2020

Exploring the biological function of carboxypeptidase O: analysis of expression and association with chylomicrons and lipid droplets

Erika Bauza Nowotny
Andrews University, bauza@andrews.edu

ABSTRACT

EXPLORING THE BIOLOGICAL FUNCTION OF CARBOXYPEPTIDASE O: ANALYSIS OF EXPRESSION AND ASSOCIATION WITH CHYLOMICRON AND LIPID DROPLETS

By

ERIKA BAUZA NOWOTNY

Chair: Peter J. Lyons, Ph.D.

ABSTRACT OF GRADUATE STUDENT RESEARCH

Thesis

Andrews University

College of Arts and Sciences

Title: EXPLORING THE BIOLOGICAL FUNCTION OF CARBOXYPEPTIDASE O: ANALYSIS OF EXPRESSION AND ASSOCIATION WITH CHYLOMICRON AND LIPID DROPLETS

Name of researcher: Erika Bauza Nowotny

Name and degree of faculty chair: Peter J. Lyons, Ph.D.

Date completed: July 2020

Carboxypeptidase O (CPO) is a membrane-bound peptidase that cleaves acidic and polar C-terminal amino acids of peptides; however, its biological function remains unknown. CPO is strongly expressed in the small intestine, where it has been proposed to participate in digestion events at the brush border by extracellular cleavage of dietary peptides. At the subcellular level, CPO is anchored to the inner leaflet of the endoplasmic reticulum (ER) membrane, where it has been shown to associate with lipid droplets (LDs). The ER membrane also represents the site of

chylomicron formation, a process that shares many similarities with that of LDs biogenesis. We sought to investigate a potential role of CPO in LD and chylomicron formation. Using a stably-expressing human colon carcinoma (Caco-2) cell model, we evaluated CPO association with LDs and apoB, a chylomicron biomarker, by immunofluorescence microscopy. The association of CPO with LDs was dependent on the time elapsed post-feeding (ANOVA, $p < 0.05$), and was lowest at 6 h (17.9%), reaching a maximum at 12 h (29.8%) and later slightly decreasing at 24 h (26.3%). Our results indicate that CPO associated with apoB, further showing a similar association pattern as that of LDs at the three time points. However, we express a limited level of confidence in these results since the immunocytochemistry images presumably suggest lower association levels than those reported in the quantification data. Overall, our findings strengthen the hypothesis that CPO plays an intracellular role in the formation of lipid droplets in enterocytes.

To better understand the potential biological role of CPO, we performed an in-depth analysis of data from several gene expression databases and platforms. RNA-Seq and microarray expression data showed high expression in the ileum epithelium and mucosa. Modulation of CPO expression, which was mainly characterized by downregulation in the ileum, was found in ulcerative colitis and Crohn's disease, with the latter showing stronger modulation. RNA-Seq data also showed prominent expression in Peyer's patches and basophils, which suggested a promising role of CPO in immunity.

Andrews University
College of Arts and Sciences

EXPLORING THE BIOLOGICAL FUNCTION OF CARBOXYPEPTIDASE O: ANALYSIS OF EXPRESSION AND ASSOCIATION WITH CHYLOMICRON AND LIPID DROPLETS

A Thesis

Presented in Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Erika Bauza Nowotny

July 2020

EXPLORING THE BIOLOGICAL FUNCTION OF CARBOXYPEPTIDASE O: ANALYSIS OF EXPRESSION DATA AND ASSOCIATION WITH CHYLOMICRON AND LIPID DROPLETS

A thesis
presented in partial fulfillment
of the requirements for the degree
Master of Science

by

Erika Bauza Nowotny

APPROVAL BY THE COMMITTEE:

Peter Lyons Digitally signed by Peter Lyons Date: 2020.07.13 09:44:39 -04'00'

Peter J. Lyons Ph.D., Chair

Denise L Smith Digitally signed by Denise L Smith Date: 2020.07.13 12:17:08 -04'00'

Denise L. Smith, Ph.D.

Benjamin Navia Digitally signed by Benjamin Navia Date: 2020.07.13 09:52:46 -04'00'

Benjamin Navia, Ph.D.

July 13, 2020

Date approved

For Opa, who left us shortly after the defense of this thesis

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ApoB	Apolipoprotein B
ApoB100	Apolipoprotein B100
ApoB48	Apolipoprotein B48
CM	Chylomicron
CP	Carboxypeptidase
CPE	Carboxypeptidase E
CPO	Carboxypeptidase O
ddH_2O	Double-distilled H_2O
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
dNTP	Deoxynucleoside triphosphates
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
G418	Geneticin
GFP	Green fluorescent protein
GPI	Glycosylphosphatidylinositol
hCPO	Human CPO
HEK293T	Human embryonic kidney 293T cell line
HPA	Human Protein Atlas
HRP	Horseradish peroxidase
LD	Lipid droplet

MDCK	Madin-Darby canine kidney
NL	Neutral Lipids
PBS	Dulbecco’s modified Phosphate Buffered Saline
pcDNA	Plasmid DNA
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PMSF	Prehenylmethylsulfonyl fluoride
ProCPB2	Procarboxypeptidase B
pTPM	Protein-coding transcripts per million
RNA-Seq	RNA sequencing
SDS	Sodium dodecyl sulfate
WB	Western blot
WT	Wild type

ACKNOWLEDGEMENTS

I want to extend my most sincere thank you to my mentor and advisor, Dr. Peter Lyons, for his continued support and guidance throughout these years. I have found your passion to advance scientific knowledge to be contagious. Your mentorship has helped me to grow as a researcher and has been crucial in my decision to pursue a Ph.D.

I also want to express my gratitude to my committee members, Dr. Benjamin Navia, with whom I had many stimulating conversations with, and Dr. Denise Smith, for being a helping hand in many occasions.

I am really grateful for the opportunity to continue to work with Dr. Daniel Gonzalez-Socoloske during my time as a graduate student.

Thank you to all of my professors for helping me grow intellectually and spiritually, for fostering my love for science, and for giving me the opportunity share my knowledge with other students. I am also grateful for my friends and family members who have encouraged me throughout the process of thesis writing.

Thank you to my dearest friends Linnea and Adriana, who have supported me endlessly during this process and have encouraged me to finish this thesis during a very challenging time.

Lastly, I would like to thank you the School of Graduate studies and Research for their financial support.

CHAPTER 1

INTRODUCTION

1. Carboxypeptidases

Carboxypeptidases are enzymes characterized for cleaving the C-terminal amino acid of peptides and proteins and are known to have diverse physiological functions such as cell signaling and digestion (Fricker, 2011). Carboxypeptidases are grouped in two major categories: serine carboxypeptidases, which contain a serine residue at the active site, and metallocarboxypeptidases, which contain a zinc atom. The M14 is an important family of metallocarboxypeptidases with unique physiological functions in humans (Fig. 1). Substrate specificity exists within the members of this family, with some members exhibiting preference for basic residues (i.e. CPB2, CPE), acidic residues (i.e. CPO), or non-polar aliphatic/aromatic residues (i.e. CPA1) (Fricker, 2011).

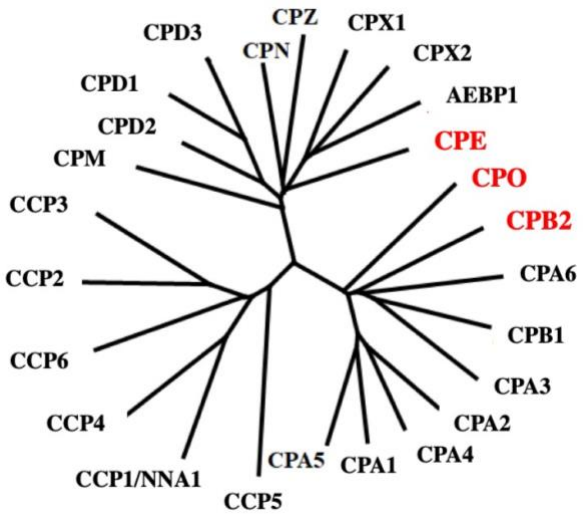

Figure 1. Phylogeny of the M14 metallocarboxypeptidase family. In red, CPE, CPB2, and CPO. Contrary to CPO, CPE and CPB2 are well-characterized carboxypeptidases with important physiological functions such as neuropeptide processing and proteolytic cleavage.

A number of studies have successfully described the physiological functions of several M14 carboxypeptidases. For example, studies by Ji et al. (2017) and Chen et al. (2001) identified the role of carboxypeptidase E (CPE) in the synthesis of neuropeptides and peptide hormones, vesicular transport, cellular signaling, and neuroprotection. CPE was originally named enkephalin convertase after it was found to proteolytically cleave the C-terminal basic amino acid residues of enkephalin precursors within adrenal chromaffin granules of cattle (Fricker and Snyder, 1982). In a subsequent study in rats, CPE was found to participate in conjunction with endoproteases in mid stages of development in rats, where they processed neuropeptide precursors in the brain as well as proproteins in multiple peripheral locations (Zheng et al., 1994). Due to its supporting role in proinsulin processing (Harding and Ron, 2002), mutations in the CPE gene have been directly associated with development of type 2 diabetes in mice and humans (Fricker et al., 1996; Harding and Ron, 2002). In addition, CPE has been reported to provide neuronal protection against oxidative stress and cell death (Cong et al., 2017). Moreover, CPE participates in anterograde transport of synaptic vesicles, which allows for their recycling or degradation (Ji et al., 2017).

Another M14 member with a significant biological role is thrombin-activatable procarboxypeptidase B (proCPB2) also known as thrombin-activatable fibrinolysis inhibitor (TAFI). Following activation by proteolytic cleavage, CPB2 mediates several physiological functions, namely fibrinolysis, coagulation, and inflammation (Novakovic et al., 2012). Its inhibitory role in the fibrinolysis cascade is executed through cleavage of the C-terminal lysine residues from partially degraded fibrin, hence preventing further fibrin degradation by other enzymes (Leung et al., 2008; Novakovic et al., 2012). In addition, CPB2 is said to work in conjunction with protein C during injury response at the vascular epithelium: protein C reduces

the clotting while CPB2 protects the fibrin clot from premature lysis (Leung et al., 2008). A potential role of CPB2 in inflammation has been proposed after observation of proteolysis of proinflammatory mediators such as bradykinin, complement C3a and C5a, and thrombin-cleaved osteopontin (Novakovic et al., 2012).

Overall, a large number of physiological functions of carboxypeptidases (i.e. CPB2 and CPE) have been described to date. While some studies continue to further investigate well-characterized members of the M14 family, a new interest has emerged to investigate less-known members. Thus, Lyons and Fricker (2011) recently identified carboxypeptidase O (CPO), a previously uncharacterized member of the M14 family and closely related to CPB2 (Fig. 1). CPO was described as a fully functional peptidase that cleaves acidic and polar C-terminal amino acids of peptides (Burke et al., 2018; Lyons and Fricker, 2011). In addition, CPO was shown to be expressed on the small intestine, more specifically on the apical surface of enterocytes, and anchored to the membrane through a glycosylphosphatidylinositol (GPI) modification (Lyons and Fricker, 2011). Taken together, these findings allude to CPO working extracellularly, likely cleaving dietary peptides in the intestine (Garcia-Guerrero et al., 2018). Overall, there is a high possibility of CPO working extracellularly given that 1) GPI anchored proteins are almost exclusively present on the outer leaflet of the plasma membrane and 2) its location on the apical surface coincides with that of most GPI-anchored proteins in polarized epithelial cells (Zurzolo and Simons, 2016).

In addition to the proposed extracellular function cleaving dietary peptides, CPO might also play an intracellular role. This possibility was suggested after observation of intracellular expression of CPO in canine-derived epithelial cells where CPO was GPI-anchored to the inner leaflet of the ER membrane and associated with LDs within this compartment (Burke et al.,

2018). Overexpression of CPO led to an increase in the number of LDs, as determined by immunofluorescence microscopy. This finding suggested a physiological function of CPO in the biogenesis of lipid droplets.

Burke et al. (2018) further hypothesized that CPO could play a role in the formation of chylomicrons. This suggestion was based on the many similarities between their formation processes. The most accepted model states that their formation begin due to an accumulation of neutral lipids (NL) on the ER membrane, grow to form a spherical structure, and ultimately bud off (Saka and Valdivia, 2012). Conceptually, the most significant difference between these processes is in the direction of budding. If the nascent droplet buds off outwards to the cytosol, it will become a cytosolic LD, whereas if it buds off into the ER lumen, the droplet will become a chylomicron. Given the similarity in the budding processes and the resulting similar composition – a triglyceride core surrounded by a phospholipid monolayer – it is legitimate to hypothesize that similar membrane proteins could influence their formation. Thus, CPO might have a role in formation of both chylomicrons and lipid droplets, namely through cleavage of the C-terminal residue of a membrane protein that is known to be associated with their formation.

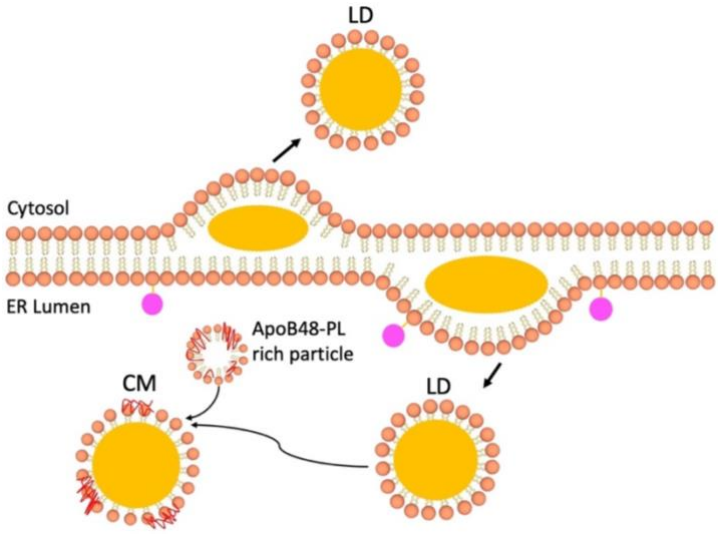

Figure 2. Model for the formation of lipid droplets from the ER membrane. The accumulation of neutral lipid (yellow spheres) between the leaflets of the membrane will form a spherical structure that buds off. If the lipid droplet buds into the lumen, it will later fuse with an apo-B48-phospholipid (PL) rich particle and form a chylomicron (CM). CPO (pink) is attached to the inner leaflet through a GPI anchor (yellow).

2. Digestion

The digestive processes involved in food breakdown and absorption involve a large number of structures, ranging from organs to accessory glands and include a wide array of enzymes. Digestion first begins in the mouth and moves through the stomach, aided by pancreatic secretions, until finally arriving to the small intestine, where chemical digestive processes continue to further break down dietary disaccharides, polypeptides, and emulsified fats (Tortora, 2009).

a. Digestion in the brush border

An important part of digestion in the small intestine occurs on the apical surface of enterocytes, more specifically at the brush border. Although the apical surface has been considered the hub for final digestion events, recent evidence points to an additional digestion

mechanism that occurs in the periapical space. Hooton et al. (2015) describe this mechanism, which lies on the vesicularization of the apical membrane and subsequent movement of the vesicles into the periapical space. This allows the brush border enzymes, now found on the membrane of the vesicles, to perform catalysis in the periapical space.

As a key structure in the end stages of digestion, the small intestinal brush border is characterized by abundant and diverse enzymatic activity. The digestive enzymes present in the brush border finalize digestion of carbohydrates, proteins and lipids, and can be categorized into three major groups: disaccharidases /oligosaccharidases, peptidases, and lipases (Hooton et al., 2015). Carbohydrate digestion is finalized in the brush border by action of disaccharidases such as sucrase, lactase, and maltase, which break down disaccharides into monosaccharides. In addition, α-dextrinase acts on small fragments of starch producing glucose (Tortora, 2009). Likewise, protein digestion is completed by aminopeptidase A, and dipeptidases, resulting in mostly free amino acids along with few dipeptides and tripeptides (Tobey et al., 1986). In contrast, lipid digestion can be considered finalized as lipids, which are assembled in spherical structures called micelles, arrive to the brush border (Tortora, 2009).

b. Absorption in the brush border

When nutrients are broken down into the most simple units (i.e. monosaccharides, fatty acids, and amino acids), they are able to pass through the enterocytes and eventually reach the blood or lymph. This process is known as absorption and it entails different forms of transport. For example, monosaccharides can be transported from the lumen through the apical membrane in two different ways. Glucose and galactose pass the apical membrane and enter the enterocytes via secondary active transport with Na^+ through Sodium-Dependent glucose transporter 1

(SGLT1) while fructose is able to enter via facilitated diffusion through Glucose Transporter 5 (Tortora, 2009). These monosaccharides leave through the basolateral membrane via facilitated diffusion using GLUT2 and are subsequently absorbed into the bloodstream through capillaries. In the case of amino acids, they may be actively transported into the enterocyte, or they will enter alongside Na^+ via a symporter. As transporters discriminate amongst different classes of amino acids, a large cohort of transporters have been identified for neutral amino acids (B^0AT1), lysine, arginine, ornithine, and cystine (rBAT, and $B^{0,+}AT$), proline and hydroxyproline (IMINO), aspartic acid and glutamic acid (EAAT3) (Broer, 2008). In contrast, dipeptides and tripeptides are transported via H^+ symporter such as PEPT1 (Hooton et al., 2015). Amino acids, dipeptides and peptides collectively diffuse out of the enterocyte and enter the blood capillaries.

