper gel displays extracts lysed without sarkosyl; lower gel extracts were lysed in presence of sarkosyl. Little variation occurred across induction conditions. Only induced extracts lysed with sarkosyl expressed CDA. 4-15% polyacrylamide gel stained in Coomassie Blue R-250.

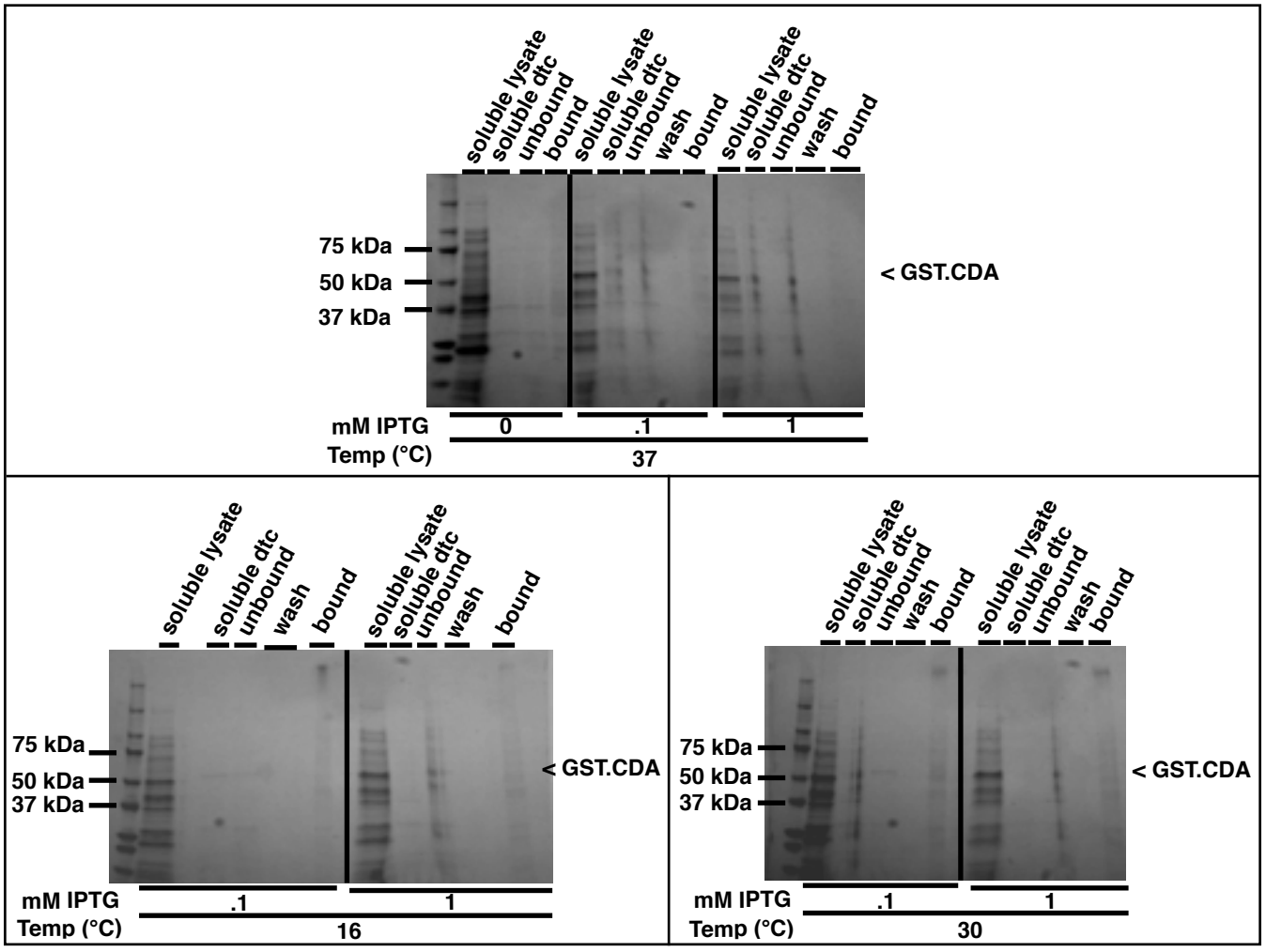

Figure 21. SDS-PAGE analysis of GSH purification of pGEX.CDAt/Rosetta2 lysate. Cells were induced at varying temperatures and concentrations of inducer (IPTG) as indicated. "Soluble lysate" represents lysates solubilized with sarkosyl. Lysates were diluted to 2% sarkosyl with 4% triton and 40 mM CHAPS in ST buffer ("soluble dtc" represents the soluble fraction of this renatured lysate) prior to incubation with beads. Renaturation resulted in low yield of soluble CDA; purification was extremely low-yield and nonspecific. 4-15% polyacrylamide gel stained in Coomassie Blue R-250.

