

5-2014

# Expression of Human Neutrophil Cathepsin G in *Pichia pastoris*

Evan T. Perry

Follow this and additional works at: <https://dc.etsu.edu/honors>

 Part of the [Biochemistry Commons](#)

---

## Recommended Citation

Perry, Evan T., "Expression of Human Neutrophil Cathepsin G in *Pichia pastoris*" (2014). *Undergraduate Honors Theses*. Paper 225.  
<https://dc.etsu.edu/honors/225>

This Honors Thesis - Open Access is brought to you for free and open access by the Student Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact [digilib@etsu.edu](mailto:digilib@etsu.edu).

EXPRESSION OF HUMAN NEUTROPHIL CATHEPSIN G IN *PICHIA PASTORIS*

Thesis submitted in partial fulfillment of Honors

By

Evan Thomas Perry  
The Honors College  
Midway Honors Scholar Program  
East Tennessee State University

May 6<sup>th</sup>, 2014

---

Evan Thomas Perry, Author

---

Dr. David Johnson, Faculty Mentor

---

Dr. Theo Hagg, Faculty Reader

---

Dr. Jonathan Peterson, Faculty Reader

## ABSTRACT

Cathepsin G (CatG), a serine protease found in the azurophil granules of neutrophils, participates in killing engulfed microorganisms. CatG is a poorly understood enzyme, in part because it can only be obtained as mature enzyme purified from human blood, and because it seems to have dual specificity for chymotrypsin-like and trypsin-like substrates. Therefore, yeast *Pichia pastoris* was used to express immature recombinant human CatG to provide a source of the enzyme free of biohazards and to allow study of its dual specificity and C-terminal processing. To avoid potential cleavage by a yeast kexin protease, the amino sequence was modified without altering CatG's biological activity to remove one glycosylation site and eight dibasic sites. An N-terminal 6-His-cytochrome B5 (CytB5) heme binding fusion domain was linked to the modified human CatG by an enteropeptidase cleavage site for activation. The DNA for this construct was codon-optimized and placed in the pPICz $\alpha$  secretion vector. After transforming *P. pastoris* strain X-33, 48 Zeocin-resistant clones were screened for relative levels of CatG activity following activation by recombinant human enteropeptidase. Recombinant human CatG was partially purified from fermentation medium by nickel affinity chromatography and its activity was confirmed by assays using the substrate Succinyl-Ala-Ala-Pro-Phe-SBzl. Supported by a Student Faculty Collaborative Grant from the ETSU Honors College and ETSU Office of Research and Sponsored Programs and by NHLB grant R15HL091770.

## INTRODUCTION

Neutrophils, also known as polymorphonuclear leukocytes, comprise the most abundant type of white blood cell in the human body, providing a rapid and localized response to infection (Borregaard & Cowland, 1997; Korkmaz et al., 2010). While typically found in circulation in the blood stream, neutrophils relocate to sites of inflammation in tissues by crossing activated vascular endothelial cells, a process called diapedesis (Borregaard et al., 2005). Highly phagocytic neutrophils kill engulfed microorganisms by the fusion of granules containing many types of antimicrobial peptides, proteases, and reactive oxygen species. While at least four types of granules have been described, the azurophil granules, also called primary granules, are characterized by the ability to stain with the basic dye azure and by the presence of myeloperoxidase (Borregaard & Cowland, 1997). The azurophil granules also contain the three neutrophil serine proteases (NSPs) cathepsin G (CatG), human neutrophil elastase (HNE), and myeloblastin (elastase-like) along with defensins, bacterial permeability-increasing protein, and azurocidin (an inactive member of the serine protease gene family) (Borregaard et al., 2005; Pham, 2006). Following phagocytosis of the pathogen, the azurophil granules fuse with the phagolysosome and release their cytotoxic contents to destroy the invader (Pham, 2006).

The proteolytic activities of NSPs provide an important non-oxidative mechanism for the neutrophil to kill pathogens that resist oxidative attacks (Bangalore et al., 1990). It has also been found that the antimicrobial properties of NSPs are not only dependent on proteolytic activity but also due to their strong positive charges. Cathepsin G

releases charged antimicrobial peptide fragments after being degraded by trypsin-like enzymes (Bangalore et al., 1990).

NSPs also participate in regulation of many inflammatory signaling pathways (Pham, 2006). The NSPs process cytokines such as interleukin 6 (IL-6) (Bank et al., 1999) and chemokines like IL-8 (Padrines et al., 1994). Proteolytic modifications to these cytokines and chemokines initiate signaling pathways that can result in both pro- and anti-inflammatory responses, as well as the recruitment of various immune cells to initiate both the innate and the adaptive immune responses (Padrines et al., 1994; Wittamer et al., 2005)

Neutrophils are protected from the proteolytic activities of their own NSPs by storage in the acidic granules as active enzymes; consequently, the release of NSPs outside of the neutrophil can pose a serious threat to surrounding tissue. In fact, many human diseases have been classified due to improper regulation of NSPs. NSPs are capable of degrading collagen and elastin, making them dangerous when exposed to connective tissue, especially lung tissue. Some commonly known diseases resulting in NSP activity outside the neutrophil are chronic obstructive pulmonary disorder (COPD), acute respiratory distress syndrome, and ischemia-reperfusion injury (Korkmaz, Moreau & Gauthier, 2008).

Given the consequences of poorly controlled NSPs there is great interest in understanding how NSPs are regulated and directed to the azurophil granules without damaging the neutrophil or surrounding tissue. It has been shown that all NSPs are synthesized as zymogens in inactive forms as pre-proteins containing an N-terminal signal peptide that directs the protein to the endoplasmic reticulum and a pro-dipeptide

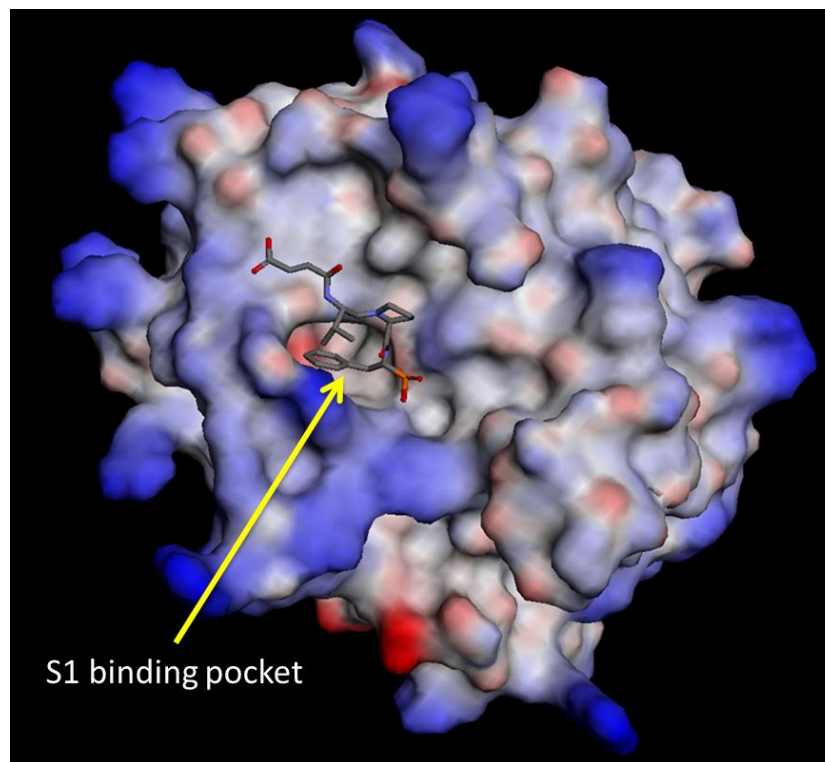
that keeps the enzyme in an inactive state until granule maturation (Korkmaz et al., 2010). It has been shown that the signal peptide is removed by signal peptidase and the N-terminal pro-dipeptide by dipeptidyl peptidase I (also called Cathepsin C) (Korkmaz et al., 2010).

In addition to the removal of N-terminal pro-peptides leading to activation, the mature NSPs in the azurophil granule are missing C-terminal peptide extensions. By studying NSP cDNA it can be seen that the amino acid sequence and length of the C-terminal extension varies between CatG, elastase, and myeloblastin. The C-terminal peptides do not play a role in activation but are removed during enzyme maturation by an unidentified protease (or proteases) (Gullberg et al., 1995). Thus there is some interest in studying what removes this C-terminal extension and what function it might serve.