At the brush border, monoglycerides and free fatty acids (e.g. 2,3-diacylglycerol and 2-monoacylglycerol) diffuse out of the micelles into the enterocyte through the apical membrane, whereas glycerol will freely diffuse in and out of the enterocyte to enter the blood capillary (Meisenberg and Simmons, 2016). Unlike the aforementioned nutrients, which are simply transported in and out the enterocytes, free fatty acids undergo a series of chemical modifications while passing through the enterocytes (Bayly, 2014). Through multiple reactions, free fatty acids are converted into triglycerides. Triglycerides collectively assemble with phospholipids and apolipoproteins, thus giving rise to chylomicrons (Tortora, 2009). Due to their considerably large size (ranging from 200-1,000 mm in diameter) chylomicrons must exit the enterocytes via exocytosis and are further unable to enter the capillaries (Randolph and Miller, 2014). Instead, chylomicrons penetrate the adjacent lacteals through large pores located on their tips and are transported through the lymphatic system to peripheral tissues such as cardiac, adipose, and muscle (Randolph and Miller, 2014).

3. Chylomicrons

Over the past couple decades, there has been an increasing interest to study the formation and metabolism of chylomicrons. This has been mostly due to the recent detection of chylomicron upregulation in widespread diseases, namely type 2 diabetes and atherosclerosis (Dash et al., 2015). Clinical conditions involving dyslipidemia such as insulin resistance and type 2 diabetes have been characterized by increased secretion of chylomicrons and apolipoprotein B, amongst other lipoproteins (Hussain, 2014). This finding has important implications, since dyslipidemia is one of the greatest risk factors for cardiovascular disease (Dash et al., 2015; Mooradian, 2009; Xiao et al., 2011), currently labeled as the number one leading cause of death in the U.S according to the most recent report by the CDC (Kochanek et al., 2019). Additionally, chylomicron dysregulation further extends to less common diseases, such as genetically transmitted familial hyperchylomicronemia syndrome (Brahm and Hegele, 2015) and familial dysbetalipoproteinemia (Koopal et al., 2017).

Chylomicrons are the largest and least dense lipoproteins particles and are composed of a core of neutral lipids, namely triacylglycerols (85% of the total weight), enclosed by a phospholipid monolayer (8%) with about 1% of associated apolipoproteins (Demignot et al., 2014). A form of apolipoprotein B, namely apoB48, is characteristic of chylomicrons and its presence serves as a marker to differentiate them from other closely-related lipoproteins (Nakajima et al., 2014). For example, while apoB48 is the main protein in chylomicrons, a longer form of the protein, apoB100, is the main protein in very-low density lipoproteins (VLDL) and low-density lipoproteins (Randolph and Miller, 2014). ApoB48 plays a key structural role in chylomicrons (Xiao, Stahel, & Lewis, 2018) while other apolipoproteins such

as ApoA-IV assemble on the surface of chylomicrons and play a role in their formation by determining their size and regulating their secretion (Dash et al., 2015; Xiao et al., 2018).

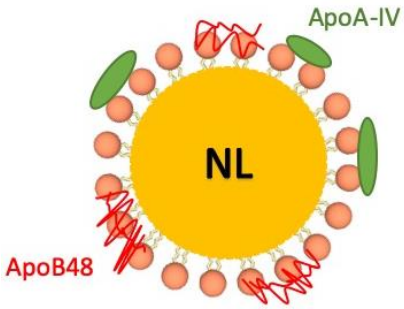

Figure 3. Schematic illustration of a chylomicron. A phospholipid monolayer surrounds a core of neutral lipids which is mainly composed of triacylglycerols and sterol esters. Apolipoproteins apoB48 and apo-AIV located on the surface of the chylomicron play important structural and regulatory roles, respectively.

While much is known about the physiology and composition of chylomicrons, their biogenesis is not entirely understood. The most recent model for chylomicron formation suggests a two-step assembly process. A lipid droplet buds off the inner leaflet of the ER membrane into the lumen, where it will fuse with a lipidated apoB48-rich particle in a reaction catalyzed by microsomal triglyceride transfer protein (MTP), eventually giving rise to a pre-chylomicron (Fig. 4; Demignot et al., 2014). After this, the pre-chylomicrons are assembled into pre-chylomicron transport vesicles (PCTV) and trafficked to the Golgi apparatus (Cartwright and Higgins, 2001; Demignot et al., 2014).

4. Lipid droplets

Lipid droplets (LDs) and chylomicrons are formed by similar mechanisms. When a lipid droplet buds off to the ER lumen, it will give rise to a chylomicron; however, if it buds off to the cytosol, it will become a cytosolic lipid droplet. LDs constitute the major type of cellular organelle that stores excess neutral lipids namely triacylglycerols and sterol esters (Vevea et al.,

2015). This storage mechanism is considered protective, since it avoids disruption of membrane integrity by over-accumulation of neutral lipids (Welte, 2015). In eukaryotes, *de novo* formation of LDs occurs due to an accumulation of neutral lipids, on defined regions between the leaflets of the ER (Roux and Loewith, 2017). The most accepted model proposes that the nascent LD forms a lens, which grows in size until the LD is fully formed and released (Saka and Valdivia, 2012). LDs that bud off to the cytosol and accumulate on the apical side of the enterocyte, where they will serve as transient lipid reservoirs (Hussain, 2014).

Nonetheless, LDs are versatile as they have been recently found to have functions beyond lipid storage. Several studies have identified many LD functions such as storage of vitamins and molecules involved in cellular signaling, as well as assembly of viruses (Saka and Valdivia, 2012; Welte and Gould, 2017). For example, Welte and Gould (2017) discuss the involvement of LDs in lipid signaling by means of storage of eicosanoid precursors. Although these precursors, namely arachidonic acid rich triglycerides, may be obtained through the breakdown of phospholipids, LDs represent an important additional source. Moreover, LDs constitute major sites of vitamin A storage (in the form of retinyl esters) with about 80 % of the total reservoir found in hepatic cells (Welte & Gould, 2017). Cholesteryl esters also stored in the LD core serve as sources of cholesterol for synthesis of steroid hormones (Kraemer et al., 2013). A well-studied case of LDs acting as assembly platforms during viral infections is that of hepatitis C. Studies have shown that localization of the virus core protein on the surface of LDs is necessary for assembly of new immature viruses (virions) on the host cell (Saka and Valdivia, 2012). Due to their many physiological roles, LD formation is tightly regulated. Accumulation of lipid droplets has been associated with major diseases (e.g. type II diabetes, obesity, atherosclerosis, and fatty

liver disease) which has led to an increasing interest in studying their physiology (Meyers et al., 2017; Welte and Gould, 2017).

Despite continuing efforts to elucidate the molecular processes behind LD biogenesis, many questions remain unanswered today (Hussain, 2014; Roux and Loewith, 2017; Wilfling et al., 2014). Yeast represents a well-studied eukaryote model for lipid metabolism, and its use in the study of this metabolism is suitable given its high conservation in both humans and yeast (Karathia et al., 2011). To date, the yeast model has identified many key proteins that participate in LD formation. For example, Lor1p, a putative homolog of the human lecithin cholesterol acyltransferase (LCAT), mediates the conversion of diacylglycerol to triacylglycerol (Oelkers et al., 2000; Roux and Loewith, 2017). The model of lipid formation in yeast is the most detailed model described to date; however, a number of studies consider this model to be incomplete since it has failed to describe crucial events in droplet formation, such as the excision of the nascent LD, thus pointing to a need for further research in this area (Meyers et al., 2017; Roux & Loewith, 2017; Wilfling et al, 2014).

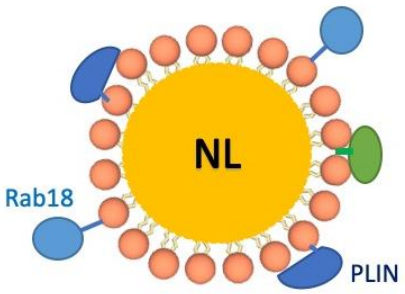

Figure 4. Schematic illustration of a generic LD. A phospholipid monolayer surrounds a core of neutral lipids (NL) which is mainly composed of triacylglycerols and sterol esters. While perilipins (PLIN) and Rab18 proteins are abundant on the LD surface, other LD-associated proteins (here depicted in green) such as SNARES can be inserted in the phospholipid monolayer.

In this study, I investigate a hypothesis initially made by Burke et al. (2018), which suggested that CPO is involved in the formation of chylomicrons. In addition, I build on previous work from Burke et al. (2018) by evaluating the association of CPO and LDs in undifferentiated human colon carcinoma (Caco-2) cells. Here, I show that CPO stably-expressing Caco-2 cells exhibit CPO association with both LD and apoB, with association in LDs being significantly dependent upon time elapsed after feeding. A comprehensive analysis of CPO expression data showed prominent expression in the ileum and modulation of expression in ulcerative colitis and Crohn's disease as well as expression in basophils and Peyer's patches, which suggest promising implications of CPO in immunity.

CHAPTER 2

METHODS

1. Compilation of CPO expression data

CPO expression data was obtained from several databases which included RNA sequencing (RNA-Seq) and microarray data. RNA-Seq data was obtained from publicly available datasets such as Human Protein Atlas (HPA; www.proteinatlas.org) which allow access to transcriptomics data, including a compilation of data from prominent datasets such as Genotype-Tissue Expression (GTEx), HPA, and FANTOM5 CAGE (Pontén et al., 2011). Additionally, data was gathered from the Bgee website (Bastian et al., 2008), which compiles RNA-seq data from the Gene Expression Omnibus (GEO; NCBI) and GTEx. Genevestigator, a powerful transcriptomics analysis tool (Hruz et al., 2008; Zimmermann et al., 2005), was used to access the mRNASeq_HUMAN_GL-1 platform, thus providing an additional RNA-Seq dataset.

Microarray data was also included in the expression analysis due to its notable prevalence in transcriptomic profiling. Microarray data, although providing less sensitivity for low-expression genes, constitutes a reliable choice (Zhao et al., 2014). The microarray data was retrieved from three platforms. The first one was the Affymetrix Human Genome U133 Plus 2.0 Array, a microarray platform in Genevestigator. The second one was the Bgee website, which compiled data from the GEO and the ArrayExpress from European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI). The third one was the Gene Expression database of Normal and Tumor Tissues (GENT2; www.gent2.appex.kr) database, which includes gene expression data from healthy and cancerous tissues. GENT2 compiles

affymetrix microarray data from U133 Plus 2.0 as well as the U133A platforms (Park et al., 2019).

2. Site-directed mutagenesis

The goal of site-directed mutagenesis was to introduce a point mutation, E310Q, into plasmid DNA (pcDNA; Fig. 5). First, custom primers were manually designed. To do this, the human CPO (hCPO) sequence was translated into its amino acid sequence using the EMBOSS explorer interface. The amino acid of interest E310 was located and aligned with the corresponding codon on the forward DNA strand. Overlapping primers were created following ThermoFisher's protocol and included the following criteria: 15 base pair complementary regions with the point mutation located in the middle along with non-overlapping bases at the 3' end (underlined). The final primer sequences were hCPO-E310Q-F3 5'-CGTTTCAGCTGAGGGACAGTGGAAC-3'; hCPO-E310Q-R3 5'-CCCTCAGCTGAAACGTATATGAGAAGGG-3'. PCR amplification was used to create the mutation, and two 50 μL reactions were prepared. Each reaction contained 1.5 μL Platinum Superfi Polymerase (Thermofisher), 2mM dNTP mix, 30 μL 5x Superfi buffer (Thermofisher), 1 μL template DNA, and 0.5 μM of each primer, and one reaction contained 3% DMSO to enhance amplification. A BioRad 100 Thermocycler was used with the following program conditions: 98°C (30 s), 98 °C (10 sec), 67.1 °C (10 sec), 72 °C (195 s) repeated for 25 cycles, and final extension at 72 °C for 5 min. The PCR products were resolved on a 1% agarose gel and visualized with ethidium bromide/UV. After confirmation of amplification, PCR products were digested with Dpn1 to remove the parental plasmid DNA, which did not contain the mutations.

3. Ethanol precipitation

To purify the digested PCR product, ethanol precipitation was performed. The PCR product (40 µL) was mixed with 0.3 M NaAC (pH 5.2), 250 µL of cold 100% ethanol (-20 °C), and 50 µL of ddH_2O. Following incubation at -80 °C for 30 min, the sample was centrifuged at 21,130 g for 30 min. The precipitate was rinsed with 500 µL of 70% ethanol, centrifuged 21,130 g for 15 min, and resuspended in 10 µL ddH_2O.

4. Bacterial transformation

DH5α competent *E.coli* cells were transferred from the -80 °C freezer and allowed to thaw on ice for 15 min. Two controls, a positive and a negative, were prepared by addition of 25 µL DH5α cells to 5 µL pcDNA [hCPO] (1 ng/µL) and 5 µL ddH_2O respectively. Additionally, two experimental samples, one with 5 µL digested PCR product and one with 5 µL of the digested + purified PCR product were prepared by addition of 50 µL DH5α cells. The samples were gently mixed, incubated on ice for 30 min, and heat shocked for 40 s at 42 °C. Following a 5 min incubation on ice, 200 µL of sterile Luria Broth (LB) were added and mixed by stirring with the pipet tip. The samples were incubated at 37°C for 60 min and then spread on LB/ampicillin plates, which were incubated at 37 °C overnight. The plates were inspected for colony growth 16 h following plating.

5. Plasmid preparation by NID method

Colonies were randomly selected from both the plates: three colonies from the plate with digested DNA and three colonies from the plate with digested + purified DNA. In culture tubes, 2 mL of sterile LB broth and 2 µL ampicillin were mixed. The colonies were individually

scooped from the plates with a flamed pipet tip and resuspended in the mixture contained within the culture tube. The tubes were later placed in a shaking incubator at 37 °C for 20 h.

A larger working stock of pcDNA-hCPO E310Q was produced by a large-scale bacterial transformation. The volumes of reagents and pcDNA (see above) were reduced by 5-fold, while bacterial cultures were grown in culture tubes with about 5-fold the volume of LB broth and ampicillin. After 12 h of incubation at 37°C, 1 mL of the culture was transferred to a flask containing 50-fold volume of LB broth and ampicillin and grown for 8 additional hours.

The contents of the culture tubes were transferred to 2 mL tubes and centrifuged at 21,130 g for 30 s. The pellets were resuspended in 150 μL of extraction buffer (5% sucrose, 50 mM EDTA, 50 mM Tris pH 8.0, 0.75 M NH_4Cl, 0.5% Triton X-100, 100 μg/mL lysozyme, 25 μg/mL RNAase A) and incubated at 65 °C for 5 min. Following centrifugation (10 min at 21,130 g), the pellet was removed with a sterile toothpick and mixed with 120 μL of isopropanol by vortexing. The mixture was centrifuged at 4,602 g and the supernatant was subsequently discarded. To the tube, added 500 μL 70% ethanol was added and it was centrifuged at 21,130 g for 10 min. In order to remove any supernatant left, the bulk of it was pipetted out and then the pellet was allowed to air dry for 10 min. The pellet was resuspended in 50 μL of EB buffer.

In the large-scale NID plasmid prep, the contents of the culture tubes were transferred to 50 mL tubes and centrifuged (3,000 g for 5 min at 0 °C). The pellet was resuspended in 5 mL of extraction buffer and the tube was incubated in a 65 °C water bath for 8 min. Following centrifugation (19,000 g for 30 min at 0 °C), the supernatant was transferred to a new tube, mixed with 7 mL of isopropanol. After the sample was centrifuged (19,000 g for 15 min at 0 °C), the pellet was resuspended in 1.5 mL of 70% ethanol and centrifuged at 15,000 rpm for 5 min. In

order to remove any supernatant left, the bulk of it was pipetted out and then the pellet was allowed to air dry for 10 min. The pellet was resuspended in 400 μL of EB buffer.

6. DNA quantification

DNA isolated from each of the six *E.coli* colonies was visualized on a 1% agarose gel. Each sample loaded contained 1 μL template DNA, 2 μL 6x loading dye and 7 μL ddH_2O. An image of the gel taken with AlphaImager imaging system was analyzed using ImageJ (NIH*).* This image processing program allowed quantification of the DNA by comparison to the marker lane (Tridye 1kb).

Spectrophotometry was used to calculate the quantity and purity of the pcDNA-hCPO E310Q, known as the ratio of absorbance at 260 nm over the absorbance at 280 nm (A260/A280). In a UV compatible cuvette, mixed 2 μL template DNA and 198 μL ddH_2O were mixed and absorbance measured at 260 nm and 280 nm.

7. Sanger sequencing

In accordance with Genescript's requirements (NJ, USA), only those DNA samples with optimal yields were selected and prepared for sequencing. Preparation was done by addition of 5 μL of T7 primer (10 μM) to 10 μL of template DNA. Sanger sequencing and subsequent alignment to the hCPO sequence (Accession: NM_173077.3) using BLAST allowed confirmation of the presence of the target mutation (Fig. 5).