Discussion:

In this study, I investigated the potential of three distinct *C. elegans* F48E3.8c bacterial expression systems to generate nematode CDA suitable for activity analysis. I subjected all three constructs to parallel protein expression and purification steps in an effort to isolate soluble, active CDA.

Ineffectiveness of Varying Induction Conditions

In the cases of all three constructs, preliminary expression of CDA resulted in high levels of CDA synthesis in comparison to other proteins. However, this protein consistently localized to the insoluble fraction of cellular lysates. It was ultimately found that manipulating various parameters of cellular induction, including temperature, duration, and concentration of inducer, generally did not affect CDA solubility. In fact, the only induction condition that appeared to affect protein production at all was temperature; lowering the temperature of the culture medium reduced the net yield of CDA over all. This was not a useful result for protein solubilization.

Problems Associated with Denaturation

All three constructs generated CDA that could only be extracted into the soluble fraction of cellular lysates with the addition of some form of denaturant. In order to solubilize CDA derived from pET22b.CDA/Rosetta(DE3) (construct I) in sufficient quantity for purification, urea was used as a denaturant. pGEX.CDA/Rosetta (construct II) and pGEX.CDAt/Rosetta2 (construct III) lysates both required the presence of sarkosyl to solubilize CDA. A recurring issue

encountered following protein solubilization via denaturation was that CDA proved extremely difficult to renature with high yield.

The column purification techniques utilized to isolate CDA from construct I extracts allowed for the purification of denatured protein, because the C-terminal His_6 tag on the recombinant protein did not need to be in any particular conformation to chelate around the cations in the column cartridges. Following purification, however, efforts to dialyze the denaturant while maintaining solubility were unsuccessful, and the protein was not successfully renatured (McNeil 2014). The effort to renature purified protein using construct I was ultimately abandoned in favor of investigating new plasmid vectors to replace pET22b. Unlike the proteins expressed using construct I, the fusion proteins generated by constructs II and III needed to be restored to renaturing conditions prior to purification. This is because the C-terminal GST fusion tag attached to the CDA produced by pGEX constructs must be in its correct conformation in order to enable purification via GSH-sepharose beads. The techniques used to renature protein derived from constructs II and III significantly reduced the yield of CDA.GST by causing it to precipitate out of solution. Regardless of whether CDA was renatured before or after purification, as dictated by the expression system, the preparation of purified, active, and soluble protein from all three constructs was either halted or severely hindered at the renaturation step. This suggests that in order to accomplish the end goal of generating soluble CDA in its proper conformation, a protocol should be designed either to optimize its renaturation, or to solubilize the protein without a denaturation step.

To address the former suggestion, optimizing CDA renaturation following solubilization via denaturation may require the presence of solubilized substrate, such as glycol chitin. The

substrate gel assay adapted from Trudel and Asselin (1990) has been found to demonstrate *C. elegans* CDA activity by diluting out the denaturant (SDS) in the presence of a milder detergent (Triton X-100) and substrate (glycol chitin). This process would be best replicated in solution if substrates are present during renaturation, which may encourage accurate protein folding.

Even if experimental outcomes continue to suggest that the development of a reliable CDA renaturation technique using transformed *E. coli* to produce *C. elegans* CDA may be unlikely, it may be possible to design a construct that consistently yields a large quantity of soluble protein. Then, even if the yield of purification is relatively low, considerable quantities of purified CDA would still be produced over all. This is most likely to be accomplished using an expression system that renders denaturation unnecessary.

Multiple expression systems have failed to achieve this end. Neither the full-length F48E3.8c gene, nor the truncated version designed to reduce the number of hydrophobic amino acid residues, translated to soluble CDA following subcloning into the pGEX-4T-2 vector. McNeil's attempt to generate soluble CDA in the absence of denaturant using the Novagen vector pET30a, which lacks the signal peptide found in pET22b, met with similar results (McNeil 2014). Soluble fractions of pET30a.CDA lysates generated without denaturant did not contain CDA (McNeil 2014). Experimentation with a greater variety of vectors and bacterial host strains should be pursued in order to find an expression system that yields soluble CDA without requiring denaturation.

Construct III: Further Assessment

Of all three constructs studied, Expression System III demonstrated the most potential to be used for the generation of soluble protein. Based on preliminary experiments, the soluble fraction of the renatured lysate appeared to contain equivalent proportions of CDA relative to other proteins both before and after renaturation. Acetone precipitation of renatured samples should be performed to verify this observation, so that the concentration is equivalent to that of the starting material and the intensities of the bands can be directly compared. If the specificity of the subsequent purification can be improved, this expression system may yield soluble, purified CDA.