Currently the only available source of the NSPs to study their characteristics is by purification from human blood or tissue sources. These purified NSPs are in the mature state and lack the C-terminal pro-peptide, making it nearly impossible to study its role in the NSP biosynthesis process. Due to the recent advances in genetics and the ability to express homologous proteins in various expression systems, however, there exists a potential to genetically engineer an expression system to make an NSP with an intact C-terminal extension peptide to aid the study of its biochemical functions.

While other recombinant NSPs have been expressed in many cell types, the expression of full length human CatG has only been reported in RBL cells and murine myeloid cell line 32D (Gullberg et al., 1995). CatG has not been reported to be expressed in yeast species. CatG is noted for its high arginine content and dual

specificity. Using the S1/P1 Schechter-Berger nomenclature, CatG shows both chymotrypsin-like P1 specificity and trypsin like-P1 specificity (Fig. 1) (Hof et al., 1996). The large, hydrophobic S1 pocket confers chymotrypsin-like specificity for the bulky hydrophobic amino acids phenylalanine (as shown) and tyrosine. Trypsin-like specificity for positively charged lysine and arginine residues is a result of the negatively charged residue glutamate 226 (chymotrypsinogen numbering) at the base of the S1 pocket.



**Figure 1:** Crystal structure of human neutrophil CatG with Suc-VPF-(OPh)<sub>2</sub> inhibitor (PDB ID: 1CGH). Blue and red regions represent positive and negative charges, respectively, while gray regions represent hydrophobic areas. The S1 pocket is labeled to show CatG's catalytic site and dual specificity (Hof et al., 1996).

CatG has a native pI of about 12.0, highest among the NSPs, and contains 34 arginine residues out of its 224 amino acid length (Hof et al., 1996). The C-terminal extension peptide of CatG ranges from 11 to 13 amino acids in length, as the precise

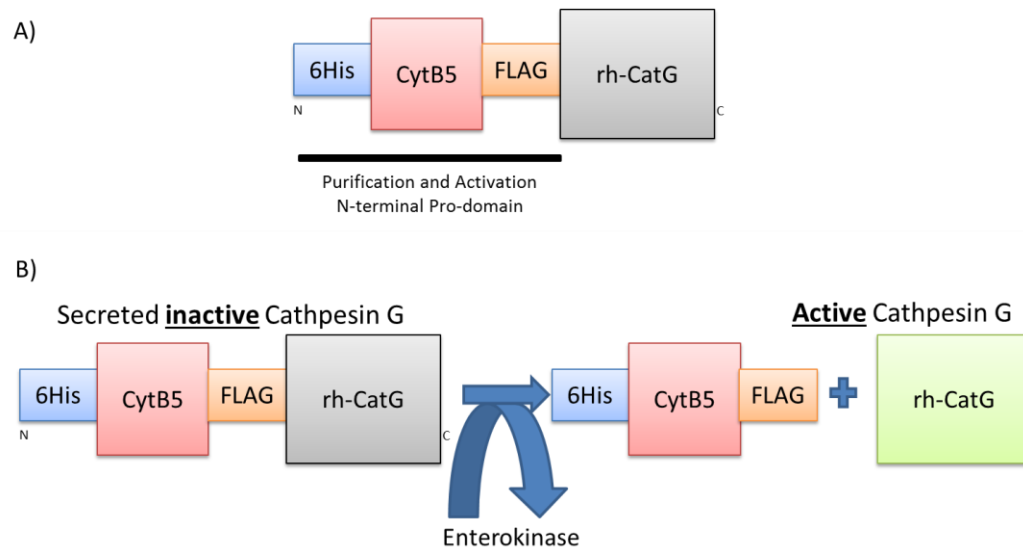
cleavage site is not clear. Previous research by crystallographic models of mature CatG shows the most common cleavage site to remove the C-terminal extension peptide is at serine 244, but one structure reported a C-terminal arginine 243, which suggests the possibility of autolysis by CatG (de Garavilla et al., 2005; Greco et al., 2002; Hof et al., 1996).

Recombinant expression of CatG has proven to have many challenges. CatG shows both antibacterial (Pham, 2006) and antifungal activity (Newman et al., 2000) making bacteria species and yeast species questionable expression hosts. While expression in insect cells and/or mammalian cells is theoretically possible, these expression hosts are not capable of reaching a high cell density and thus could have lower expression amounts. Popular bacterial expression hosts, such as *E. coli*, would not be ideal due to CatG's strong antimicrobial properties by both proteolytic activity and by positivity charged amino acid sequences (Bangalore et al., 1990). The yeast species *Pichia pastoris* was chosen as the expression system due to the eukaryotic protein expression, folding, co- and post-translation modification capabilities, and secretion mechanisms, relative to prokaryotic expression hosts (Smith, 2013). However to avoid CatG's antifungal activity, recombinant CatG has been designed to include a fusion domain at its N-terminus to keep CatG in an inactive zymogen form in the growth media.

Several fusion proteins were used in previous attempts to express CatG, including human serum albumin (HSA), enhanced green fluorescent protein (eGFP), rubredoxin, and small ubiquitin-like modifying protein (SUMO); however, all expression trials with these fusion proteins yielded low expression amounts. Due to previous



success expressing HNE with the fusion protein cytochrome B5 (CytB5) heme binding domain from the rat species *Rattus norvegicus*, CytB5 was used as a fusion domain in this expression trial. The CytB5 domain was chosen because, when bound to heme, it has distinct absorbance at 410 nm and produces visible red color if present in high enough concentration (Mitra et al., 2005). Thus the CytB5 fusion protein can be traced spectrophotometrically at 410 nm and visually during the purification and processing steps. The fusion design (Fig. 2) also contains a six histidine (6His) sequence at the N-terminal of the CytB5 domain to aid in purification. To link the C-terminus of the CytB5 domain and the N-terminus of the CatG domain, a FLAG tag linker with the sequence DYKDDDDK is included that is capable of being cleaved by enteropeptidase (EP) after the lysine to expose CatG's native N-terminus (Fig. 2).



**Figure 2:** A) Block diagram showing the fusion protein design of Cathepsin G and cytochrome B5 with a six histidine tag for purification and FLAG tag for specific cleavage by enteropeptidase to separate the domains after purification. B) Activation schematic showing enteropeptidase cleaves after lysine in the FLAG tag to release the two domains and activate human recombinant Cathepsin G.

The amino acid sequence of Cathepsin G was modified in an attempt to enhance expression in *P. pastoris*. Arginines (R) were replaced with glutamines (Q) at 8 sites to reduce the net positive charge on CatG and to block potential cleavage by the endogenous yeast Kex2 protease, used in the secretory mechanism (Fig 3). While glutamine is uncharged it maintains arginine's general shape, so the pI of the protein should be decreased while its overall structure and function theoretically should not be dramatically affected. The other amino acid substitution made was a glutamine (Q) for an asparagine (N) to block a potential N-linked glycosylation site, as yeast species are known to hyperglycosylate proteins (Niles et al., 1998).

rhCatG	I IGGRESRPHSRPYMAYLQIQSPAGQSRGGFLVREDFVLTAAHCWGSNI <b>Q</b> VTLGAHNIQ
nativeCatG	I IGGRESRPHSRPYMAYLQIQSPAGQSRGGFLVREDFVLTAAHCWGSNI <b>N</b> VTLGAHNIQ
rhCatG	R <b>Q</b> ENTQQHITA <b>Q</b> RAIRHPQYNQRTIQNDIMLLQLSR <b>Q</b> VR <b>Q</b> NRRNVNPVALPRAQEGLRPGT
nativeCatG	R <b>R</b> ENTQQHITA <b>R</b> RAIRHPQYNQRTIQNDIMLLQLSR <b>R</b> VR <b>R</b> NRRNVNPVALPRAQEGLRPGT
rhCatG	LCTVAGWGRVSMR <b>Q</b> GTDTLREVQLRVQRDRQCLRIFGSYDP <b>Q</b> RQICVGDR <b>Q</b> E <b>Q</b> KAAFKGD
nativeCatG	LCTVAGWGRVSMR <b>R</b> GTDTLREVQLRVQRDRQCLRIFGSYDP <b>R</b> RQICVGDR <b>R</b> E <b>R</b> KAAFKGD
rhCatG	SGGPLLCNNVAHGIVSYGKSSGVPPEVFTRVSSFLPWIRTTMRS <b>FKLLDQMETPL</b>
nativeCatG	SGGPLLCNNVAHGIVSYGKSSGVPPEVFTRVSSFLPWIRTTMRS <b>FKLLDQMETPL</b>

**Figure 3:** Sequence alignment of native CatG (bottom rows) and recombinant human CatG (top rows). Eight amino acid substitutions (blue) were made from arginine (R) to glutamine (Q) to reduce overall pI of CatG and avoid potential cleavage by yeast internal protease Kex2 in attempts to enhance expression. The other mutation of asparagine (N) to glutamine (Q) was made to block potential a glycosylation site. The C-terminal pro-peptide extension is highlighted in red.