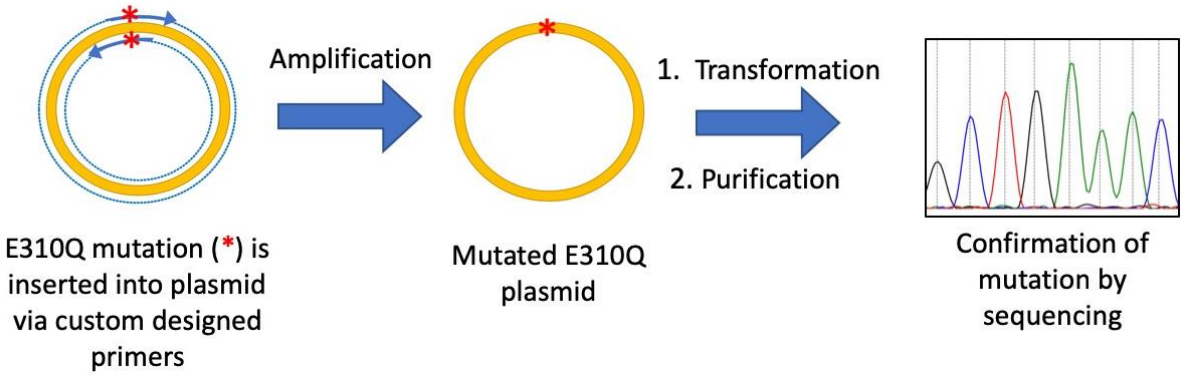

Figure 5. Site-directed mutagenesis is used to create the mutant pcDNA-hCPO E310Q. A set of custom partially overlapping primers with 3' overhangs (blue arrows) containing single point mutations (red) were used to amplify the plasmid. Transformation into competent cells allows to produce large quantities of the primer, which are subsequently purified. Presence of mutation is confirmed via Sanger sequencing

8. Maintenance of Caco-2 cell line

Caco-2 were grown in an incubator set to 37°C and 5% CO_2 and maintained in Dulbecco's Modified Eagle Medium supplemented with 10% Fetal Bovine Serum (Sigma Aldrich) and 1% Penicillin/Streptomycin (Sigma Aldrich). The cells were passaged at 1:5 or 1:10 (vol/vol) when ≅ 80 % confluency was reached. To passage the cells, the media was removed and the cells were washed twice with Dulbecco's modified Phosphate Buffered Saline (PBS; Sigma) and trypsinized (Trypsin-EDTA 1x; Sigma) for 6 min in the incubator. The cells were subsequently resuspended in fresh complete medium and seeded at desired density. Medium was regularly changed every 48 h.

9. Transfection of Caco-2 cells

Caco-2 cells were seeded in a 6-well plate at a density of 1.5 x10_5/well with 2 mL of complete medium 24 h before transfection. Once at room temperature, the reagents were mixed in a sterile 1.5 mL tube as follows: first, 200 µL of serum-free DMEM were mixed with 3 µg

total plasmid DNA (Fig. 6). Second, the diluted DNA was mixed with 9 μL Polyethylenimine (25kD linear; Polysciences, cat# 23966-2). After incubation at room temperature for 15 min, the mixture was added to the cells and the plate was mixed 6-8 times. Cells were incubated at 37°C and 5% CO_2 for 48 h.

10. Selection of stable transfectants

Post-transfection, the cells were counted on a hemocytometer and subcultured at the desired density in 100 mm dishes. Selection of stable transfectants began 24 h after with addition of Geneticin (G418) at 0.8 mg/mL to the medium (Fig. 6). Medium was changed and supplemented with G418 three times a week for three weeks. Resistant clones were observed under a bright-field microscope and those exhibiting healthy morphologies were marked for further expansion. To isolate the individual clones, the dishes were washed twice with 5 mL of Dulbecco's PBS and a cloning disc previously soaked on trypsin for 15 min was placed directly on top of each selected clone. After incubation at 37°C for 5-6 min, the discs were transferred to a 24-well plate containing 0.5 mL medium and mixed 6-8 times. Medium was changed 2-3 times a week. As G418-resistant clones reached moderate confluences (50-70%), they were transferred to 6-well plates, where they were maintained with addition of fresh medium three times a week. A population of stable transfectants were maintained in 100 mm dishes with complete medium supplemented with 0.8 mg/mL G418.

11. Preparation of cell extracts for western blotting analysis

G418-resistant clones of stably transfected cells were trypsinized in 1 mL of trypsin and resuspended in 1 mL of complete medium. Resuspended cells were transferred to a 2 mL tube

and kept on ice, and 0.1 mL were used for cell counting. The resuspended cells were centrifuged for 1 min at 5,000 g. After removing the supernatant, 2 mL of PBS were added and the tubes were vortexed and centrifuged for an additional minute. The volume of cold lysis buffer (150 mM NaCl, 20 mM Tris pH 7.5-8.0, 1% Triton X-100, 2 mM EDTA, 1 mM PMSF) added to the pellet was proportional to the number of cells in the extracts. Samples with the highest cell count were resuspended in 300-400 μL of buffer, while 70-100 μL of buffer was used for samples with the lowest cell counts. The mixtures were passed through a 22 G ½ needle 8-10 times. After centrifugation at 21,130 g for 2 min, the supernatant (containing the cell extracts) was transferred to a new tube and stored at -20°C overnight.

12. Western blotting

Samples for western blotting by mixing 20 μL cell extracts with 5 μL 5x sample buffer (0.375 M Tris pH 6.8, 12% SDS, 60% glycerol, 0.6 M DTT, 0.06% bromophenol blue). The samples were heated for 3 min at 95 °C and 15-20 μL were loaded onto a 10% SDS-PAGE gel and ran at 160V for 65 min. The gel was subsequently incubated in 1X transfer buffer (25 mM Tris Base, 192 mM glycine, 10% MeOH) and transferred for 60 min at 100 V onto a nitrocellulose membrane, then blocked in a solution of 5% non-fat milk in TBST (Tris-buffered saline, 0.1% Tween 20). Incubation of the nitrocellulose membrane was done with rabbit RP_3-CPO (Triple Point Biologics; 1:5000 dilution) as the primary antibody and HRP conjugated anti-Rabbit IgG (Cell Signaling Technology; 1:2000 dilution) the secondary antibody. Nitrocellulose was treated with Lumi-GLO chemiluminescent reagent (Cell Signaling Technology) for 1 min and exposed to an X-ray film in a cassette, followed by manual development.

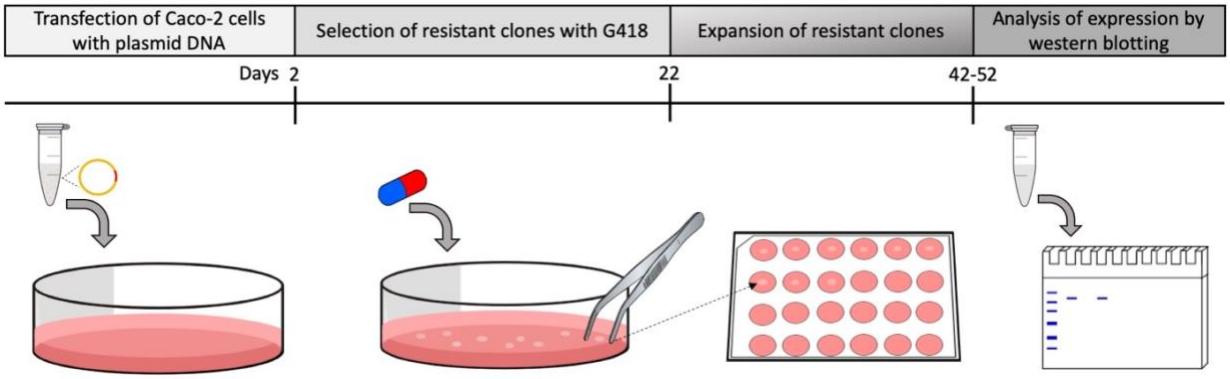

Figure 6. Illustration of transfection and selection protocols. Plasmid DNA containing WT or E310Q CPO (in red) is added to the cell cultures (pictured: 100mm culture dish). Cell cultures are treated with geneticin (G418) 2-3 times a week. A population of resistant clones is transferred to 24-well plates and expanded. Cell extracts from clones that were successfully expanded are analyzed for CPO expression via western blotting.

13. Immunofluorescence

Coverslips were incubated with Poly-D-lysine in PBS (1 mg/mL) for 1 h, washed with sterile H_2O, and dried under UV light for 1h. Caco-2 cells from a pooled polyclonal culture containing hCPO WT transfectants were seeded at 1.5 $x10^5$ cells/well in 6-well plates containing poly-D-lysine coated coverslips. When cells reached 60-80% confluency, they were washed three times with PBS, fixed in 4% paraformaldehyde for 10 min at room temperature, and then washed three times for 5 min on the rocker. The cells were permeabilized in 0.1% Triton X-100 in PBS for 15 min, washed, and blocked with 5% bovine serum albumin in PBS for 35 min. Primary antibodies RP3-CPO (Triple Point Biologics, cat.# MAB41241-100; 1:1000 dilution) or human Apolipoprotein B (R&D Systems, 1:25 dilution) were diluted in 5% bovine serum albumin in PBS and added directly onto the coverslip and incubated for one hour. After washing three times with PBS, the cells were incubated with secondary antibodies anti-mouse Alexa Fluor 488 and/or anti-rabbit Alexa Fluor 555 (Cell Signaling Technology, 1:1000 dilutions). Coverslips were washed three times and then mounted with 12 µL of aqueous mounting medium

(PBS containing 1 mg/mL p-phenylenediamine hydrochloride and 1 μg/ml DAPI). To stain lipid droplets, 4 μL BODIPY (Invitrogen) was added to the coverslip during the penultimate washing step and incubated on the rocker for 13 min. The slides were allowed to dry in a cool, dark place overnight and were later stored at 4°C.

Images of thirteen fields of view containing at least one CPO-expressing cell were captured with a fluorescence microscope. To evaluate association, % values (LD:CPO or apoB:CPO) were normalized using the arcsin transformation and subsequently analyzed by a one-way ANOVA.

14. Carboxypeptidase Assay

Twenty microliters of HEK293T cell extracts were incubated with 100 μL 0.5mM Fa-EE carboxypeptidase substrate in a 96 well plate and their A340 was measured at 24°C every min for a total of 30 min.

CHAPTER 3

RESULTS

1. Analysis of CPO expression

In order to better understand the potential biological role of CPO, an in-depth analysis of data from several gene expression databases was performed. Investigating the levels of expression across organs and tissues can potentially shed some light on the biological function of this protein. Thus, a comprehensive collection of CPO expression data is an appropriate approach to accomplish this.

Protein expression is commonly similar to RNA expression which can be quantified using transcriptomic techniques (Wang et al., 2009). These transcriptomic techniques present a high-throughput, highly sensitive sequencing method that allows quantitative assessment of the expression of a gene in a given tissue or cell line (McGettigan, 2013). This is achieved by measuring the RNA transcript abundance found for a given protein-coding gene. For the purpose of the expression analysis in this study, RNA sequencing (RNA-Seq) and microarray data were compiled.

a. CPO expression in the small intestine

CPO exhibits its highest levels of expression in the small intestine, as found in RNA-Seq and microarray databases. The consensus expression found in proteinatlas.org, which includes data from HPA, GTEx and FANTOM 5, collectively denote the small intestine as the tissue with highest expression in humans (Fig. 7A-C). In addition, HPA indicates that CPO expression is tissue enriched, as it is at least 4x greater than in any other tissue. More specifically, microarray

data from the Affymetrix Human Genome U133 Plus 2.0 Array platform revealed that CPO is overwhelmingly expressed in the ileum (Fig. 8B). The affymetrix microarray data was found to closely match the data found in the mRNASeq_HUMAN_GL-1 platform, both referring to location and relative levels of expression (Fig. 8A). Notably, in the mRNASeq platform the aforementioned ileal epithelium is referred to as the terminal ileal epithelium. High levels of expression were detected in the ileal epithelium, also known as mucosa. RNA-Seq data from Bgee further suggests that within the ileal mucosa, lymphoid structures known as Peyer's patches, are the anatomical entity with the highest relative CPO expression (Fig. 7D).

b. CPO expression in other tissues

In addition to the small intestine, CPO exhibits varying levels of expression in other tissues including male reproductive tissues, brain, adrenal, and thyroid (Figs. 7,8). According to the HPA and FANTOM 5 databases, the epididymis is the tissue with the second highest CPO expression (Fig. 7A,C). This claim is supported by the Affymetrix Human Genome U133 Plus 2.0 Array platform, which names two regions with medium expression levels, namely the epididymal cauda and corpus (Fig. 8B). Although the mRNASeq platform fails to include the epididymis, it mentions another male reproductive structure, the testis, with medium-low expression (Fig. 8A).

The brain exhibits widely spread CPO expression with varying levels of expression. For example, the Affymetrix Human Genome U133 Plus 2.0 Array platform indicates that medium expression levels are found in a variety of neural structures. Within the medium expression level range, the cingulate cortex neuron had the highest expression, followed by the hippocampus pyramidal neuron, basal ganglia, entorhinal cortex large stellate neuron, and the hypothalamus

(Fig. 8B). While the Affymetrix U133 platform suggests medium expression in multiple brain tissues, RNA-Seq data suggests low to background level expression in most of those tissues (Fig. 8A).

RNA-Seq data from Bgee and mRNASeq_HUMAN_GL-1 report that the caudate nucleus and caudate-putamen complex, respectively, as the sites with highest expression in the brain; nonetheless, the expression levels in both structures appears to be modest (Figs. 7D, 8A). Further, the mRNASeq_HUMAN_GL-1 platform reports medium expression in the hippocampus, which contrasts with the background level expression detected by RNA-Seq from GTEx (Fig. 8A, 7B, respectively). GTEx and FANTOM 5 do not report regional specificity nor specific distribution of CPO in the human brain. Nonetheless, both databases report low expression in the ganglia while GTEx also reports expression in the pituitary gland (Fig. 7A,C). Taken together, the data points to an overall low presence in the brain.

Low levels of CPO expression are seen in adrenal tissues. The adrenal gland cortex presents a moderate expression score, as determined by RNA-Seq databases from Bgee (Fig. 7D). Expression in the adrenal gland is also reported in the mRNASeq_HUMAN_GL-1, Affymetrix Human Genome U133 Plus 2.0 Array, and GTEx platforms (Fig. 8A,B; Fig. 7A). While overall expression in the adrenal gland is moderate, it appears to be slightly greater in the fetal stage than in adulthood (Fig. 8A,B). Nonetheless, it is important to note that none of the aforementioned platforms report expression in both fetal and adult, thus making it challenging to establish a fair comparison across platforms. Additionally, low expression is detected in the thyroid gland, as reported in and GTEx, HPA, Bgee, and Affymetrix Human Genome U133 Plus 2.0 Array, (Fig.7 A, B, D; Fig. 8B).