One possible cause of the lack of specificity observed in preliminary purification of construct III lysates is that the GSH sepharose stock used in the experiment contained fragmented beads that trapped proteins nonspecifically during centrifugation (Figure 13). Although the beads were washed in sterile PBS prior to sampling of the "bound" fraction, the purifications performed thus far did not involve glutathione elution. Following the binding and wash steps of the purification, the beads were simply heated in reducing sample buffer to obtain the "bound" fraction. This harsh elution process likely released proteins trapped between the beads that did not actually have the capacity to bind GSH. This would explain why the bound fractions analyzed by SDS-PAGE presented bands at such a wide range of mobilities (Figure 13). To remedy this source of ambiguity, the purifications performed thus far should be repeated using a fresh stock of GSH sepharose, or using glutathione to release the bound fraction of cellular lysates rather than simply heating the beads in reducing sample buffer. Additionally,

column purification could be attempted as opposed to the pull-down method, which may allow for higher levels of specific binding and an overall greater yield of purified protein.

Alternative Routes to CDA Activity Analysis

If CDA cannot be solubilized, it may be possible to proceed with activity assays without producing soluble CDA. Insoluble CDA could simply be purified in bulk through substrate gel electrophoresis under reducing conditions. The gel could be stained for total protein with a nonspecific, non-denaturing stain, and the band running at the mobility of CDA could be excised from the gel. Crudely purified protein could then be transferred from the excised gel to nitrocellulose membrane. Strips of nitrocellulose bearing CDA could be placed into a reaction mix containing soluble substrate, and evidence of enzymatic activity could be quantified using an indicator such as the organic reagent fluorescamine. Nitrocellulose binding the CDA could potentially even be ground to a powder and then permitted to react with substrate.

E. coli has successfully been utilized as an expression host for recombinant CDAs derived from multiple species, including the bacterium *Streptomyces lividans* (Tokuyasu *et al.* 1999), the fungus *C. lundemuthanium* (Blair *et al.* 2006) and the insect *Mammestra configurata* (although this last study did not require protein solubilization) (Toprak *et al.* 2008). However, if it continues to become apparent that *C. elegans* CDA expressed in a bacterial expression system is unlikely to be soluble without some form of denaturation, and if renaturation similarly continues to present challenges, it may become necessary to explore the option of choosing a new host cell line. The *C. elegans* CDA gene might be cloned into a eukaryotic cell line, such as

insect cells or CHO cells. It is possible that the translational machinery native to the host cell strains used across these studies are simply not conducive to the generation of nematode CDA.

Concluding Remarks

Based on what is known from Heustis *et al.* (2012) about the localization of F48E3.8 in the developing embryonic pharynx, the product of this gene must be secreted upon synthesis, enabling the enzyme to localize to the chitinous pharyngeal lining and participate in chitin metabolism. If CDA were insoluble, its transport to the extracellular, acellular, chitinous structures in the nematode would be complicated and metabolically expensive. It is thus highly unlikely that *C. elegans* CDA precipitates immediately upon synthesis in its native environment. By pursuing avenues of protein expression that evade the need to denature CDA prior to purification, it is likely that an expression system can be designed to generate soluble enzyme for activity analysis. Furthering our understanding of *C. elegans* CDA activity will help us to define CDA function in nematodes at large, and ultimately to develop therapeutics that can be used to treat parasitic nematode infections.

References:

African Programme for Onchocerciasis Control. [Internet]. World Health Organization; 2015. Available from http://www.who.int/apoc/onchocerciasis/en/.

Arakane Y, Dixit R, Begum K, Park Y, Specht CA, et al. (2009) Analysis of functions of the chitin deacetylase gene family in Tribolium castaneum. Insect Biochem Mol Biol 39: 355-365

Blair DE, Hekmat O, Schüttelkopf AW, Shrestha B, Tokuyasu K (2006). Structure and Mechanism of Chitin Deacetylase from the Fungal Pathogen *Colletotrichum lindemuthianum.* Biochemistry 45: 9416-9426.

Blair DE, Schuttelkopf AW, MacRae JI, van Aalten DMF (2005) Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. PNAS 102: 15429 – 15434

Carbohydrate Esterase Family 4. [Internet]. Carbohydrate-Active Enzymes; 2015. Available from: http://www.cazy.org/CE4.html.

Chirunomula M (2013) Utilizing fluorescamine to monitor chitin deacetylase activity *in vitro*. (Unpublished work at Tufts University.)