The resulting amino acid sequence was then combined with the cytochrome B5 domain, FLAG tag and 6His tag to give the overall amino acid sequence (Fig. 4). The cDNA for this amino acid sequence was codon optimized for *P. pastoris* to optimize expression by using the codons that appear most frequently in the *P. pastoris* genome.

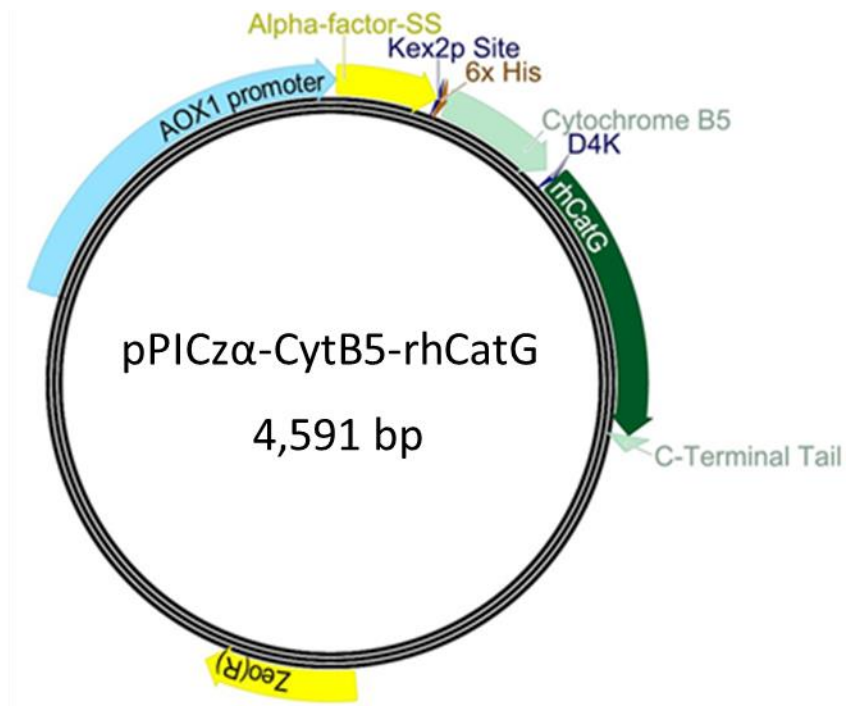
>CytB5-rhCatG

HHHHHAEQSDKDVKYITLEEIQKHKDSKSTWVILHHKVYDLTKFLEEHPPGEEVLRQAGGDATENFEDVGHSTD  
ARELSKTYIIGELHPDDRSKIAKPSETLGTDYKDDDDKIIGGRESRPHSRPYMAYLQIQSPAGQSRCGGFLVREDFVLT  
AAHCWGSNIQVTLGAHNIQRQENTQQHITAQRAIRHPQYNQRTIQNDIMLLQLSRQVRQNRNVNPVALPRAQEG  
LRPGTLCTVAGWGRVSMRQGTDTLREVQLRVQRDRQCLRIFGSYDPQRQICVGDREKQAAFKGDSGGPLLCENN  
VAHGIVSYGKSSGVPPEVFTRVSSFLPWIRTTMRSFKLLDQMETPL

C-terminal propeptide

**Figure 4:** Amino acid sequence of full length 6His-CytB5-FLAG-CatG. The 6His sequence is shown in red and following it is the cytochrome B5 domain. The FLAG tag linker is shown in blue. The recombinant human CatG gene is shown after the FLAG tag with its C-terminal pro-peptide extension underlined in red.

The resulting DNA sequence of codon optimized 6His-CytB5-FLAG-CatG was ligated into the pPIC $\alpha$  plasmid from Invitrogen (Fig. 5). This plasmid was chosen because of its alpha factor domain that signals for secretion of the fusion protein into growth media for easy recovery and the production is regulated by the methanol-inducible AOX1 promoter. The Zeo(R) gene confers Zeocin resistance, making selection for transformed colonies possible on Zeocin treated plates. This plasmid was linearized and introduced to the yeast genome by electroporation and sequential homologous recombination.



**Figure 5:** The pPICzα plasmid with the 6His-CytB5-FLAG-CatG ligated. Key features of the plasmid include the expression of the alpha factor that signals for secretion, the methanol-inducible AOX1 promoter and the Zeo(R) gene that gives the *P. pastoris* resistance to Zeocin for selection of transformed cells.

## MATERIALS AND METHODS

### MOLECULAR CLONING / TRANSFORMATION / SCREENING

The molecular cloning, transformation, and expression screening for CytB5 CatG were performed by Dr. Eliot Smith and all credit is thereby given to his efforts (Smith, 2013). The native amino acid sequence for *Homo sapiens* CatG (NP\_001902.1) and the soluble, heme-binding domain of *Rattus norvegicus* cytochrome B5 (NP\_071581.1) were obtained from the National Center for Biotechnology Information (NCBI). The sequences were aligned and modified in the computer software Geneious® 6.1.4. In

the CatG sequence, amino acid mutations were made by replacing asparagine with glutamine at position 65 and replacing arginine with glutamine at positions 76, 86, 111, 114, 148, 178, 186 and 188. Codons for the FLAG tag sequence (DYKDDDDK) were added to the 5' end of the Cathepsin G gene. The resulting recombinant human CatG (rhCatG) sequence was codon optimized (GenScript) for expression in *P. pastoris* and provided in the sub-cloning plasmid pUC57 with 5' KpnI and 3' NotI restriction sites flanking the FLAG-CatG sequence. A 6Hisamino acid sequence was added to the N-terminus of the CytB5 domain. This sequence was codon optimized (GenScript) and provided in a separate pUC57 sub-cloning plasmid with the gene flanked by restriction sites 5' XhoI and 3' KpnI. The pUC57 sub-cloning plasmids were introduced to Zymo DH5 $\alpha$  "Easy Comp" *E. coli* for stock and were purified for further cloning by miniprep purification technique (Zymo Research).

The plasmids were treated with their corresponding restriction enzymes (Thermo Scientific, Fermentas FastDigest®) and the desired bands were recovered by gel extraction using the Zymoclean™ gel DNA recovery kit (Zymo Research). The ligation process happened in two-steps where the pPICz $\alpha$  and pUC57-6His-CytB5 plasmids were cut with XhoI and KpnI and mixed and ligated with Quick Stick Ligase (Bioline) resulting in the insertion of the 6His-CytB5 gene into the pPICz $\alpha$  plasmid. This resulted ligated pPICz $\alpha$ -6His-CytB5 plasmid was cloned into DH5 $\alpha$  *E. coli*. For the final ligation the pPICz $\alpha$ -6His-CytB5 and pUC57-FLAG-rhCatG plasmids were cut with restriction enzymes KpnI and NotI and mixed resulting in the ligation to form the desired the pPICz $\alpha$ -6His-CytB5-FLAG-rhCatG plasmid which was again cloned into DH5 $\alpha$  *E. coli*.

The presence of the insert was confirmed by amplification by PCR and the size was confirmed using gel electrophoresis.

Transformation was done per the EasySelect™ Pichia Expression Kit and protocol (Life Technologies). The plasmid was linearized for electroporation by the restriction enzyme BstXI. The linearized plasmid was further purified by a DNA cleanup Kit (Zymo Research). The purified plasmid was then concentrated by precipitating the DNA with ethanol, then evaporating the ethanol and resuspending the DNA in a small amount of nuclease free water. Purified pPICzα-6His-CytB5-FLAG-rhCatG plasmid was electroporated into *P. pastoris* strain X-33 cells by the Life Technologies electroporation protocol using a 2 mm gap cuvette and BTX ECM 600 electroporation system, with the settings of 1.5 kV and 186 Ω producing a 7.5 millisecond pulse length. Successful transformants were selected by plating on YPDS (10 g/L yeast extract, 20 g/L peptone, 182.2 g/L sorbitol, 2% glucose) agar plates treated with 100 µg/mL Zeocin™ (Life Technologies) and 100 µg/mL ampicillin. Colonies that showed resistance to Zeocin™ were transferred to numbered location spots on fresh YPD agar plates with 100 µg/mL Zeocin™ and 100 µg/mL ampicillin.