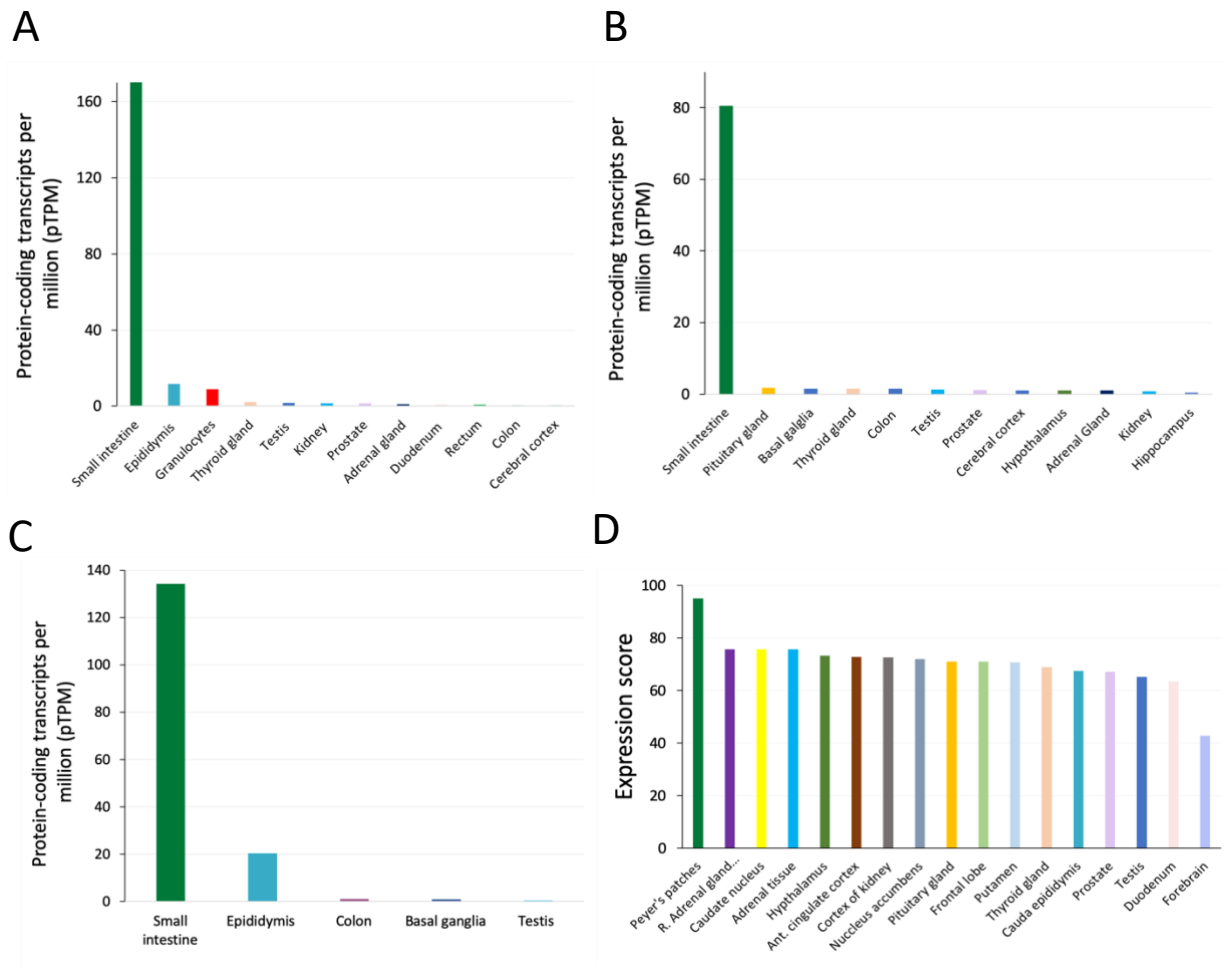

Figure 7. CPO is largely expressed in the small intestine and has low expression in a number of other tissues. CPO expression levels were determined by RNA-Seq data from HPA (A), GTEx (B), FANTOM 5 (C) datasets and both RNA-Seq/Affymetrix from the BGEE website (D). The small intestine is collectively reported as the tissue with highest CPO expression (A-D), and expression in other tissues is considerably lower than in the small intestine (A,B,C). Expression scores from Bgee have been normalized (0-100) from expression call values across genes and conditions within the species, with scores below ~80 arguably considered as moderate expression (D). Modified from http:// proteinatlas.org and http://bgee.org.

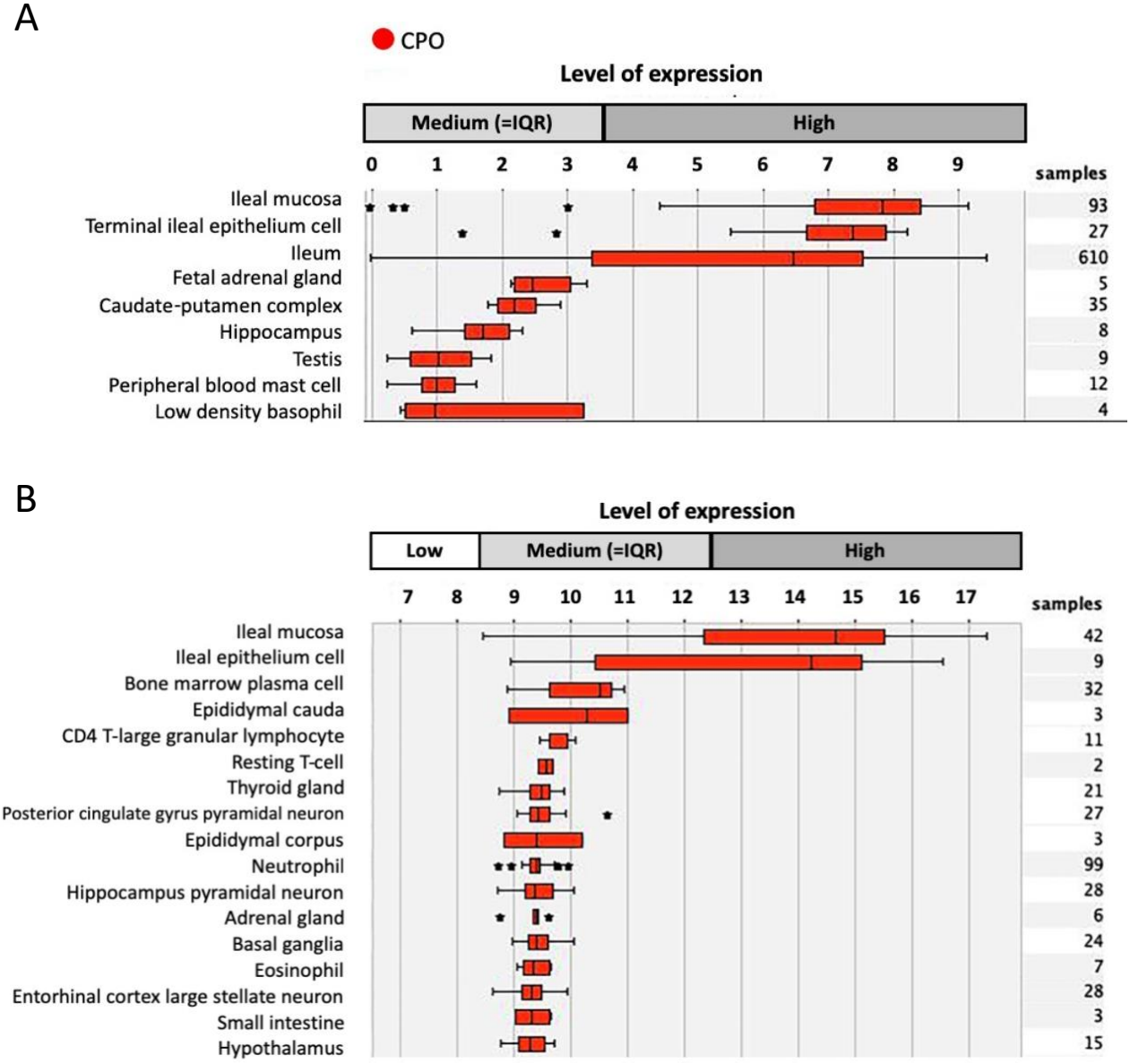

Figure 8. CPO is expressed in blood cells, male reproductive, brain, and adrenal tissues, with highest expression in the ileum mucosa and ileal epithelial cells. Gene expression is reported as absolute levels in log_2 scale, with low, medium, and high ranges specific to the mRNASeq_HUMAN_GL-1 platform (A) and the Affymetrix Human Genome U133 Plus 2.0 Array platform (B). In the affymetrix microarray platform (B), the mean value of the ileal mucosa "13.92" (shown as log_2) equals "35,075" in linear scale, which compares to "7.25" (shown as log_2) and "231.77" in linear scale according to the mRNASeq platform (A). Data is modified from Genevestigator.

c. CPO expression in the Blood

The blood was found to be the third highest tissue displaying significant CPO expression. mRNA expression data from the HPA database revealed enrichment of CPO expression in granulocytes, specifically the basophils (Fig. 9). Similarly, the mRNASeq_HUMAN_GL-1 platform also reports medium levels of CPO expression in low density basophils (Fig. 8A). In contrast, the Affymetrix Human Genome U133 Plus 2.0 Array platform reports a wider variety of blood cells with CPO expression ranging in the medium levels, with bone marrow plasma cells at the highest, then CD4 t-large granular lymphocyte, followed by T-cells, neutrophils, and eosinophils (Fig. 8B). A small, almost negligible level of expression in eosinophils is also found in the HPA database (Fig. 9).

Interestingly, granulocytes are white blood cells, known to play a crucial role in immunity. More specifically, basophils are involved in inflammatory responses such as allergic reactions (Wedemeyer et al., 2000). Thus, CPO expression in the basophils alludes to a potential role of CPO in immunity. Moreover, the expression in Peyer's patches, known to play a role in immunity by creation of B-lymphocytes also alludes to a link between CPO and immunity.

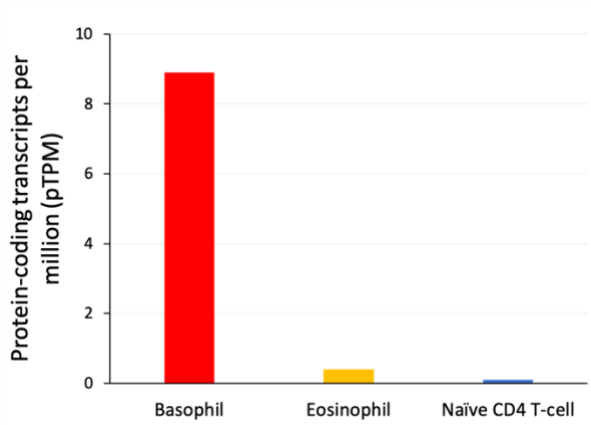

Figure 9. CPO expression in blood cells. Out of 29 total cell types analyzed by RNA-Seq, only three showed CPO expression, with basophils exhibiting the highest levels of expression at 8.9 pTPM. Modified from http://proteinatlas.org.

d. CPO expression in cell lines

While CPO is expressed across a variety of cell lines, its highest levels of expression are detected in cancer derived cell lines. RNA-seq data from the HPA database reports four cell lines where CPO is moderately expressed, and three of which are human cancer cell lines (Fig. 10A). These cancer cell lines are NB-4, from the bone marrow, ASC diff, from sarcoma, BEWO, from the brain, and the non-cancerous telomerase immortalized cell line hTCEpi, originating from the cornea. The mRNASeq_HUMAN_GL-1 platform cites four different cell lines with medium expression, with HD-iPS4-HTT at the highest followed by NCl-446, SNU-349, and WA09 (Fig. 10B). HD-iPS4-HTT constitutes induced pluripotent stem cells from Huntington's disease patient, while WA09 is a human embryonic stem cell line. The remaining two, NCl-446 and SNU-349 are cancer cell lines, from the lung and renal carcinoma respectively. A lung cancer metastatic cell line is reported to have high levels of CPO expression, according to the Affymetrix Human Genome U133 Plus 2.0 Array platform (Fig. 10C). Additionally, medium CPO expression levels are reported in embryonic stem cell WA07, mammary epithelial cell line 76N-RHOA, and bone marrow metastatic cancer cell line Rh30 (Fig. 10C).

Because CPO has prominent expression in the small intestine and basophils, we sought to investigate this expression in cell culture models, namely Caco-2, which has an intestinal phenotype, and a human basophil cell line, KU812. Both cell lines presented low to no expression according to mRNASeq_HUMAN_GL-1 and Affymetrix Human Genome U133 Plus 2.0 Array platforms. Overall, Caco-2 has a more consistent record of expression, with medium-low expression reported by the Affymetrix Human Genome U133 Plus 2.0 Array platform and low expression reported by the mRNASeq_HUMAN_GL-1 platform (Fig. 10B,C, respectively).

In turn, KU812 has medium CPO expression according to the affymetrix microarray platform, whereas the mRNA-seq platform notes no expression (Fig. 10C; 10B, respectively).

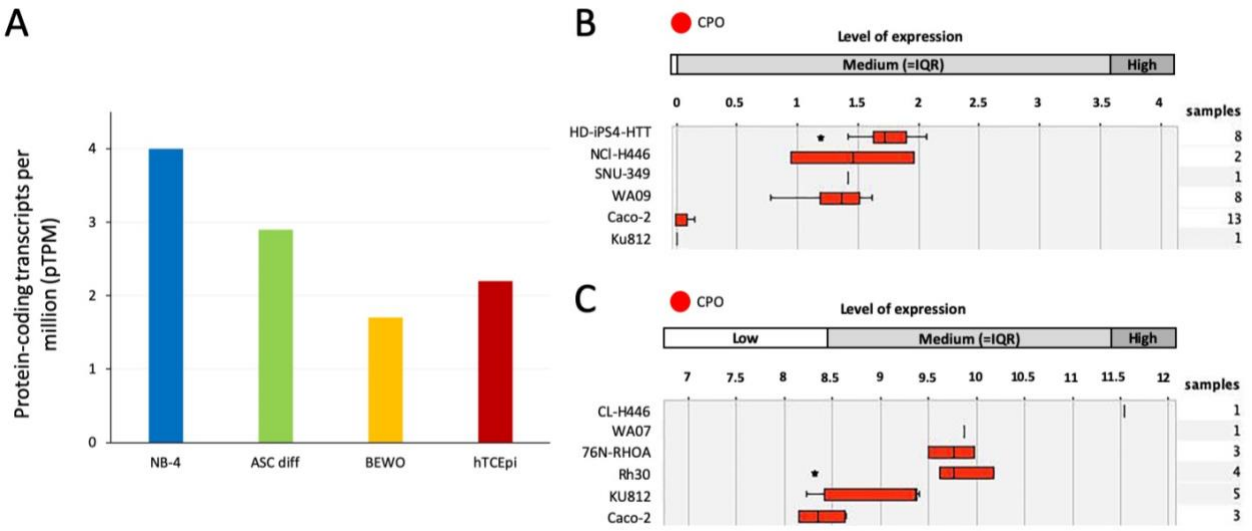

Figure 10. A diverse number of cell lines, mostly cancer derived, have moderate levels of CPO expression. RNA-seq data from HPA (A) shows three cancer cell lines with important CPO expression (in pTPM), NB-4, ASC diff and BEWO, while the mRNASeq_HUMAN_GL-1 platform (B) includes two, NCI-H446 and WA09 with medium expression (in absolute levels in log_2 scale). The Affymetrix Human Genome U133 Plus 2.0 Array (C) is the only platform to report a cell line (CL-H446) with high CPO expression. Data modified from Genevestigator and http://proteinatlas.org.

e. CPO expression and associated pathology

A number of pathological disorders characterized by chronic inflammation of the digestive tract have shown abnormal CPO expression levels. These pathological disorders are considered subtypes of the inflammatory bowel disease (IBD) and include Crohn's disease and ulcerative colitis (Baumgart and Sandborn, 2012). Since they share many symptoms, these two pathologies are mainly differentiated by their localizations. Crohn's disease can develop anywhere along the digestive tract, while ulcerative colitis is confined to the colon (Baumgart and Sandborn, 2012; Head and Jurenka, 2003). Clinical data retrieved from Genevestigator

mRNA-Seq (mRNASeq_HUMAN_GL-1) and affymetrix microarray (Human Genome U133 Plus 2.0 Array) platforms revealed two main sites with high CPO expression where such pathological conditions have been reported.

The first site was identified as the ileal mucosa and was characterized by the highest CPO expression (Fig. 11A, 12A). A study by Arijs et al. (2009; Genevestigator ID (G-ID): HS-00598) assessed expression in healthy patients and Crohn's disease patients who were receiving Infliximab treatment. The results of the study suggest a pattern of decreased expression in Crohn's disease patients (Fig. 11B).

Similarly, the ileal epithelium, which was ranked second highest in expression in the affymetrix microarray platform, reported high-medium expression in all patients (n = 6) diagnosed with Crohn's disease (Funke et al., 2009; G-ID: HS-01050; Fig. 11C). These levels of expression in diseased patients appear to be less than those detected in healthy patients, coinciding with the presumed trend observed by Arijs et al. (2009; G-ID: HS-00598) in the ileal mucosa. Furthermore, a large study conducted by Howell et al. (2018; G-ID: HS-03228) reported expression ranging from high to medium in 17 patients with IBD (n = 17) (Fig. 12B). Upon analysis of the individual cases, it was found that all the patients with ulcerative colitis (n = 7) exhibited higher expression levels than those with Crohn's disease (n = 10; Fig. 12C). Likewise, most Crohn's patients exhibited decreased expression versus healthy individuals (n = 6).

Additionally, ulcerative colitis appears to modulate CPO expression, as seen by a differential modulation in ileal tissue. For example, one study (Haberman et al., 2014; G-ID: HS-01562) reported CPO expression was modestly downregulated in ileal tissue of young patients (aged 0-10 years old) with ulcerative colitis (Fold Change (FC): -1.63; $p < 0.001$) (Fig. 12C). In the same study, older diseased patients (10-17 years old) showed increased expression than

younger patients (FC: +1.61; $p < 0.001$), indicating that CPO expression specifically in the ileum of diseased patients may be age dependent.

In contrast, CPO is more drastically regulated in Crohn's disease. One study by Peck et al. (2015; G-ID: HS-01505) found a very large upregulation of the CPO gene in colon tissue of patients with bowel stricture (B2) and penetrating (B3) Crohn's disease (FC: 12.15 and 4.28, respectively; $p < 0.001$) (Fig. 12C). Conversely, CPO was significantly downregulated in the ileum of diseased pediatric patients, as reported in study G-ID: HS-01562 (Haberman et al., 2014). More specifically, the downregulation was greater in patients aged 10-17 (FC: -4.45; $p < 0.001$) than in patients aged 0-10 (FC: -3.57; $p < 0.001$).

Although CPO is modulated in both IBD conditions, ulcerative colitis and Crohn's disease, the data indicates a much more evident relationship between CPO and Crohn's disease. When comparing CPO expression in both pathologies, a study (G-ID: HS-01562; Haberman et al., 2014) found that CPO is significantly more downregulated in ileal tissue in Crohn's disease versus ulcerative colitis (Fig. 12C). In fact, this difference is greater in Crohn's patients who exhibit deep ulcers (FC: -4.51; $p < 0.001$) than in those with a clinically affected colon (FC: -2.13; $p < 0.001$) (Fig. 12C). Downregulation in Crohn's disease was more moderate in younger patients (FC: -1.38; $p = 0.033$; Fig. 12C).