Deng DM, Urch JE, Cate JM, Rao VA, van Aalten DMF, Crielaard W (2009). *Streptococcus mutans* SMU.623c Codes for a Functional, Metal-Dependent Polysaccharide Deacetylase That Modulates Interactions with Salivary Agglutinin. Journal of Bacteriology 191: 394-402

Heustis RJ, Ng HK, Brand KJ, Rogers MC, Le LT, Specht CA, Fuhrman JA (2012) Pharyngeal polysaccharide deacetylases affect development in the nematode C. elegans and deacetylate chitin *in vitro*. PLoS ONE 7: 1-12

Kafetzopoulos D, Thireos G, Vournakis JN, Bouriotis V (1993). The primary structure of a fungal chitin deacetylase reveals the function for two bacterial gene products. Proc. Natl. Acad. Sci. USA 90: 8005-8008

Lentz CS, Halls V, Hannam JS, Niebel B, Strübing U et al. (2013) A Selective Inhibitor of Heme Biosynthesis in Endosymbiotic Bacteria Elicits Antifilarial Activity In Vitro. Chemistry and Biology 20: 177-187

Luschnig S, Bätz T, Armbruster K, Krasnow MJ (2006). Serpentine and vermiform Encode Matrix Proteins with Chitin Binding and Deacetylation Domains that Limit Tracheal Tube Length in Drosophila. Current Biology 16: 186-194.

Lymphatic Filariasis. [Internet]. World Health Organization; 2015. Available from http://www.who.int/lymphatic_filariasis/en/.

McNeil W (2013) Expression and purification of a recombinant chitin deacetylase from Caenorhabiditis elegans. (Unpublished work at Tufts University.)

McNeil W (2014) Expression and purification of a recombinant chitin deacetylase from Caenorhabiditis elegans. (Unpublished Senior Thesis at Tufts University.)

Park DW, Kim SS, Nam MK, Kim GY, Kim J, Rhim H (2011) Improved recovery of active GST-fusion proteins from insoluble aggregates: solubilization and purification conditions using PKM2 and HtrA2 as model proteins. BMB Reports 44: 279 - 284.

Psylinakis E, Boneca IG, Mavromatis K, Deli A, Hayhurst E, *et al.* (2005). Peptidoglycan *N*-Acetylglucosamine Deacetylases from *Bacillus cereus*, Highly Conserved Proteins in *Bacillus anthracis.* Journal of Biological Chemistry 280: 30856–30863

Tao H, Liu W, Simmons BN, Harris HK, *et al.* (2010) Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS. BioTechniques 48:61-64.

Tokuyasu K, Kaneko S, Hayashi K, Mori Y (1999). Production of a recombinant chitin deacetylase in the culture medium of *Escherichia coli* cells. *FEBS Lett. 458:* 23-26

Toprak U, Baldwin D, Erlandson M, Gillott C, Hou X, *et al.* (2008). A chitin deacetylase and putative insect intestinal lipases are components of the *Mamestra configurata* (Lepidoptera: Noctuidae) peritrophic matrix. Insect Molecular Biology 17: 573-585

Trudel J, Asselin A (1990) Detection of chitin deacetylase activity after polyacrylamide gel electrophoresis. Anal Biochem 189: 249-253

Urch JE, Hurtado-Guerrero R, Brosson D, Liu Z, Eijsink VGH, *et al.* (2009) Structural and functional characterization of a putative polysaccharide deacetylase of the human parasite *Encephalitozoon cuniculi*. Protein Science 2009 18: 1197-1209

Zhong X, Wang X, Tan X, Xia Q, Xiang Z, Zhao P (2014). Identification and Molecular Characterization of a Chitin Deacetylase from *Bombyx mori* Peritrophic Membrane. Int. J. Mol. Sci 15: 1946-1961

Acknowledgements:

I sincerely thank Dr. Juliet Fuhrman for being my PI, research advisor, and teacher over the past three years, and for guiding me to develop my knowledge of biology and its practical applications beyond the scope of my coursework. Thank you Dr. Joshua Kritzer and Dr. Michael Romero for serving on my thesis committee. Thank you William McNeil and Brittany Ruhland for your collaboration on select projects presented in this paper, and for being wonderful to work with in lab; thanks also to Danielle Gillard, Sara Balch, Harris Fitzgerel, and Sabrina Khan.

Thank you Anne Moore and the Tufts Summer Scholars program for funding a portion of this work and providing multiple opportunities to present it at Tufts. Finally, thank you to the Tufts University Biology Department for sustaining our lab and making projects like these possible.