To screen for enzyme production, 44 colonies were picked and grown in 2 mL cultures in 12-well plates. The cells were grown in synthetic minimal medium (SMM) (10 g/L monosodium glutamate (MSG), 2.5 mg/L ammonium sulfate (NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>), 1.34% yeast nitrogen base, 4 x 10<sup>-5</sup>% biotin, 100 mM potassium phosphate (KH<sub>2</sub>PO<sub>4</sub>), pH = 5.0) with 100 µg/mL of both Zeocin™ and ampicillin. The 12-well plates were placed on an orbital shaking platform and grown at 25 °C for 4 days. The cultures received 1% feedings of methanol at inoculation and every 24 hours after. After 4 days,

the cells were harvested by centrifugation. The supernatants were removed from the cell pellets and screened for activity. The cell pellets were weighed to give a wet cell weight (WCW) for quantification by comparison of the culture activity per gram WCW.

## FERMENTATION

The fermentation was run in a BioFlo 110 bioreactor with 7.5 liter capacity (New Brunswick). The fermentation vessel was equipped with a pH probe and dissolved oxygen ( $dO_2$ ) sensor (Mettler-Toledo) to monitor culture conditions. The most successful growth medium employed the basal salts medium (BSM) (26.7 mL/L phosphoric acid ( $H_3PO_4$ ), 0.93 g/L calcium phosphate ( $CaSO_4$ ), 18.2 g/L dipotassium phosphate ( $K_2SO_4$ ), 14.9 g/L magnesium sulfate heptahydrate ( $MgSO_4 \cdot 7H_2O$ ), 40 g/L glycerol) suggested by Life Technologies. BSM was supplemented with 10 g/L MSG and 1 mL/L PTM4 trace element supplements ( 2 g/L  $CuSO_4 \cdot 5H_2O$ , 0.08 g/L NaI, 3.0 g/L  $MnSO_4 \cdot 2H_2O$ , 0.2 g/L  $Na_2MoO_4 \cdot 2H_2O$ , 0.02 g/L  $H_3BO_3$ , 0.5 g/L  $CoCl_2$ , 7 g/L  $ZnCl_2$ , 22 g/L  $Fe(II)SO_4 \cdot 7H_2O$ , 0.2 g/L Biotin, 1 mL/L  $H_2SO_4$ , 0.5 g/L  $CaSO_4 \cdot 2H_2O$ ) (Stratton, Chiruvolu & Meagher, 1998). A supplementary vitamin mix was also added from a 10X concentrate containing 0.1 g/L L-His-HCl, 0.2 g/L DL-Met, 0.2 g/L DL-Trp, 0.0002 g/L biotin, 0.0002 g/L folic acid, 0.2 g/L inositol, 0.04 g/L niacin, 0.02 g/L p-aminobenzoic acid, 0.04 g/L pyridoxine HCl, 0.02 g/L riboflavin, 0.04 g/L thiamine HCl, 0.04 g/L calcium pantothenate, and 10 g/L  $CaCl_2$ .

BSM was assembled in the fermentation vessel with calculations made for a final volume of 3.5 liters at the time of inoculation. While the amounts of ingredients were added for 3.5 liters, only 2.5 liters of volume was added due to the fact other ingredients

would be added after the medium was autoclaved. After the autoclaving process and once the medium was cooled, 15 mL of filter-sterilized PTM4 salts and 35 mL of filter-sterilized 100X vitamin mix. The fermenter vessel was hooked up to an oxygen tank and the oxygen probe was calibrated by achieving 100% dO<sub>2</sub> with a blast of oxygen at 4 L/min and 1200 RPM. The pH probe was calibrated using pH = 4 and pH = 7 buffering solutions. X-33 *P. pastoris* colony #43 containing the incorporated pPICzα-6xHis-CytB5-FLAG-CatG vector was grown two days before inoculation in a 10mL “seed” culture containing 10mL YPD with 100 µg/mL Zeocin and ampicillin. Then the seed culture was used to inoculate a fresh 500mL culture containing SMM media with glycerol. This 500 mL starter culture was allowed to grow overnight and was then used to inoculate the fermenter.

To prepare for inoculation the air flow was turned down to 0.4 L/min and the fermenter was hooked up with a feed of 12.5 % ammonia controlled by variations in the pH readings. The initial medium was intentionally placed at very low pH (pH < 2) with phosphoric acid so that the pH was raised by the addition of ammonia (approximately 250 mL) before inoculation to give the cells both a phosphate source (phosphoric acid) and nitrogen (ammonia) source to start the fermentation process. As the cells metabolize the carbon source the pH of the media decreases; consequently, the auto adjustment of pH by ammonia allows for constant addition of nitrogen source for the cells. Approximately 200 mL of sterile ultrapure water was added to the medium to bring the volume to 3 liters. The fermentation medium was inoculated with 500 mL of the starter culture bringing the final volume to 3.5 L.



The fermentation was run in three distinct phases. The glycerol fed-batch phase allowed the cells to exhaust the 40 g/L glycerol already present in the fermentation media. During the glycerol fed-batch phase, the cells were fed a 40% (w/v) glycerol solution drop wise at 1 mL/min to increase cell density and to ensure that the culture was carbon-limited. Daily 25 mL samples were collected for centrifugation and analysis. The supernatant was removed and cell pellet was weighed to monitor cell density. When the fermenter had reached sufficient density the culture was transitioned to methanol to start the methanol fed-batch phase to induce expression of the 6His-FLAG-rhCatG. Drop wise addition of 95% Methanol with 5 mL/L PTM4 was introduced following Life Technologies protocol. The addition of methanol had to be carefully monitored as the accumulation of too much methanol can be toxic to the cells. The cell density was monitored as well as spikes in the  $dO_2$  content when adjusting the methanol feed rate. The cells were grown until they entered stationary phase when the weight cell weight leveled off. The cells were harvested at 140 hours post-induction by centrifugation. The supernatant was pooled and kept for further processing while the cell pellets were discarded.

#### TREATMENT OF MEDIUM

Two liters of fermentation medium was filtered first by glass fiber pre-filters (Millipore Cat#AP2004200) then by 0.45  $\mu$ m regenerated cellulose filters (Pall P/N 60173). Clarified medium was then concentrated by ultrafiltration using two 10 kDa molecular weight cut off tangential flow filtration cartridges (Sartorius Vivaflow50 P/N VF05P0) in series. Once concentrated to approximately 50 mL, the medium was diluted

1:20 with 500 mM acetic acid (pH unadjusted) to eliminate phosphate and calcium that would otherwise precipitate at pH levels above 5. After two 1:20 dilutions with acetic acid, the solution was further diluted three times using a 1:20 dilution each time with IMAC equilibration buffer (25 mM sodium phosphate, 300 mM NaCl, pH = 8). The sample was then concentrated to approximately 50 mL to be loaded on an IMAC column for purification.

#### PURIFICATION (IMMOBILIZED METAL AFFINITY CHROMATOGRAPHY)

The column was set up using Profinity™ Ni<sup>2+</sup> IMAC Resin (Bio-Rad) resin suspended in IMAC equilibration buffer. The column was loaded with the concentrated fermentation sample at 1 mL/min. The flow through from the loaded sample was collected in a clean container. After all the sample had been loaded, the column was washed with IMAC Wash Buffer (200 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, pH = 8) at a rate of 5 mL/min. The wash step flow through volume was collected in 5 mL fractions to watch for the A280 and A410 peaks. For the elution step, the IMAC Elution Buffer (200mM sodium phosphate, 300 mM NaCl, 20 mM imidazole at pH = 8) was loaded at a rate of 2 mL/min into 2 mL fractions. To monitor protein elution an absorbance spectrum on 100 µL of each fraction sample was measured from 250 nm to 600 nm using a Biotek PowerWave™ XS2 plate reader. The resulting absorbance values at 280nm and 410nm were recorded.

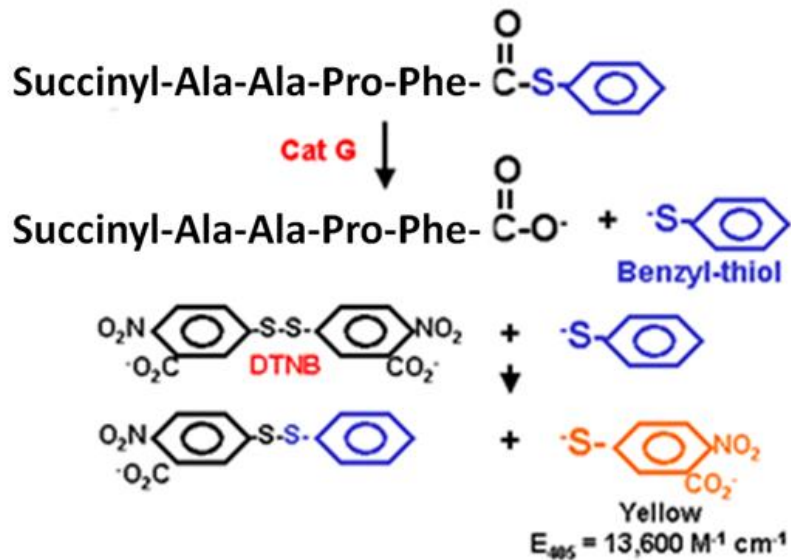
## SAMPLE ACTIVATION BY ENTEROPEPTIDASE

For all liquid enteropeptidase (EP) activations 100  $\mu\text{L}$  of the sample of interest was combined with 2  $\mu\text{L}$  of EP. The sample was then incubated in a 37  $^{\circ}\text{C}$  water bath for 24 hours.