In addition, moderate CPO expression has been reported in several pathologies, including cancer. In an analysis of expression across 72 paired tissues, the GENT2 database showed that CPO was significantly overexpressed in cancerous small intestinal and blood tissue ($p = 0.026$ and $p < 0.001$, respectively) (Fig. 13). Colon carcinoma and lymphoma, although based on a very small sample size ($n = 2$), exhibit high CPO expression (Fig. 11A). Other pathologies with medium expression levels were the monoclonal gammopathy of undetermined

significance (MGUS) and multiple myeloma. Furthermore, a study on adrenal cortical carcinoma (G-ID: HS-01616; Zheng et al., 2016) found a moderate downregulation in patients on stage IV compared to those on stage I (FC: -2.2; $p < 0.001$). (Fig. 12C).

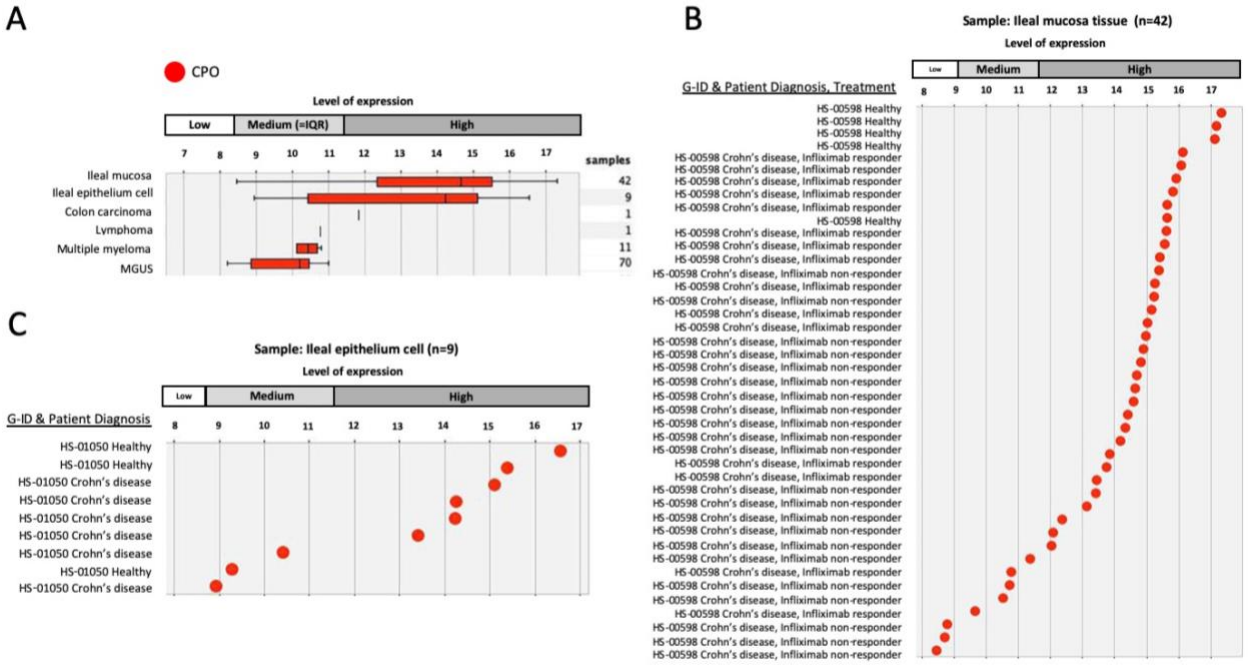

Figure 11. CPO expression data from the Affymetrix Human Genome U133 Plus 2.0 Array platform. The platform identifies high levels of CPO expression the ileal mucosa and ileal epithelium cells, with a trend indicating decreased expression in Crohn's disease patients. Microarray data from Genevestigator includes CPO expression levels across 1,066 anatomy/cancer entries, and identifies the ileum as the primary site of expression (A). Two studies, G-IDs: HS-01050 and HS-00598, found decreased expression in the ileal epithelium cells (B) and ileal mucosa (C) of patients with Crohn's disease. The experimental treatment in study HS-00598 consisted of a single Infliximab dose (5mg/kg body weight) (C). Values shown are absolute expression levels in log_2 scale. Modified from Genevestigator.

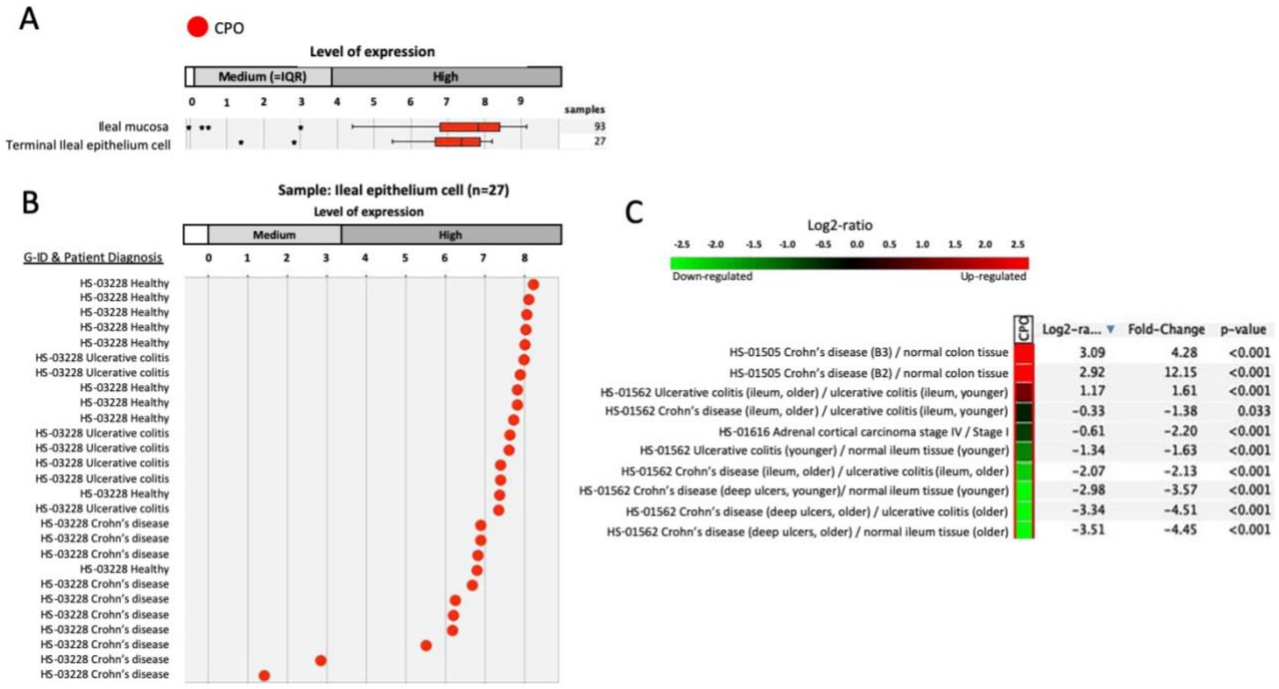

Figure 12. CPO expression data from the mRNASeq_HUMAN_GL-1 platform. This RNA-Seq platform reports CPO expression levels across 624 anatomy/cancer entries, and identifies the ileum as the primary site of expression (A). A study (G-ID: HS-03228) reported decreased levels of expression in IBD patients, with ulcerative colitis patients exhibiting higher expression than Crohn's (B). Values shown are absolute expression levels (TPM) in log_2 scale. C. CPO expression is differentially modulated in IBD patients, with older patients (aged 10-17) presenting stronger regulations than younger patients (aged 0-10). CPO expression is regulated more strongly in Crohn's disease patients than in ulcerative colitis patients. Log_2-ratio represents the difference between the average log_2 expression for experimental and control samples. Modified from Genevestigator.

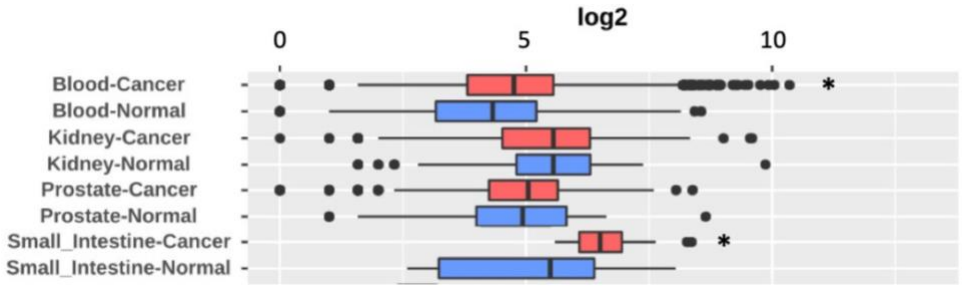

Figure 13. CPO is overexpressed in cancerous small intestinal and blood tissues. In an analysis of expression across 72 healthy and cancerous "paired" tissues, the GENT2 database reported a significantly increased log_2 fold-change in small intestine ($p = 0.026$) and blood tissue ($p < 0.001$). Modified from http://gent2.appex.kr/gent2.

2. Creation of the hCPO E310Q mutant

Unlike other members of the M14 family, CPO is expressed as a constitutively active enzyme, as it does not need to undergo proteolytic cleavage to be activated (Garcia-Guerrero et al., 2018). Thus, to control for enzyme activity and protein expression, it is useful to employ an inactive CPO mutant as a means of comparison against the fully active (WT) CPO. This may be achieved through mutagenesis of a key catalytic residue, namely E310, which will effectively suppress the catalytic ability of CPO. An hCPO E310Q mutant was created with the ultimate goal of comparing the association of CPO with lipid droplets (LDs) and apoB48 in WT and E310Q CPO expressing Caco-2 cells. Differences in association would be considered in light of CPO activity (in WT) or lack thereof (in E310Q).

a. Why E310Q?

The selected mutation was on a key catalytic residue, E310. To assess any potential steric strain in the hCPO E310Q mutant, a 3-D model of the molecule was visualized using PyMOL (Fig. 14A). A close-up on the catalytic residue area revealed minimal change in structure due to the similarity of the two amino acids, glutamic acid (E) and glutamine (Q; Fig. 14B). Essentially, these two residues only differ in the terminal side chain, where E contains a -COOH group and Q has a $-CONH_2$. Thus, modeling of the point mutation E310Q in hCPO verified that the steric hindrance produced by the mutagenesis was negligible.

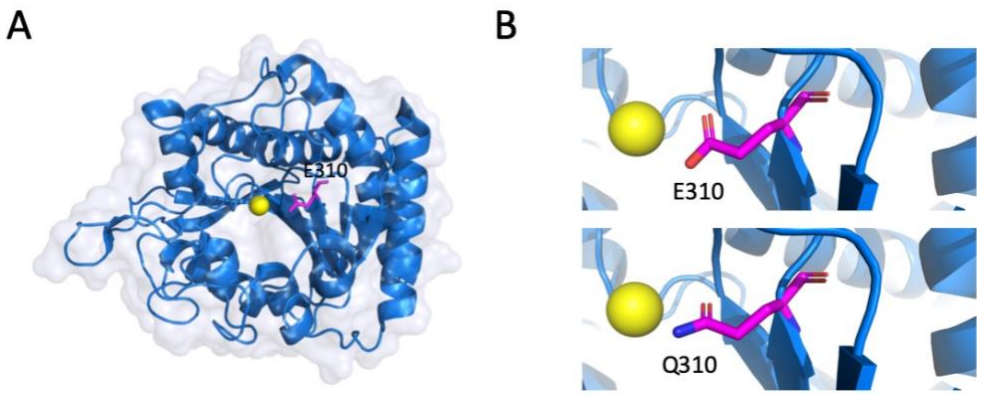

Figure 14. The E310Q mutation substitutes a key catalytic residue in hCPO without contributing steric hindrance. A. Three-dimensional structure of hCPO in its monomeric form (PDB: 5RMV). E310, a key catalytic residue, is shown in magenta. Zinc cofactor is depicted as a yellow sphere. B. A close-up of the 3-D hCPO model showing a key catalytic residue E310 (top) and its mutated form Q310 (bottom). While the E310Q mutation does not add any steric hindrance, it effectively inhibits hCPO catalytic activity.

b. Site-directed mutagenesis

A PCR-based mutagenesis procedure was used to perform site-directed mutagenesis. Amplification of hCPO DNA using mutagenic primers was visualized via gel electrophoresis (Fig. 15). Bands of similar sizes were seen in +DMSO and -DMSO PCR products, indicating that addition of DMSO did not enhance amplification.

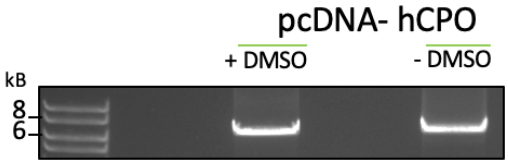

Figure 15. HCPO DNA was successfully amplified and resolved using gel electrophoresis. PCR products with and without DMSO exhibited similar bands.

Prior to performing ethanol precipitation, the PCR products (± DMSO) were mixed with each other resulting in two products that contained about half of the +DMSO PCR reaction and half of the -DMSO. Both 50/50 products were digested by addition of 2.5% Dpn1. Ethanol

precipitation was performed only on one of the samples to purify it by concentrating the DNA and remove unwanted substances (Soares et al., 2012). Because of the complexity of the technique and the risk of losing DNA, it was decided that it would be appropriate to perform this only on one sample. Subsequent assessment of bacterial growth confirmed that the ethanol precipitation step was unnecessary, as colonies containing digested DNA were larger and more abundant than those containing ethanol purified DNA (n = 16 and n = 14, respectively).

After transformation, plasmid from *E.coli* colonies was purified and visualized on a 1% agarose gel (Fig. 16A). DNA quantification showed that only colonies 2 and 3 had optimal amounts of plasmid required for sequencing.

Sanger sequencing was used to confirm the presence of the novel mutation (C→G) in both samples. Presence of the E310Q mutation (*) was confirmed using sequence alignment against the hCPO WT (BLAST). The best mutant was selected on the basis of presence of the mutation and overall DNA quality, as determined by the base calls in the chromatogram (Fig. 16B).

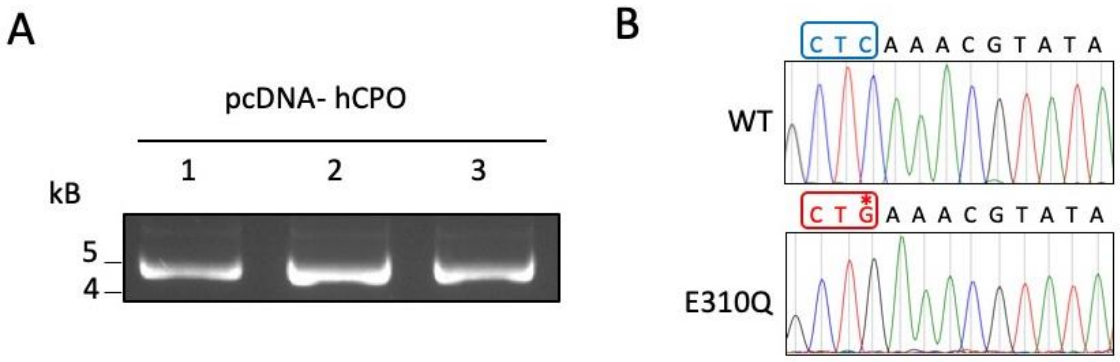

Figure 16. Confirmation of the mutagenesis in bacterial cells. A. After transformation, pcDNA from three E. coli colonies (1-3) was purified and visualized on a 1% agarose gel. B. Presence of the E310Q mutation (*) was confirmed using sequence alignment against the hCPO WT (BLAST). The missense mutation results in the coding of glutamine (Q) instead of glutamic acid (E). The best mutant was selected on the basis of presence of the mutation and overall DNA quality, as determined by the base calls in the chromatogram. In blue, reverse compliment of codon coding for E; in red, reverse compliment of codon coding for Q.

c. Testing the expression and catalytic activity of the hCPO E310Q mutant

i. Western blotting

To determine expression of the hCPO E310Q mutant, HEK293T cells were transiently transfected with pcDNA-hCPO E310Q. HEK293T cells were also transfected with empty vector pcDNA3.1(-) and pcDNA-hCPO WT, which were used as negative and positive controls, respectively. Each transfection was performed in duplicate. The cell extracts were subsequently collected and analyzed via western blotting. Analysis via western blotting revealed that the E310Q mutant had similar hCPO expression as the WT hCPO (Fig. 17A).

ii. Carboxypeptidase Assay

Two carboxypeptidase (CP) assays were run independently. Successful proteolytic cleavage of the substrate by CPO can be measured by a decrease in absorbance at 340 nm (Fig. 17B). The average activity of the E310Q mutant (0.051 milliUnits/min) was significantly smaller than that of the WT (1.59 milliUnits/min) as determined by Student's t-test ($p < 0.005$). In contrast, E310Q and empty vector had similar activity, suggesting that hCPO E310Q lacked catalytic activity.