## ACTIVITY ASSAYS

All assays testing for the activity of rhCatG were performed using the synthetic substrate Suc-AAPF-SBzl (LifeTein custom synthesis). This synthetic substrate is cleaved by CatG after the phenylalanine and the benzyl-thiol is released which then outcompetes for one of the substituted benzene rings on DTNB (Fig 6). The displaced substituted benzene generated as the final product in the reaction produces a yellow color and has a defined extension coefficient at 410 nm ( $\epsilon_{410 \text{ nm}} = 13600 \text{ M}^{-1} \text{ cm}^{-1}$ ) and by monitoring its absorbance over time, the amount of activity can be quantified (Johnson, 2006).

### CatG Thioster Substrate



**Figure 6:** Synthetic substrate succinyl-AAPF-thiobenzyl ester used to measure the activity of CatG. CatG is specific to cleave after the phenylalanine (F) and results in the release of the benzyl-thiol group which out completes one of the substitutive containing benzene rings releases a carbo,nitro benzene ring that produced yellow color and absorbance can be tracked spectrophotometrically at 410 nm over time (Johnson, 2006).

The assay conditions were in a 100  $\mu\text{L}$  total volume reaction in 96-well plates in the Biotek PowerWave™ XS2 plate reader which monitored the increasing absorbance at 410 nm every 30 seconds for a 10 minute total assay time. Each sample contained 25  $\mu\text{L}$  of CatG Assay Buffer (100 mM HEPES, 1 M NaCl, 10 % glycerol, 0.01 % Triton X-100, and 0.01 %  $\text{NaN}_3$  at pH = 8.0). To the buffer, 25  $\mu\text{L}$  of sample containing rhCatG was added. Immediately before the assay was started, 50  $\mu\text{L}$  of substrate buffer (300  $\mu\text{M}$  Suc-AAPF-SBzl, 100 mM DTNB in CatG assay buffer) was added to bring the reaction volume to 100  $\mu\text{L}$ .

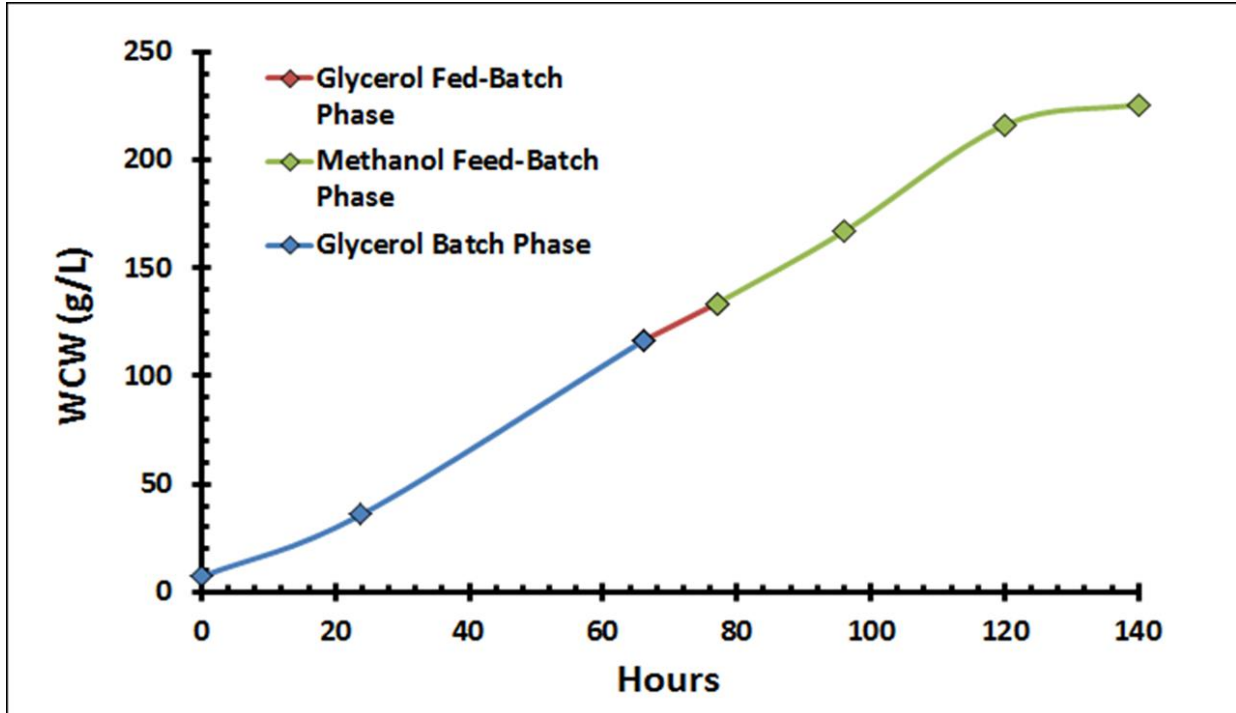
## DOT BLOT

The dot blot was prepared by pipetting very small volumes on a dry piece of nitrocellulose membrane. Human purified CatG was loaded in amounts of 2  $\mu\text{g}$ , 0.5  $\mu\text{g}$ , 0.1  $\mu\text{g}$ , and 0.05  $\mu\text{g}$  to serve as CatG positive controls for the primary antibody. Also blotted was a 10  $\mu\text{L}$  sample from the IMAC purified pooled 6His-CytB5-FLAG-CatG sample. To test the primary rabbit-anti-CatG polyclonal IgG antibody and secondary HRP-conjugated goat-anti-rabbit IgG polyclonal antibody efficacy 2  $\mu\text{L}$  of a 1:4000 dilution was blotted to the membrane. Two samples of secondary antibody were tested on the blot and the second (#2) was used to develop the blot. After loading the samples the nitrocellulose membrane was blocked overnight in a 1% BSA solution in TBST (10 mM Tris, 150 mM NaCl, 0.05% Tween20 pH= 7.5) at 4 °C. After 24 hours, the blot was rinsed with 25 mL of TBST 4 times and the blot was allowed to sit in the primary antibody (CatG-anti-rabbit) at a 1:4,000 dilution in TBST overnight at 4 °C. After another 24 hours the blot was rinsed with 4 more rinses of 25 mL TBST. The secondary antibody #2 (HRP-anti-rabbit) was added at a 1:4,000 dilution in TBST and was allowed to react for 1 hour at 4 °C. The membrane was rinsed with the 4 washes of TBST. The image was exposed to HRP ECL SuperSignal® West Pico Chemiluminescent Substrate (Thermo) immediately prior to imaging.

## RESULTS

### FERMENTATION CELL DENSITY

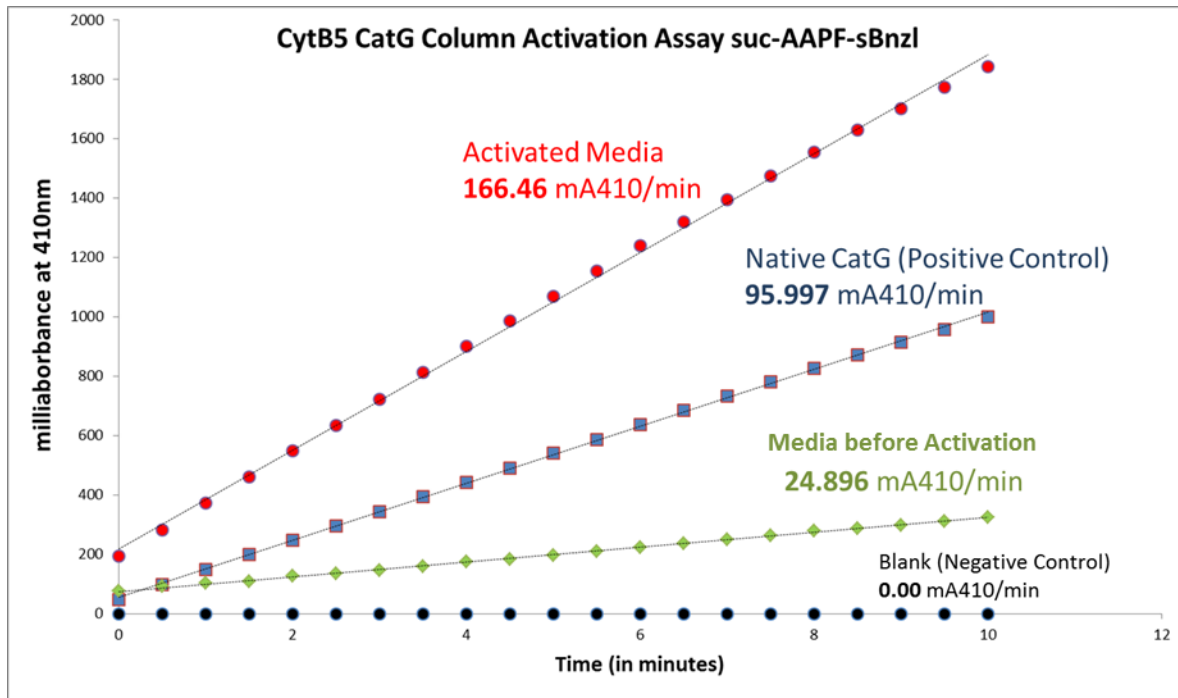
The large scale growth of the best expressing clone of rhCatG #43 yielded a final cell density of approximately 225 g/L with a final volume of approximately 6 liters (Fig. 7). The culture was on the glycerol fed-batch phase until a cell density of 140 g/L was reached. The culture was then transitioned to the methanol fed-batch phase after 77 hours post inoculation. The cells reached a final density of 225.76 g/L before the culture growth leveled off in stationary phase at 140 hours post inoculation.



**Figure 7:** Fermentation cell density graph in units of wet cell weight (WCW) in g/L versus the time in hours. The culture grew to sufficient density (140 g/L) on glycerol during the glycerol batch phase and glycerol fed batch phase. The culture was transitioned to the methanol fed batch phase (i.e. induced) after 77 hours post inoculation. The culture expanded for 120 hours before entering stationary phase. The culture was harvested 140 hours post-inoculation at a final cell density of 225.76 g/L.

## CATG ACTIVATION

Supernatants from the fermentation medium were tested for CatG activity using the synthetic peptide substrate suc-AAPF-SBzl. The results consisted of a positive control of native CatG purified from human blood neutrophils and a negative control without CatG to check for substrate autolysis. An activity assay (Fig. 8) confirmed the presence of activatable CatG in the fermentation medium and the increase in activity between the sample before activation and the activated sample was 6 fold. The media before activation still had some activity towards the synthetic substrate, suggesting that either some rhCatG is auto-activating in the fermentation media or that *Pichia pastoris* is secreting a yeast chymotrypsin-like enzyme. A separate control sample containing only EP produced minimal activity (< 20 mA410) (data not shown) and its contribution, if any, to the overall activity was negligible in these assays.

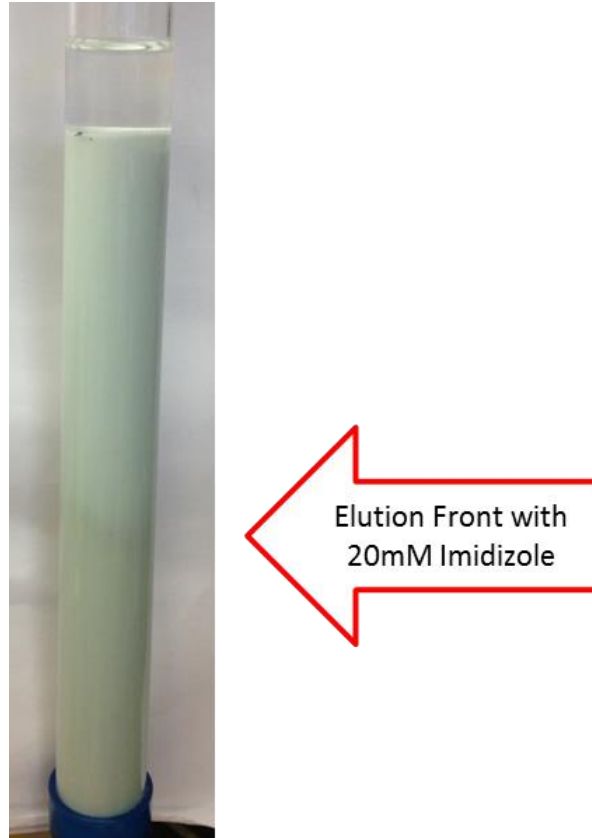


**Figure 8:** Fermentation media supernatant screened for activity using suc-AAPF-SBzl substrate. Positive control native CatG (purple) displays activity against the synthetic substrate while the negative control with no enzyme added serves as a blank. The activated sample (red) has approximately 6-fold greater activity than the equivalent sample prior to activation (green), confirming expression of the activatable 6His-CytB5-FLAG-rhCatG fusion protein design.

## IMAC COLUMN

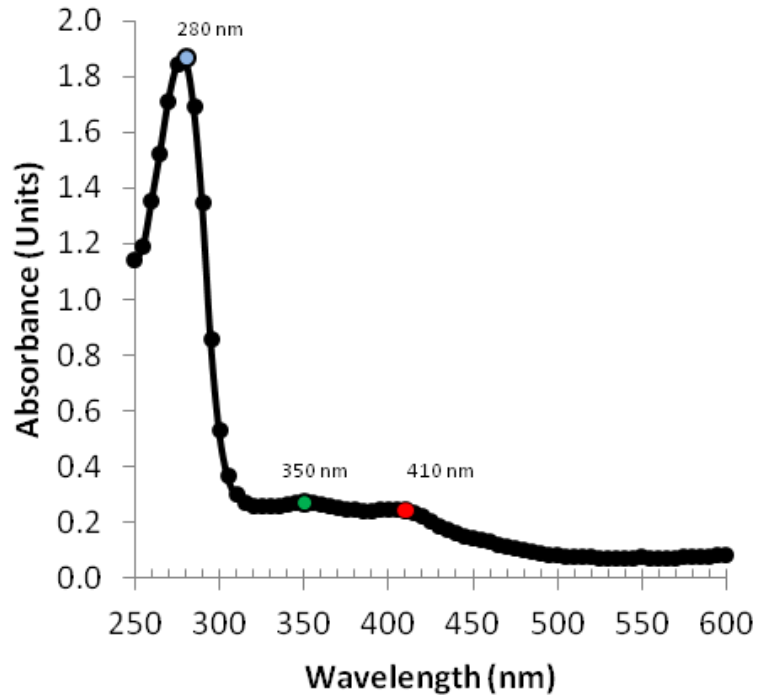
The expressed 6His-CytB5-FLAG-rhCatG fusion protein in the fermentation medium supernatant was purified by its 6His tag using a nickel IMAC column. The concentrated fermentation medium supernatant for the rhCatG clone was consistently a dark green color that is abnormal compared to other fermentation media from parallel serine protease expression projects. While most of the green color was seen in the flow through, some of the color adsorbed to the column and eluted with 20 mM imidazole (Figure 10).