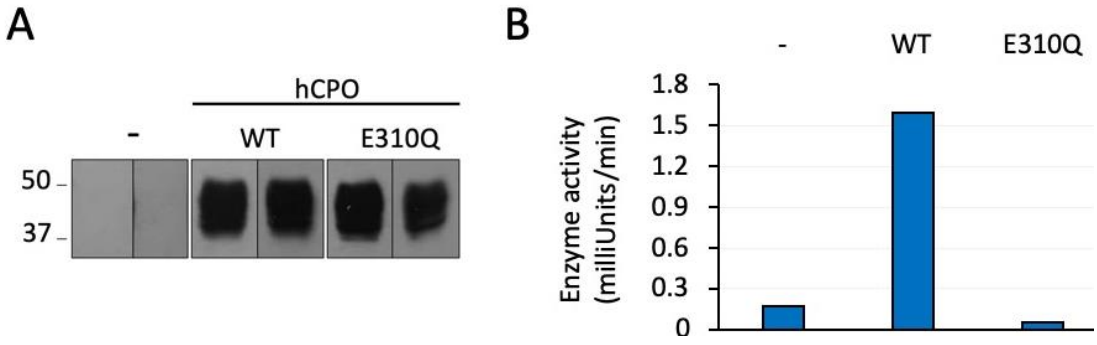

Figure 17. Analysis of hCPO E310Q expression and catalytic activity in mammalian cells. A. Extracts from HEK293T cells transiently transfected with empty vector (-), pcDNA-hCPO WT and pcDNA-hCPO E310Q were analyzed via western blotting. The E310Q mutant had similar hCPO expression as the WT. B. To measure enzymatic activity, HEK293T cell extracts were

incubated with 0.5mM Fa-EE substrate and their A340 was measured at 24°C every min for a total of 30 min. Values shown are averages from two separate carboxypeptidase (CP) assays.

3. Cell transfection and selection with hCPO WT and hCPO E310Q

Caco-2 cells are considered a good model to study intestinal physiology, as they mimic molecular and physiological characteristics of the small intestinal epithelium (Natoli et al., 2012). Although the Caco-2 cell line is an optimal choice for this study, its endogenous expression of CPO is low and thus insufficient to accurately assess enzyme activity (Fig. 10B,C). Therefore, to be able to study the activity of CPO in Caco-2 cells, it was necessary to produce a hCPO stably expressing cell line.

a. Control experiments

Prior to performing the transfection with pcDNA-hCPO WT and pcDNA-hCPO E310Q, several control and optimization experiments were conducted. First, to control for transfection efficiency, Caco-2 cells were transiently and stably transfected with a fluorescent plasmid. Also, the selection protocol was optimized by experimentally determining optimal G418 concentration. Finally, an attempt to optimize the transfection protocol included the testing of two transfection reagents, PEI and FuGENE6.

i. Determining effective G418 concentration

In order to serve as a good selective agent, G148 must be used at an optimal concentration to induce cell lethality before cells reach confluency. Thus, G418 concentration as well as seeding cell density are key factors to consider when optimizing the selection protocol. Thus, an experiment was designed where Caco-2 cells were seeded at 43,000 cells/ml in a 6-well

plate. G418 was added at three different concentrations: 0.6 mg/mL, 1mg/mL, and 4mg/mL. One well that received no G418 supplementation was used as a control. Medium was changed and supplemented with fresh G418 every 48 h. Cell toxicity was visually assessed at 48 h, 72 h, and 92 h following the first G418 addition.

Visual assessment of cell toxicity revealed that the minimum lethal concentration was between 0.6-1mg/mL. Approximately 48 h after the first G418 addition, cell toxicity was similar at 0.6 mg/mL and 1mg/mL of G418, whereas it was much more severe at 4 mg/mL, as seen by a considerably larger number of floating cells. At 92 h, the confluence at 1 mg/mL was half of that observed at 0.6 mg/mL. In contrast, the confluency at 4 mg/mL G418 was extremely low (10 %) with large cohorts of dead cells as compared to the control. These results suggest that a concentration of 4 mg/mL is not optimal, since large amounts of dead cells will correlate to secretion of undesired substances that may in turn affect the resistant cells leading to their death. Treatment with 0.6 mg/mL and 1 mg/mL G418 successfully killed non-resistant cells within 92 h without massive cell death. Thus, optimal lethal concentration required for optimal selection of resistant clones lies in the 0.6-1mg/mL range.

ii. Transient transfection using fluorescent plasmids

To assess the efficiency of transfecting Caco-2 cells using the PEI transfection reagent, Caco-2 cells were transiently transfected with various fluorescently tagged plasmids. Three plasmids with fluorescent tags were used for transfection: pEGFP-N_2, pcDNA-GFP-CAAX, and pcDNA-mCherry-CAAX. The cells were observed under a fluorescent microscope 48 h post transfection. Quantification of transfection efficiency was calculated by dividing the number of fluorescent cells per field of view by that area (0.35 mm_2) and multiplying by surface area of the

6-well plate (9.6 cm_2). The total number of transfected cells was reported as the average of three fields of view. While all three transfections were successful, quantification of the observed fluorescence revealed varying transfection efficiencies (Fig. 18). The pEGFP-N_2 had the highest transfection efficiency at ≅ 21,000/250,000 transfected cells per well, while pcDNA-mCherry-CAAX had the lowest with ≅ 15,500/250,000. This suggests a transfection efficiency of 5-10% for this system.

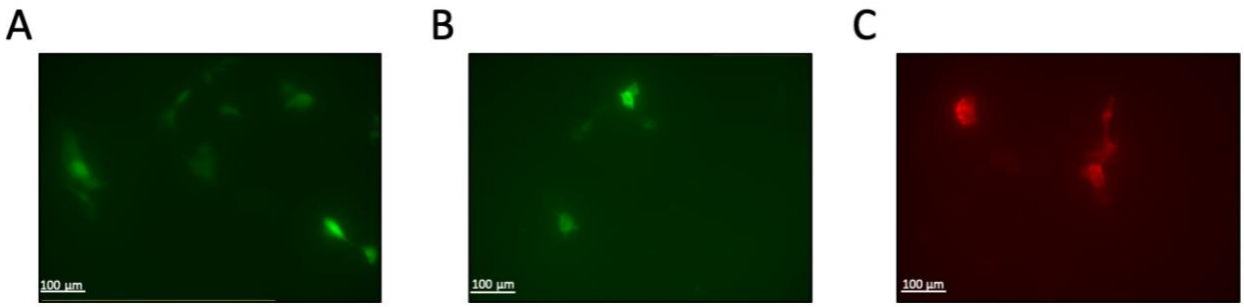

Figure 18. Caco-2 cells transiently transfected with fluorescently tagged plasmids. Plasmids pEGFP-N_2 (A), pcDNA-GFP-CAAX (B), and pcDNA-mCherry-CAAX (C) were used as transfection markers to assess the transfection efficiency in Caco-2 cells. Fluorescent cells in three fields of view were manually counted and the average was used to estimate the transfection efficiency for each of the plasmids. Images show fluorescent cells, n = 8 in (A), n = 4 in (B), and n = 3 in (C). With a total cell number of about 250,000 cells per 9.6cm_2 , 21,000 cells were transfected with pEGFP-N_2 ,20,800 with pcDNA-GFP-CAAX, and 15,500 with pcDNA-mCherry-CAAX.

iii. Stable transfection using a fluorescent plasmid

Caco-2 cells were seeded in a 6-well plate at a density of 3 x10_5/mL and subsequently transfected with 3 µg total pEGFP-N_2 pcDNA 24 h after (see Chapter 2, section 9). 48 h post-transfection, the cells were subcultured at ≃ 54,500 cells/mL in 100 mm dishes with fresh medium. After selection with G418 (see above), cells were observed under a fluorescent microscope and three resistant clones presenting uniform, bright fluorescence were selected by marking their location on the plate. Two isolated resistant clones were grown to ≃70 %

confluency in a 24-well plate at 37°C and 5% CO_2, and expanded into 60 mm dishes. Fluorescence was observed three weeks post-selection, suggesting that the transfection and selection protocols were optimal for producing stable transfectants (Fig. 19).

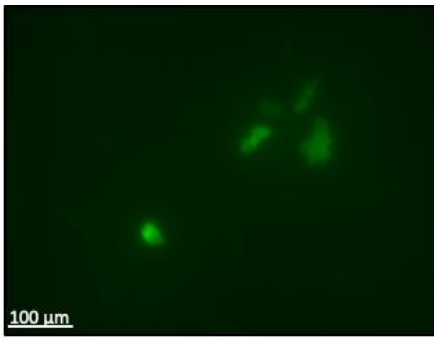

Figure 19. Caco-2 cells stably transfected with a fluorescently tagged plasmid, pEGFP-N_2. Here pictured are four fluorescent cells in a field of view. Two stably transfected clones were isolated and visualized with a fluorescent microscope to confirm fluorescence three weeks post-selection with G418.

iv. Semi-transient transfection with pcDNA-hCPO WT and hCPO E310Q

Caco-2 cells were transfected with 3 µg total pcDNA-hCPO WT and pcDNA-hCPO E310Q (see Chapter 2, section 9). 48 h post-transfection, the cells were subcultured at ≃ 10,600 cells/mL for WT and 19,500 cells/mL for E310Q in 100 mm dishes. The cells were grown to confluency at 37°C and 5% CO_2 for seven days, and media was changed every other day. After trypsinization and resuspension, a small volume of the cells was set aside to be counted. The remainder of the cells (2,012,800 cells WT and 3,180,800 cells E310Q) were harvested and lysed in lysis buffer (see Chapter 2, section 11). Cell extracts from transient hCPO WT and E310Q transfectants were analyzed via western botting with rabbit RP_3-CPO as the primary antibody and HRP conjugated anti-Rabbit IgG as the secondary antibody (see Chapter 2, section 12). After a 55 min exposure, both transfectants exhibited bands at the expected product size for CPO

(about 42.5 kDa; Fig. 20). Nonetheless, the presence of a shadow in an empty lane adjacent to the mutant sample "Caco-2 hCPO E310Q" argues to consider the possibility of spillover.

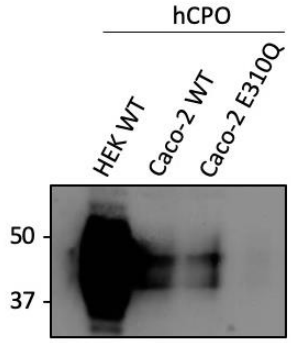

Figure 20. Caco-2 cells transiently transfected with pcDNA-hCPO WT and pcDNA-hCPO E310Q. Transient hCPO expression was assessed seven days post-transfection. Cell extracts of Caco-2 hCPO WT and E310Q transfectants were harvested, lysed, and analyzed via western blotting. Cell extracts from HEK293T cells transfected with pcDNA-hCPO WT were used as a positive control.

v. Optimization of transfection with PEI and FuGENE6

Since the semi-transient transfection results did not confidently indicate a successful transfection, it was decided to revisit the transfection protocol to investigate the efficiency of the transfection reagent, PEI. Given that transfection efficiency is largely determined by the ratio of the transfection reagent to the DNA, I attempted to optimize the transfection protocol by testing the transfection efficiency using multiple PEI:DNA ratios. To do this, Caco-2 cells were seeded at $3x10^5$ cells/well in 6-well plates and transfected with WT hCPO pcDNA at varying PEI:DNA ratios (µL PEI, and µg for DNA, respectively): 1:3, 3:3, 9:3, 15:3, 20:3, 1:1, 3:1, 9:1, 15:1, 20:1. Transfection with fluorescent plasmid pEGFP-N_2 at a 9:3 ratio was used as a control.

Cell extracts were harvested, lysed, and analyzed via western blotting (see Chapter 2, sections 11,12). None of the cell extracts showed expression (not shown). While these results may be suggestive of some form of human error in the experiment, the control sample showed abundant fluorescence, thus ruling out that possibility. An additional attempt to optimize the

transfection efficiency with CPO involved the use of a high efficiency transfection reagent, FuGENE 6. The transfection was done according to the manufacturer's protocol (Roche) and pEGFP-N_2 was used as a control. Assessment of the control sample post-transfection with FuGENE6 revealed minimal fluorescence, and it was concluded that this fluorescence was less than that observed in the previous optimization experiment with PEI. Thus, the low levels of fluorescence in the control sample suggested low transfection efficiency with FuGENE6. In contrast, high levels of fluorescence demonstrated that transfection with pEGFP-N_2 using PEI was successful; however, low hCPO expression, possibly below the western blot detection limits, prevented confirmation of transfection. Thus, failure to detect hCPO expression via western blotting did not necessarily indicate failure of the transfection, but rather suggested low hCPO expression levels.

b. Isolation of stable transfectants expressing hCPO WT and hCPO E310Q

Caco-2 cells were seeded at 0.8 x10^5 cells/well in a 6-well plate and subsequently transfected with plasmids expressing hCPO WT, hCPO E310Q, and pcDNA3.1(-) empty vector (see Chapter 2, section 9). 48 h post-transfection, cells were passaged at 1:2 (vol/vol) and seeded into 100 mm dishes. Twelve hCPO WT resistant clones, thirteen hCPO E310Q, and nine empty vector resistant clones were grown in 24-well plates using the aforementioned protocol (see Chapter 2, section 10). Cell extracts were centrifuged at 5,000 g for 1 min and the pellet was resuspended in 1x sample buffer. Volumes of sample buffer added ranged from 100 to 300 μL and were directly proportional to the size of the pellet. Cell extracts from resistant clones (four hCPO WT, eight hCPO E310Q, and three empty vector) were analyzed by western blotting (see Chapter 2, section 12). Western blotting analysis of four hCPO WT and eight E310Q resistant

clones showed that all lacked CPO expression (not shown). A closer inspection of the western blotting results revealed that some resistant clones presented a single band of faint intensity at around 65 kDa, which corresponded to bovine serum albumin (BSA) protein.

Resistant clones with BSA expression indicated protein presence in the sample, which argued for potential presence of hCPO. If hCPO expression was presumably too low to be detectable via western blotting, then further expanding the clones would increase the number of hCPO expressing cells thus allowing detection of expression. Thus, resistant clones with BSA expression (four hCPO WT and four hCPO E310Q) were transferred from their respective 6-well plates to 100 mm dishes for further expansion. After reaching 70-100% confluency, the resistant clones were harvested and lysed in lysis buffer (see Chapter 2, section 11). Eight resistant clones (four hCPO WT and four hCPO E310Q) showed no CPO expression, as determined by western blotting analysis (Fig. 21A)

In a second attempt to isolate stably expressing hCPO WT and E310Q transfectants, Caco-2 cells were seeded at $3x10^5$ cells/well in 6-well plate and transfected with WT, E310Q and 3.1 pcDNA (see above). Resistant clones with healthy morphologies (nineteen hCPO WT, nineteen hCPO E310Q, and three empty vector) were harvested from the 100 mm dishes and expanded in 24-well plates. Twenty-one resistant clones (eleven hCPO WT and ten hCPO E310Q) showed no CPO expression, as determined by western blotting analysis (Fig. 21B).

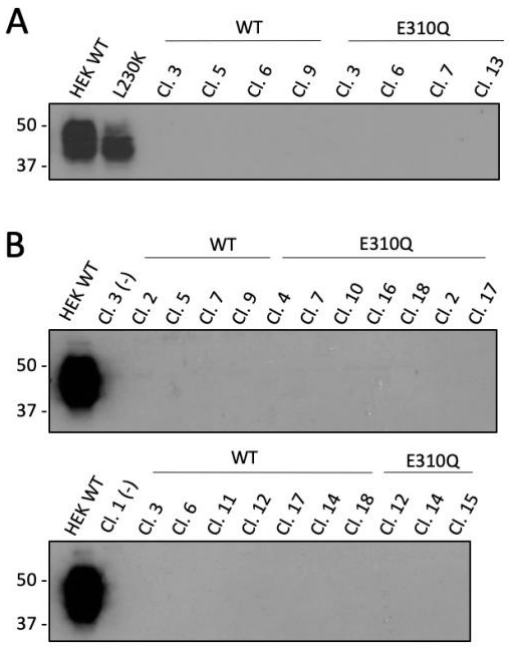

Figure 21. Resistant clones from stably expressing hCPO WT and hCPO E310Q transfectants showed no CPO expression. Caco-2 cells were transfected and subsequently selected with G418 for three weeks. WT (n = 15) and E310Q (n = 14) resistant clones were harvested and analyzed via western blotting to test for hCPO expression. Extracts from HEK293T cells transfected with plasmids expressing hCPO, either WT (A,B) or L230K (A), and two empty vector clones Cl.3, 1 (-) (B) were used as controls.

4. Immunocytochemistry

Previous research has shown an association between CPO and LDs in canine kidney cells and has further suggested the possibility of an association between CPO and chylomicrons. In this study, we investigated the association between CPO, LDs, and chylomicrons in Caco-2 cells, which are known to closely mimic intestinal physiology and anatomy. To test the potential association between CPO and LDs or chylomicrons, immunocytochemistry was performed using a culture of Caco-2 cells that contained pooled hCPO WT stable transfectants. Notably, the usage of a Caco-2 cell line that stably expressed both hCPO WT and E310Q would have been more optimal than a pooled culture containing a small number of hCPO WT expressing cells.

Nonetheless, transitioning to using this pooled culture was deemed as the next logical step after numerous unsuccessful attempts to create hCPO WT and E310Q stable transfectants.

In order to compare any differences in association in differentiated and undifferentiated cells, a batch of Caco-2 cells from the pooled culture was differentiated. Caco-2 cells were differentiated into an epithelial monolayer of mature enterocytes by growing them in 6-well plates with coated coverslips (see Chapter 2, section 13) for 21 days. Prior to being analyzed via immunocytochemistry, undifferentiated and differentiated cells were fixed at 12 h and 48 h post-feeding. For immunocytochemistry, the samples were divided into two groups: LD/CPO and apoB/CPO. Following permeabilization and blocking (see above), cells in the LD/CPO group were labeled with RP3-CPO primary antibody, whereas cells in the apoB/CPO group were labeled with both RP3-CPO and Apolipoprotein B primary antibodies. Similarly, secondary antibody anti-rabbit Alexa Fluor 555 was used for the LD/CPO and the same antibody in conjunction with anti-mouse Alexa Fluor 488 was used for the apoB/CPO group.

Observation of the slides under the fluorescent microscope revealed that the number CPO expressing cells was low (not shown). This reduced number of transfected cells presented a challenge to assess association of CPO with LDs and chylomicrons. In addition, differentiated cells presented blurry staining, possibly as a result of the thickness of the cell monolayer (not shown). In light of these results, it was decided to transiently transfect the pooled culture of undifferentiated Caco-2 to increase the number of CPO expressing cells. Transiently transfected cells from the pooled Caco-2 cell culture were fixed at 6 h, 12 h, and 24 h post-feeding and analyzed via immunocytochemistry (Figs. 22,23).

Thirteen fields of view of each experimental condition were analyzed, with at least one hCPO transfected cell in each field of view (range 1-4, $\bar{x}$: 1.6). LDs, apoB, and CPO puncta were

manually counted, and the association was computed by calculating how many of the total LD or apoB molecules were seen associated with CPO (expressed as % of total LDs or apoB). Similarly, association was computed in non-transfected cells to evaluate baseline background staining.

CPO was associated with both LDs and apoB, with the highest % association found 12 h post-feeding (29.8% and 30.4%, respectively; Figs. 22A; 23A). Both LD and apoB presented the lowest association with CPO at 6 h (17.9% and 22.8%, respectively), reaching a maximum at 12 h, and later slightly decreasing at 24 h (26.3% and 29.1%, respectively; Figs. 22B; 23B). CPO-expressing cells exhibited defined, bright red puncta, which appeared as fuzzier and dimmer background spots in non-expressing cells. Non-transfected cells consistently exhibited lower association values, suggesting that the computed CPO, LD, and apoB puncta in transfected cells likely were not background staining (not shown). However, we note that the distribution of CPO and apoB puncta in Fig. 23A does not appear to correlate with the association values reported in Fig. 23B. Thus, we express a limited level of confidence regarding the apoB:CPO association reported herein.

LD:CPO association was found to be significantly different at 6, 12, and 24 h ($F = 4.27$, $p = 0.02$), whereas apoB:CPO association was found to be non-significant between the time points ($F = 1.8$, $p = 0.179$).

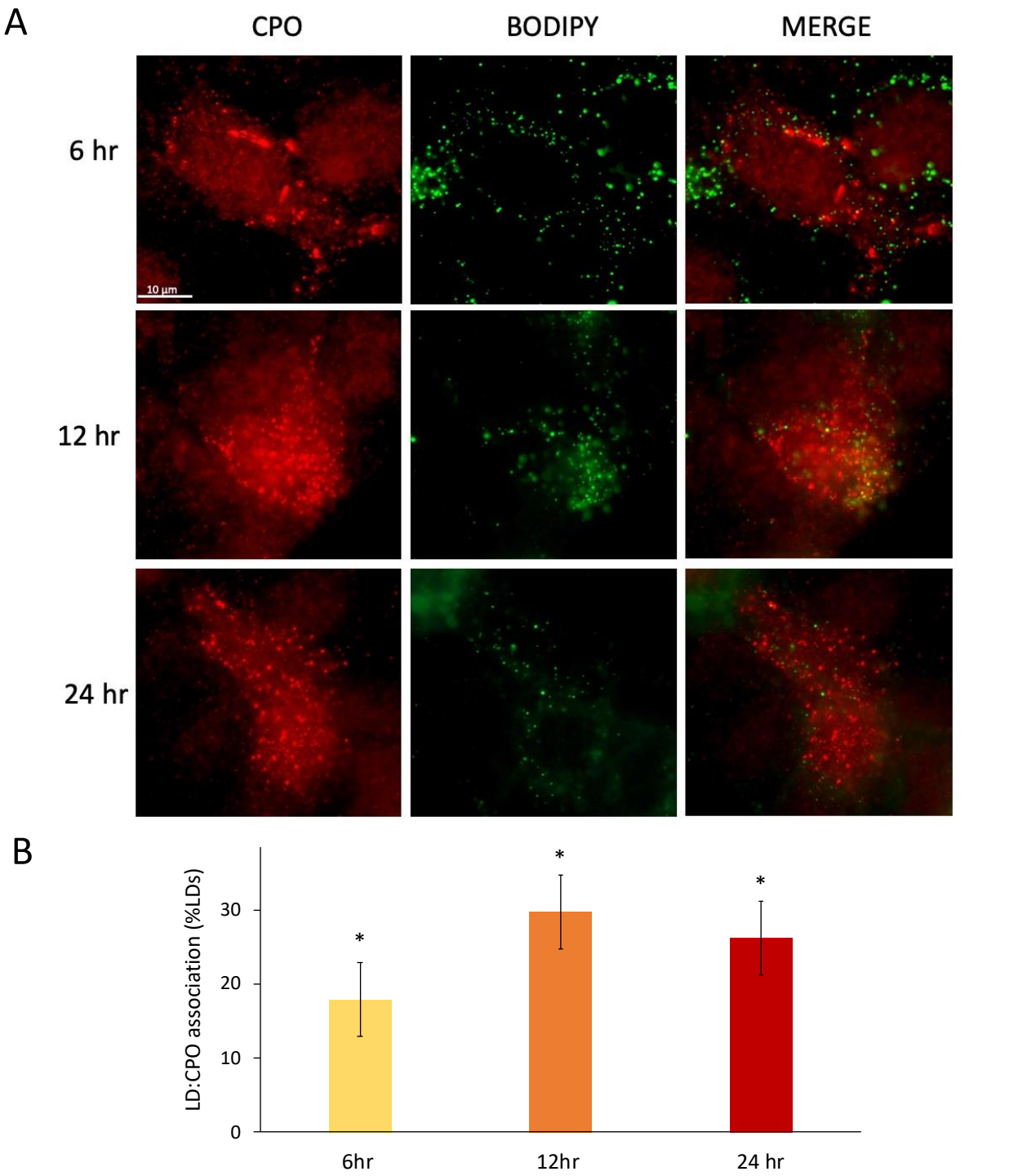

Figure 22. CPO associated with lipid droplets in transiently transfected Caco-2 cells. A pooled culture of stable CPO transfectants was transiently transfected and fixed at 6, 12, and 24 h post feeding. A. The cells were immunostained with CPO antibody and BODIPY to observe lipid droplets. CPO-expressing cells exhibited defined, bright red puncta (A, cell on the left of upper left panel), which appeared as fuzzier and dimmer background spots in non-expressing cells (A, cell on the right of upper left panel). B. The LD:CPO association (%LDs) was determined using thirteen fields of view, and was found to be the highest at the 12 h mark. The LD:CPO association was significantly different at 6 h, 12 h, and 24, h ($p = 0.02$), by one-way ANOVA.

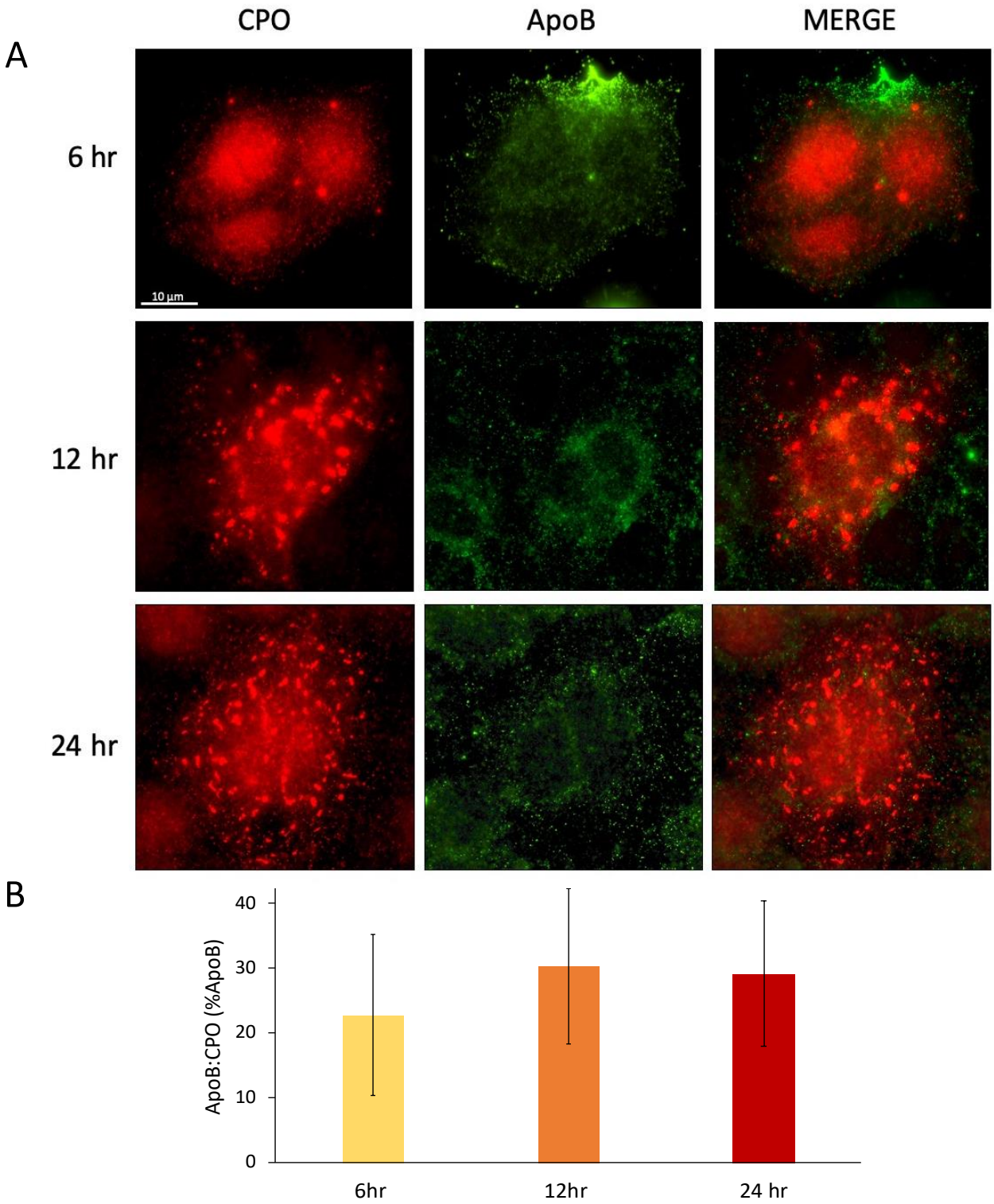

Figure 23. CPO associated with apoB, a marker for chylomicrons, in transiently transfected Caco-2 cells. A pooled culture of stable CPO transfectants was transiently transfected and fixed at 6, 12, and 24 h post feeding. A. The cells were immunostained with CPO and apoB antibodies. B. The apoB:CPO association (%apoB) was determined using thirteen fields of view, and was found to be the highest at the 12 h mark. ApoB:CPO association did not significantly vary amongst the three time points.

CHAPTER 4

DISCUSSION

1. Evaluation of expression data

Protein expression data collected from several RNA-Seq and microarray databases (see Chapter 3) collectively identify the small intestine, more specifically the ileum epithelium, as the site of highest CPO expression in the human body. In this case, RNA-Seq and microarray data indicate the same site of maximum expression; however, these expression scores vary greatly from each other. More importantly, microarray data appears to show rather high levels of background expression. This becomes evident when comparing the expression scores for the anatomical structure with the lowest expression in the Affymetrix Human Genome U133 Plus 2.0 Array microarray platform and the mRNASeq_HUMAN_GL-1 platform. As it would be expected, the lowest expression reported in the mRNASeq platform is 0, which corresponds to the spinal cord motor neuron. In contrast, the lowest mean score in the affymetrix microarray platform is 179, which corresponds to the myeloid progenitor cell.

For many years following the creation of the first Affymetrix GeneChip® technology in 1994, microarrays have continued to play an essential role in research (Lenoir and Giannella, 2006). However, microarrays are being gradually replaced by newer, more complex sequencing technologies such as RNA-Seq. This transitioning to RNA-Seq can be understood in light of the numerous disadvantages associated with microarrays such as limited detection range, cross-hybridization, and high background noise (Song et al., 2015; Wang et al., 2009; Zhao et al., 2014). My findings show high levels of background noise in data collected from Bgee and Affymetrix Human Genome U133 Plus 2.0 Array platform. While it is important to identify

background noise, it may not drastically affect the interpretation of expression data with high expression levels. In contrast, interpretation of medium and low expression data can be easily affected by background noise, thus leading to an overestimation of expression. The present study found medium CPO expression in a myriad of anatomical entities such as the epididymal cauda and adrenal gland (Figs. 7D; 8B). In light of the large background noise (i.e. 231 in Affymetrix Human Genome U133 Plus 2.0), a corrected expression would place the aforementioned entities in the low expression range. This correction of the microarray expression data would then match the values of similar structures reported by RNA-Seq (Figs. 7A-C; 8A).

Additionally, analysis of CPO expression in IBD pathologies revealed several patterns that could provide insight into the physiological function of CPO. For example, Crohn's disease presents significant upregulation in colon tissue and downregulation in ileal tissue. Interestingly, the upregulation in colon tissue, reported to be upwards of 12-fold ($p < 0.001$; Fig. 12C), is the largest upregulation of CPO across all pathologies shown in the Genevestigator database (Peck et al., 2015). Preliminarily, this finding can raise questions regarding the underlaying mechanisms of Crohn's disease. Such mechanisms can presumably lead to overexpression of CPO in the colon, or alternatively, rely on a transport and subsequent relocation of CPO from ileal epithelium to the colon. Although it would be exciting to investigate these possibilities, low expression of CPO in the colon (1.5 pTPM in HPA, 1 pTPM in GTEx, and 0.4 pTPM in FANTOM 5; Fig. 7A-C, respectively) seems to undermine the relevance of such a large fold change.

The findings of Haberman et al. (2014) point to modulation of CPO expression in the ileum with promising implications. The study, which included Crohn's disease patients aged 0-17 (Satsangi et al., 2006) showed a significant downregulation of CPO in ileal tissue of ($p < 0.001$;

Fig. 12C). Since the ileum is the site of highest CPO expression (Figs. 7, 8), a downregulation of about 4-fold is considered prominent, and further suggestive of a function of CPO in this disease.

Crohn's disease is characterized by chronic intestinal inflammation emerging at an interface of genetic susceptibility, environmental factors, imbalance of enteric bacteria, and a strong immune response led by T-cells (Sartor, 2006; Shanahan, 2002). This pathology presents a variety of phenotypes, which have been classified into age (A1- 3), location (L1- 4) and behavior (B1- 3) categories according to the Montreal guidelines (Satsangi et al., 2006). Thia et al. (2010) investigated the clinical characteristics of 306 patients who had been diagnosed with Crohn's disease. They showed that 56% of diagnoses corresponded to patients aged 17 – 40 (A2), and only 11% corresponded to patients aged 0-16 (A1). Moreover, the study showed that the terminal ileum was the most commonly affected site, accounting for 45% of the total. While the disease may shift amongst behavior phenotypes (B1-3), it reportedly maintains a constant location (Baumgart and Sandborn, 2012). This would suggest that the findings of Haberman et al. (2014), although pertaining to a younger patient cohort, may likely apply to older patients as well.

Although a series of pathologies are expressed at varying levels across phenotypes (Antoni et al., 2014), a common feature in the pathogenesis of Crohn's disease is a defective intestinal mucosal barrier. The mucosal barrier represents the first line of immunological defense in the small intestine, so the emergence of disease associated perturbations directly affects intestinal function and integrity (Sartor, 2006). Such defects include, but are not restricted to, decreased presence of mucus biofilm, alterations in expression of intercellular junction proteins, increased transepithelial permeability linked to migration of leukocytes, and decreased production of antimicrobial peptides (Antoni et al., 2014; Baumgart and Sandborn, 2012).