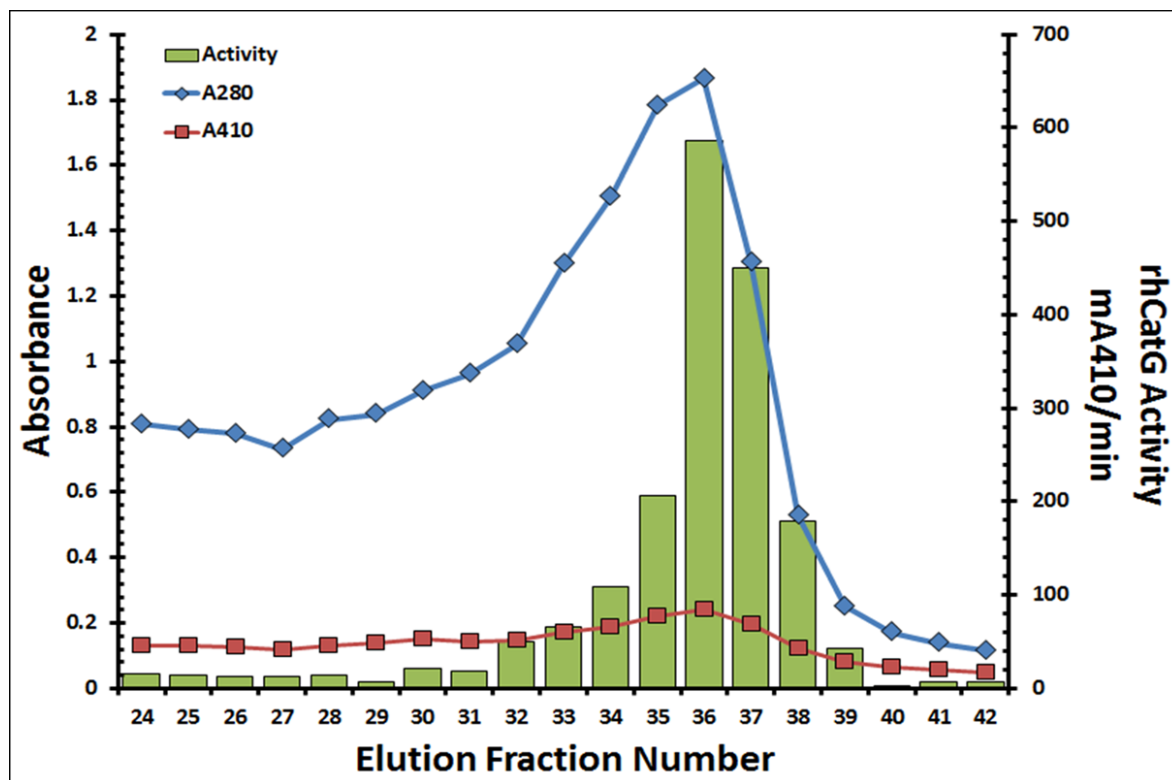
**Figure 10:** The elution front during step of the IMAC purification as tracked by the green color coming off the column during the elution step using 20mM imidazole buffer.

The green color was seen in the wash buffer and in the elution fractions; however, after fraction 33 the fractions shifted from a green color to a brownish color. This shift in color was also detected spectrophotometrically by an absorbance spectrum 250 nm to 600 nm of each fraction and fraction #36, with the highest 410 nm absorbance, had three distinct peaks at 280 nm, 350 nm, and 410 nm (Fig. 11). The 280 nm and 410 nm absorbance peaks correspond to the general protein absorbance at 280 nm and the CytB5 absorbance at 410 nm: however, the peak absorbance at 350 nm was not expected and could be due to the green color seen co-eluting with the fusion protein.



**Figure 11:** Spectrophotometric absorbance spectrum from 250 nm to 600 nm in 5 nm increments of fraction #36 from the IMAC purification column. Three distinct peaks were present at 280 nm, 350 nm, and 410 nm. The absorbance peak at 350 nm is from an unidentified contaminate protein that was seen to have a green color and was found to co-elute with the 6His-CytB5-FLAG-CatG fusion protein (Smith, 2013).

Absorbance values at 280 nm and 410nm were plotted along with the activity of each fraction to produce an elution profile (Fig 12). It can be seen from the elution profile that the activity purified partially in fractions 32 through 39. This spike in activity accompanies the peak in absorbance at 280 nm and a slight peak at 410 nm. The 280 nm absorbance values from fractions 24 to 32 correspond to some unidentified protein that is co-eluting with 6His-CytB5-FLAG-CatG and from the appearance of the fractions it appears to be the protein also responsible for the green color and absorbance at 350 nm in the wavescans. Fractions 32 through 39 were pooled for further processing.

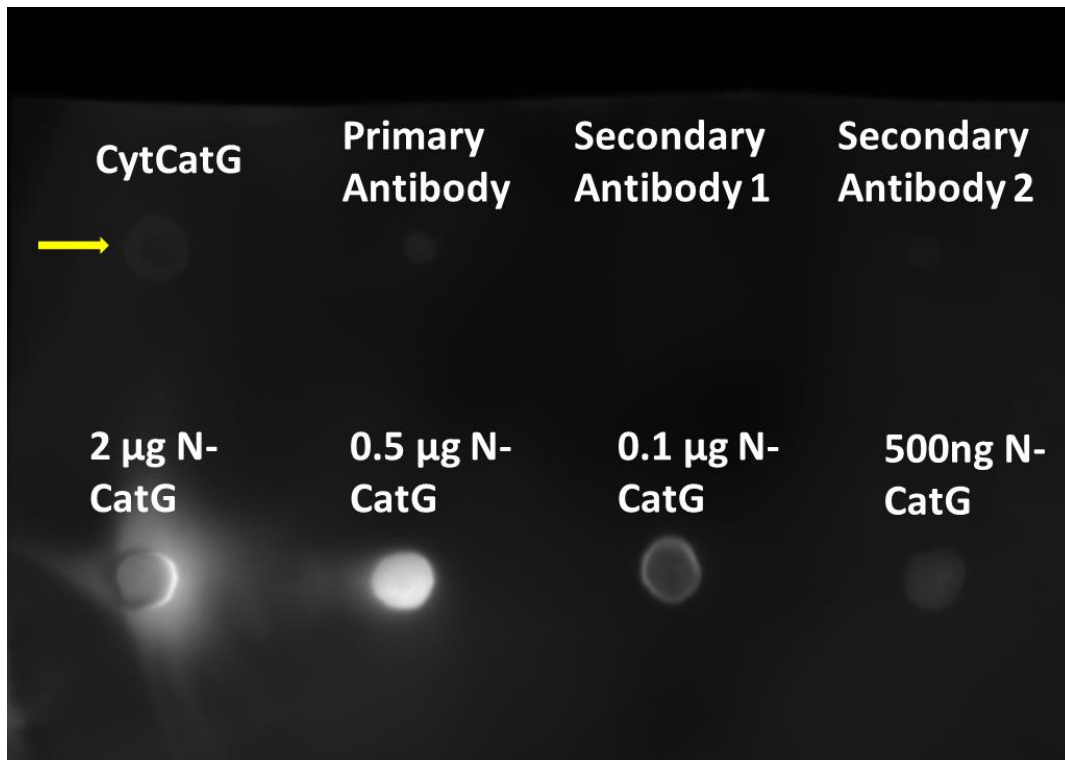


**Figure 12:** The elution profile from nickel IMAC purification of the 6His tag on the rhCatG fusion protein design. The absorbance at 280 nm (blue) and 410 nm (red) is graphed with the absorbance scale on the left and activity scale on the right. The activity (green) is graphed against its absorbance at 410 nm over time from activation assays. The high A280 readings from fractions 24-35 indicate another protein is co-purifying with the 6His-CytB5-FLAG-CatG fusion protein. The active fractions 32 through 39 were pooled.

## DOT BLOT

After several failed Western blots with no signal produced from the partially purified 6His-CytB5-FLAG-CatG fusion protein and native CatG positive control, a dot blot was performed to test for the efficacy of our primary and secondary antibodies. Native CatG samples were loaded to test for antibody signal, along with blotting the primary rabbit anti-CatG antiserum (Athens Research & Technology) and two secondary HRP-conjugated goat anti-rabbit polyclonal antibodies directly to the nitrocellulose membrane. Finally, a sample from the pooled IMAC fractions was also

analyzed. The results (Figure 13) prove that secondary antibody 1 (Proteintech) that had been used in the Western blotting attempts was not able to produce signal. The secondary antibody 2 (Pierce Biotechnology) functioned as expected in this dot blot.



**Figure 13:** Dot blot testing the effectiveness of our secondary antibodies, signal strength from a series of dilutions of CatG, and a sample of IMAC purified 6His-CytB5-FLAG-CatG. The only spot that did not show signal was the secondary antibody 1 (ProteinTech). The 6His-CytB5-FLAG-CatG signal was much weaker than the 500 ng sample of CatG suggesting a very small amount of expressed and purified 6His-CytB5-FLAG-CatG.

## CONCLUSIONS

Large scale fermentation of *P. pastoris* transformed with 6His-CytB5-FLAG-rhCatG resulted in the secretion of activatable rhCatG in the growth media, but in very low amounts ( $\mu\text{g/L}$ ). The activatable fusion protein was able to be partially purified with

Ni<sup>2+</sup> IMAC column and its elution was shown to correlate with absorbance readings at 280 nm and 410 nm as well as by activity for each fraction. A green colored protein from the media, however, was seen to co-elute with our rhCatG fusion protein, especially in the early fractions and it is thought to have an absorbance at 350 nm. A likely candidate for the identity of the green protein is alcohol oxidase 1 that is being expressed under the AOX1 promoter as the green color in the fermentation medium accumulates after the addition of methanol. After the sequential purification steps that would be needed to further purify rhCatG fusion protein from the green contaminate protein the final amount of rhCatG would be too little to do kinetic analysis, characterization, and further experiments with the C-terminal extension peptide.

The problem of low expression and troubleshooting at the purification step has been a reoccurring theme with the CatG project as similar fusion protein designs have yielded similar results. CatG's antifungal activity does not seem to be an issue due to the fact that the yeast cells grow sufficient and in a linear fashion as expected. The expression and secretion mechanisms of *P. pastoris* could be getting clogged, due to the challenge of keeping up with the demand of expression of CatG due to its high pI. In the neutrophil elastase project it was seen that the *P. pastoris* cells had a pink color and when the cells were lysed only minimal elastase activity was seen. Thus it could be entirely possible that the CytB5 domain is being expressed first (N-terminus) and expression is halted due to troubles in expression of the serine proteases. Another possibility of low expression could be enzyme degradation in the media by yeast secreted proteases. As these causes are merely speculation, further experimentation is needed to provide insight to the cause of low expression and thus bettering attempts to

enhance the expression of neutrophil serine proteases. Further studies should be done using qPCR to determine mRNA levels. By knowing the mRNA levels it could be determined if the low protein expression was due to the transcription/translation process or due to posttranslational degradation. By using the functional secondary antibody, further tests should be done by western blots to confirm the molecular weight and presence of rhCatG through the expression, purification, and activation steps.

This project using the CytB5 fusion domain has provided us with ideas for designing alternative expression constructs. Subsequent studies will employ attempted expression of an N-glycosylated rhCatG in hope of improving secretion and will use a newer *P. pastoris* strain SuperMan<sub>5</sub> (Biogrammatix) instead of X-33 *P. pastoris*, which tend to hyperglycosylate proteins. SuperMan<sub>5</sub> has been engineered to provide glycosylation patterns similar to those of human cells. The next fusion protein to try is a prancer purple colored protein (DNA 2.0). Prancer purple protein is known for its distinct purple color, instead of the red color of CytB5. Instead of using a 6His tag, a 10His tag is being engineered at its N-terminus to enhance the binding to the Ni<sup>2+</sup> IMAC column for better separation during purification. Additionally, expression of this new fusion protein gene construct will be placed under control of the constitutively active GAP promoter thus the fusion protein will be expressed and secreted at all times. It is thought that slower and constant expression will result in better secretion of the 10His-PrancerPurple-FLAG-rhCatG fusion. Without the need for transitioning to the methanol fed-batch phase the cells can theoretically grow to a higher density on glycerol as the sole carbon source from start to finish and produce more secreted product (Love et al., 2012). This new 10His-Prancer-FLAG-rhCatG fusion protein design has been

successfully ligated into the pJ915 plasmid and has been electroporated into the SuperMan<sub>5</sub> *P. pastoris*. Expression trials for this construct are currently in process.

## References

- BANGALORE, N., TRAVIS, J., ONUNKA, V. C., POHL, J. & SHAFER, W. M. (1990). Identification of the primary antimicrobial domains in human neutrophil cathepsin G. *J Biol Chem* **265**, 13584-8.
- BANK, U., KÜPPER, B., REINHOLD, D., HOFFMANN, T. & ANSORGE, S. (1999). Evidence for a crucial role of neutrophil-derived serine proteases in the inactivation of interleukin-6 at sites of inflammation. *FEBS Lett* **461**, 235-40.
- BORREGAARD, N. & COWLAND, J. B. (1997). Granules of the human neutrophilic polymorphonuclear leukocyte. *Blood* **89**, 3503-21.
- BORREGAARD, N., THEILGAARD-MÖNCH, K., COWLAND, J. B., STÄHLE, M. & SØRENSEN, O. E. (2005). Neutrophils and keratinocytes in innate immunity--cooperative actions to provide antimicrobial defense at the right time and place. *J Leukoc Biol* **77**, 439-43.
- DE GARAVILLA, L., GRECO, M. N., SUKUMAR, N., CHEN, Z. W., PINEDA, A. O., MATHEWS, F. S., DI CERA, E., GIARDINO, E. C., WELLS, G. I., HAERTLEIN, B. J., KAUFFMAN, J. A., CORCORAN, T. W., DERIAN, C. K., ECKARDT, A. J., DAMIANO, B. P., ANDRADE-GORDON, P. & MARYANOFF, B. E. (2005). A novel, potent dual inhibitor of the leukocyte proteases cathepsin G and chymase: molecular mechanisms and anti-inflammatory activity in vivo. *J Biol Chem* **280**, 18001-7.
- GRECO, M. N., HAWKINS, M. J., POWELL, E. T., ALMOND, H. R., CORCORAN, T. W., DE GARAVILLA, L., KAUFFMAN, J. A., RECACHA, R., CHATTOPADHYAY, D., ANDRADE-GORDON, P. & MARYANOFF, B. E. (2002). Nonpeptide inhibitors of cathepsin G: optimization of a novel beta-ketophosphonic acid lead by structure-based drug design. *J Am Chem Soc* **124**, 3810-1.
- GULLBERG, U., LINDMARK, A., LINDGREN, G., PERSSON, A. M., NILSSON, E. & OLSSON, I. (1995). Carboxyl-terminal prodomain-deleted human leukocyte elastase and cathepsin G are efficiently targeted to granules and enzymatically activated in the rat basophilic/mast cell line RBL. *J Biol Chem* **270**, 12912-8.
- HOF, P., MAYR, I., HUBER, R., KORZUS, E., POTEPA, J., TRAVIS, J., POWERS, J. C. & BODE, W. (1996). The 1.8 Å crystal structure of human cathepsin G in complex with Suc-Val-Pro-PheP-(OPh)<sub>2</sub>: a Janus-faced proteinase with two opposite specificities. *EMBO J* **15**, 5481-91.
- JOHNSON, D. A. (2006). Human mast cell proteases: activity assays using thiobenzyl ester substrates. *Methods Mol Biol* **315**, 193-202.
- KORKMAZ, B., HORWITZ, M. S., JENNE, D. E. & GAUTHIER, F. (2010). Neutrophil elastase, proteinase 3, and cathepsin G as therapeutic targets in human diseases. *Pharmacol Rev* **62**, 726-59.
- KORKMAZ, B., MOREAU, T. & GAUTHIER, F. (2008). Neutrophil elastase, proteinase 3 and cathepsin G: physicochemical properties, activity and physiopathological functions. *Biochimie* **90**, 227-42.
- LOVE, K. R., POLITANO, T. J., PANAGIOTOU, V., JIANG, B., STADHEIM, T. A. & LOVE, J. C. (2012). Systematic single-cell analysis of *Pichia pastoris* reveals secretory capacity limits productivity. *PLoS One* **7**, e37915.



- MITRA, A., CHAKRABARTI, K. S., SHAHUL HAMEED, M. S., SRINIVAS, K. V., SENTHIL KUMAR, G. & SARMA, S. P. (2005). High level expression of peptides and proteins using cytochrome b5 as a fusion host. *Protein Expr Purif* **41**, 84-97.
- NEWMAN, S. L., GOOTEE, L., GABAY, J. E. & SELSTED, M. E. (2000). Identification of constituents of human neutrophil azurophil granules that mediate fungistasis against *Histoplasma capsulatum*. *Infect Immun* **68**, 5668-72.
- NILES, A. L., MAFFITT, M., HAAK-FRENDSCHO, M., WHEELESS, C. J. & JOHNSON, D. A. (1998). Recombinant human mast cell tryptase beta: stable expression in *Pichia pastoris* and purification of fully active enzyme. *Biotechnol Appl Biochem* **28** ( Pt 2), 125-31.
- PADRINES, M., WOLF, M., WALZ, A. & BAGGIOLINI, M. (1994). Interleukin-8 processing by neutrophil elastase, cathepsin G and proteinase-3. *FEBS Lett* **352**, 231-5.
- PHAM, C. T. (2006). Neutrophil serine proteases: specific regulators of inflammation. *Nat Rev Immunol* **6**, 541-50.
- SMITH, E. (2013). Bioengineering the Expression of Active Recombinant Human Cathepsin G, Enteropeptidase, Neutrophil Elastase, and C-Reactive Protein in Yeast, East Tennessee State University Quillen College of Medicine.
- STRATTON, J., CHIRUVOLU, V. & MEAGHER, M. (1998). High cell-density fermentation. *Methods Mol Biol* **103**, 107-20.
- WITTAMER, V., BONDUE, B., GUILLABERT, A., VASSART, G., PARMENTIER, M. & COMMUNI, D. (2005). Neutrophil-mediated maturation of chemerin: a link between innate and adaptive immunity. *J Immunol* **175**, 487-93.