Ultimately, these factors have been shown to produce epithelial apoptosis in Crohn's disease patients (Antoni et al., 2014; Zeissig et al., 2007). Thus, destruction of mucosal barrier epithelium, with appearance of gaps and microerosions (Antoni et al., 2014), would explain the downregulation of CPO in the ileal epithelium reported by Haberman et al. (2014).

Additionally, RNA-Seq data from Bgee showed high expression of CPO in Peyer's patches (Fig. 7D), which are intimately associated with the immunological response of Crohn's disease. These structures represent the first-line of adaptive immunological defense in the intestine (Kogan and von Andrian, 2008) and have significant presence in the ileum (Shah and Misra, 2011). In an effort to assess the presence of Peyer's patches in the small intestine, a preliminary study by Cornes (1965) set out to quantify these organs across selected age groups ranging from gestational period to adolescence. The study showed that in adolescents, the number of patches ranged from 100 to 239, but most importantly, that the number and size of the patches increased with age (Cornes, 1965). Thus, an increased abundance of Peyer's patches in later developmental stages could imply a similar pattern in CPO expression. Interestingly, RNA-Seq data from Bgee shows highest expression in early adulthood (25 – 44 years old), and moderate expression in late adulthood, with no reports of earlier developmental stages. Taken together, these reports may indicate that Peyer's patches and CPO expression could follow a bell-shaped curve; increasing from childhood until early adulthood, and subsequently decreasing after. A modulation of CPO expression resulting from variation in abundance of Peyer's patches could suggest that CPO activity is relevant in these organs, possibly aiding in their proliferation. Another potential role of CPO in immunity is suggested by significant expression in basophils (Fig. 9). Although basophils are the least abundant granulocytes, they are known to play important roles in adaptive immunity, namely allergic reactions (Min et al., 2004). Future studies

should further investigate these promising associations to gain more insight into the physiological function of CPO.

2. Evaluation of Caco-2 transfection attempts

We set out to investigate the function of CPO in a stably-expressing Caco-2 cell model. Two separate attempts were made to stably transfect and isolate G418 resistant hCPO WT and E310Q clones; however, we found no CPO expression amongst the analyzed clones (n = 38; Fig. 21). Analysis of CPO expression following a semi-transient transfection presumably showed expression in both hCPO WT and E310Q samples (Fig. 20). Nonetheless, this possibility was refuted after close inspection of the X-ray film, which showed a faint shadow on an empty lane, thus hinting at the possibility of spillover.

These results prompted us to question if the inability to produce stable transfectants was related to CPO or rather cell-line specific. Lyons and Fricker (2011) reported creation of an epithelial Madin-Darby canine kidney (MDCK) cell line stably expressing CPO. MDCK cells can be easily transfected (Di et al., 2011) unlike Caco-2 cells, which are well-known for being difficult to transfect (Cerda et al., 2015). Other studies have successfully stably transfected Caco-2 cells with a myriad of proteins including cyclooxygenase-2 (Tsujii et al., 1997), insulin-like growth factor I (IGF-I) receptors (Di Popolo et al., 2000), and integrin-initiated extracellular signal-regulated (ERK) kinase (Sanders and Basson, 2000). Notably, these studies used high-end transfection reagents (Lipofectamine and Lipofectamine PLUS), which contrasts with the cost-effective reagent, namely PEI, used in this study. To our knowledge, we are the first study to report use of PEI to produce transiently and stably expressing Caco-2 cells. The reliability of the PEI reagent was demonstrated in separate transfection experiments with GFP (Figs. 18,19).

Given the abundant evidence showing that Caco-2 cells can be stably transfected, we wondered whether the inability to produce stable transfectants was CPO-specific. More specifically, we hypothesized that CPO overexpression in Caco-2 cells could be cytotoxic. Due to the low endogenous CPO presence in Caco-2 cells (Fig. 10), an increased expression post-transfection could be toxic to the cells. Reeves et al. (2002) reported numerous cases of cytotoxicity linked to protein expression when attempting to create stably expressing cell lines. In this regard, Andréll and Tate (2013) argue against constitutive expression of membrane proteins when seeking to create stably expressing mammalian cell lines. Protein overexpression induces cellular stress, further leading to loss of expression over time and delayed cell growth. To address these issues, Andréll and Tate (2013) propose the use of an inducible system such as tetracycline, which enables inducible expression of the protein for short time periods. In our study, CPO was presumably constitutively expressed for 42-52 days, from time of transfection until harvesting of the cells. It is possible that CPO was expressed upon transfection and that factors such as extended length of the G418 selection (40-50 days), excessive stress and toxicity could have resulted in low protein expression at the time of analysis.

3. CPO association with LDs and apoB in Caco-2 cells

Using a culture with pooled stably expressing Caco-2 transfectants, we show that CPO associated with LDs and apoB; nonetheless, we acknowledge a discrepancy between the quantified apoB:CPO association, which ranges from 22-30%, and the association shown in the immunocytochemistry images (Fig. 23A), which appears lower. CPO association with LDs was found to be dependent on the time elapsed post-feeding, as seen by lowest association at 6 h (18%) and highest at 24 h (30%), followed by a small decrease at 48 h (26%) ($p = 0.02$; Fig.

22B). A similar, yet not statistically significant association pattern was shown by apoB (Fig. 23B).

The observed association between CPO and LDs expand on findings previously reported by Burke et al. (2018), which showed this association in differentiated MDCK cells. Burke et al. (2018) found an association that, similarly to ours, varied upon time elapsed from the last feeding until fixation; however, two notable differences were found when comparing association patterns. First, we report an increase in association from 6 h and 12 h time points (12%) which is 6-fold higher than that reported by Burke et al. (2018). Secondly, the decrease in association from 12 h and 48 h time points (4%) represents a fourth of that reported in Burke et al. (2018). Despite use of similar methods by Burke et al. (2018), the variability between association levels could be due to differences in quantification techniques. Overall, the association of CPO with LDs confirms results previously reported by Burke et al. (2018) in a more optimal intestinal model, and further strengthens the hypothesis that CPO plays a role in lipid formation in enterocytes.

We also report association of CPO with apoB in stably expressing undifferentiated Caco-2 cells; nonetheless, we express a limited level of confidence in these results due to an apparent discrepancy between the immunocytochemistry images and the quantification data. Although not statistically significant, our results point to a similar pattern of association to that of LDs at 6 h, 12 h, and 48 h time points (Fig. 23B). A rather weak fluorescence of apoB detected through immunocytochemistry (Fig. 23A) suggested lower abundance than initially expected. Thus, it appeared that low detection could correlate to low apoB secretion in undifferentiated Caco-2 cells. Early work by Reisher et al. (1993) confirmed secretion of apoB in undifferentiated Caco-2 cells, an event previously thought to occur only after differentiation. In a subsequent study,

Luchoomun et al. (1997) stably expressed the apoB48 isoform in Caco-2 cells and found that over 50% of the apoB48 was degraded intracellularly.

Since the present study sought to investigate the association between CPO and chylomicrons, we wondered if levels of apoB48, a well-known chylomicron biomarker (Nakajima et al., 2014), could predict the presence of chylomicrons in undifferentiated Caco-2 cells. The findings of Luchoomun et al. (1997) indicate that undifferentiated Caco-2 cells can secrete chylomicron-like particles if subject to conditions that mimic a postprandial state, which can be achieved by supplementation of media with oleic acid. This evidence suggests that apoB48 metabolism, and to a lesser extent chylomicrons, can be studied in undifferentiated Caco-2 cells. Nonetheless, it appears that the use of differentiated Caco-2 cells may be more appropriate, as demonstrated by their use in a large amount of studies on chylomicron and apoB48 metabolism (Jiao et al., 1990; Nauli et al., 2014; van Greevenbroek et al., 1996).

In addition, it is uncertain what percentage of the apoB detected in the immunocytochemistry assay pertained to apoB48 and/or apoB100, since the primary antibody appears to detect both isoforms of the apolipoprotein. Therefore, future studies should consider the use of western blotting analysis to resolve both apoB isoforms, or alternatively, the use of an apoB48-specific antibody. In addition, it is suggested to mimic postprandial state conditions in differentiated cells by use of oleic acid and/or lecithin to further enhance chylomicron production (Nauli et al., 2014). Overall, the evidence suggests that the conditions used in this study were not ideal for an adequate study of chylomicron metabolism.

4. Conclusion

In conclusion, the results of the present study show that CPO associates with LDs in undifferentiated Caco-2 cells and that this association is dependent on time elapsed post-feeding. Furthermore, this study used an intestinal model to expand on the findings of Burke et al. (2018), thus strengthening the hypothesis that CPO plays an intracellular role in the formation of lipid droplets in enterocytes. Future studies aiming to analyze intracellular CPO substrates and binding partners could provide insight into the potential physiological function of CPO. We also report that CPO associates with apoB in undifferentiated Caco-2 cells; however, we express a limited level of confidence in these results since the immunocytochemistry images presumably suggest lower association levels than those reported in the quantification data. Further, we acknowledge that the parameters used to study chylomicron metabolism were not optimal and hence should be revised ahead of future studies. Further work should investigate the association of CPO and chylomicrons in differentiated Caco-2 cells in the presence and absence of oleic acid. Lastly, a comprehensive analysis of CPO expression data revealed modulation of CPO expression in Crohn's disease as well as expression in basophils and Peyer's patches, all of which could be indicative of a function of CPO in immunity.

REFERENCES

Andréll, J., and Tate, C.G. (2013). Overexpression of membrane proteins in mammalian cells for structural studies. Molecular membrane biology *30*, 52-63.

Antoni, L., Nuding, S., Wehkamp, J., and Stange, E.F. (2014). Intestinal barrier in inflammatory bowel disease. World journal of gastroenterology: WJG *20*, 1165.

Arijs, I., De Hertogh, G., Lemaire, K., Quintens, R., Van Lommel, L., Van Steen, K., Leemans, P., Cleynen, I., Van Assche, G., and Vermeire, S. (2009). Mucosal gene expression of antimicrobial peptides in inflammatory bowel disease before and after first infliximab treatment. PloS one *4*.

Bastian, F., Parmentier, G., Roux, J., Moretti, S., Laudet, V., and Robinson-Rechavi, M. (2008). Bgee: integrating and comparing heterogeneous transcriptome data among species. Paper presented at: International Workshop on Data Integration in the Life Sciences (Springer).

Baumgart, D.C., and Sandborn, W.J. (2012). Crohn's disease. The Lancet *380*, 1590-1605.

Bayly, G.R. (2014). Lipids and disorders of lipoprotein metabolism. In Clinical Biochemistry: Metabolic and Clinical Aspects (Elsevier), pp. 702-736.

Brahm, A.J., and Hegele, R.A. (2015). Chylomicronaemia—current diagnosis and future therapies. Nature Reviews Endocrinology *11*, 352.

Broer, S. (2008). Apical transporters for neutral amino acids: physiology and pathophysiology. Physiology *23*, 95-103.

Burke, L.C., Ezeribe, H.O., Kwon, A.Y., Dockery, D., and Lyons, P.J. (2018). Carboxypeptidase O is a lipid droplet-associated enzyme able to cleave both acidic and polar C-terminal amino acids. PLoS One *13*, e0206824.

Cartwright, I.J., and Higgins, J.A. (2001). Direct evidence for a two-step assembly of ApoB48-containing lipoproteins in the lumen of the smooth endoplasmic reticulum of rabbit enterocytes. J Biol Chem *276*, 48048-48057.

Cerda, M.B., Batalla, M., Anton, M., Cafferata, E., Podhajcer, O., Plank, C., Mykhaylyk, O., and Policastro, L. (2015). Enhancement of nucleic acid delivery to hard-to-transfect human colorectal cancer cells by magnetofection at laminin coated substrates and promotion of the endosomal/lysosomal escape. RSC Advances *5*, 58345-58354.

Chen, H., Jawahar, S., Qian, Y., Duong, Q., Chan, G., Parker, A., Meyer, J.M., Moore, K.J., Chayen, S., and Gross, D.J. (2001). Missense polymorphism in the human carboxypeptidase E gene alters enzymatic activity. Human mutation *18*, 120-131.

Cong, L., Cheng, Y., Cawley, N.X., Murthy, S.R., and Loh, Y.P. (2017). A Novel Single Nucleotide T980C Polymorphism in the Human Carboxypeptidase E Gene Results in Loss of Neuroprotective Function. PLoS One *12*, e0170169.

Cornes, J. (1965). Number, size, and distribution of Peyer's patches in the human small intestine: Part I The development of Peyer's patches. Gut *6*, 225.

Dash, S., Xiao, C., Morgantini, C., and Lewis, G.F. (2015). New Insights into the Regulation of Chylomicron Production. Annu Rev Nutr *35*, 265-294.

Demignot, S., Beilstein, F., and Morel, E. (2014). Triglyceride-rich lipoproteins and cytosolic lipid droplets in enterocytes: key players in intestinal physiology and metabolic disorders. Biochimie *96*, 48-55.

Di, L., Whitney-Pickett, C., Umland, J.P., Zhang, H., Zhang, X., Gebhard, D.F., Lai, Y., Federico III, J.J., Davidson, R.E., and Smith, R. (2011). Development of a new

permeability assay using low-efflux MDCKII cells. Journal of pharmaceutical sciences *100*, 4974-4985.

Di Popolo, A., Memoli, A., Apicella, A., Tuccillo, C., di Palma, A., Ricchi, P., Acquaviva, A.M., and Zarrilli, R. (2000). IGF-II/IGF-I receptor pathway up-regulates COX-2 mRNA expression and PGE 2 synthesis in Caco-2 human colon carcinoma cells. Oncogene *19*, 5517-5524.

Fricker, L.D. (2011). Carboxypeptidases. In xPharm: The Comprehensive Pharmacology Reference (Elsevier Inc.), pp. 1-4.

Fricker, L.D., Berman, Y.L., Leiter, E.H., and Devi, L.A. (1996). Carboxypeptidase E activity is deficient in mice with the fat mutation Effect on peptide processing. Journal of Biological Chemistry *271*, 30619-30624.

Fricker, L.D., and Snyder, S.H. (1982). Enkephalin convertase: purification and characterization of a specific enkephalin-synthesizing carboxypeptidase localized to adrenal chromaffin granules. Proceedings of the National Academy of Sciences *79*, 3886-3890.

Funke, B., Autschbach, F., Kim, S., Lasitschka, F., Strauch, U., Rogler, G., Gdynia, G., Li, L., Gretz, N., and Macher-Goeppinger, S. (2009). Functional characterisation of decoy receptor 3 in Crohn's disease. Gut *58*, 483-491.

Garcia-Guerrero, M.C., Garcia-Pardo, J., Berenguer, E., Fernandez-Alvarez, R., Barfi, G.B., Lyons, P.J., Aviles, F.X., Huber, R., Lorenzo, J., and Reverter, D. (2018). Crystal structure and mechanism of human carboxypeptidase O: Insights into its specific activity for acidic residues. Proc Natl Acad Sci U S A *115*, E3932-E3939.

Haberman, Y., Tickle, T.L., Dexheimer, P.J., Kim, M.-O., Tang, D., Karns, R., Baldassano, R.N., Noe, J.D., Rosh, J., and Markowitz, J. (2014). Pediatric Crohn disease patients

exhibit specific ileal transcriptome and microbiome signature. The Journal of clinical investigation *124*, 3617-3633.

Harding, H.P., and Ron, D. (2002). Endoplasmic reticulum stress and the development of diabetes: a review. Diabetes *51*, S455-S461.

Head, K., and Jurenka, J. (2003). Inflammatory bowel disease Part 1: ulcerative colitis--pathophysiology and conventional and alternative treatment options. Alternative medicine review: a journal of clinical therapeutic *8*, 247.

Hooton, D., Lentle, R., Monro, J., Wickham, M., and Simpson, R. (2015). The secretion and action of brush border enzymes in the mammalian small intestine. In Reviews of physiology, biochemistry and pharmacology (Springer), pp. 59-118.

Howell, K.J., Kraiczy, J., Nayak, K.M., Gasparetto, M., Ross, A., Lee, C., Mak, T.N., Koo, B.-K., Kumar, N., and Lawley, T. (2018). DNA methylation and transcription patterns in intestinal epithelial cells from pediatric patients with inflammatory bowel diseases differentiate disease subtypes and associate with outcome. Gastroenterology *154*, 585-598.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W., and Zimmermann, P. (2008). Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. Advances in bioinformatics *2008*.

Hussain, M.M. (2014). Intestinal lipid absorption and lipoprotein formation. Current opinion in lipidology *25*, 200.

Ji, L., Wu, H.T., Qin, X.Y., and Lan, R. (2017). Dissecting carboxypeptidase E: properties, functions and pathophysiological roles in disease. Endocr Connect *6*, R18-R38.

Jiao, S., Moberly, J.B., and Schonfeld, G. (1990). Editing of apolipoprotein B messenger RNA in differentiated Caco-2 cells. Journal of Lipid Research *31*, 695-700.

Karathia, H., Vilaprinyo, E., Sorribas, A., and Alves, R. (2011). Saccharomyces cerevisiae as a model organism: a comparative study. PloS one *6*, e16015.

Kochanek, K.D., Murphy, S.L., Xu, J., and Arias, E. (2019). Deaths: final data for 2017.

Kogan, A.N., and von Andrian, U.H. (2008). Lymphocyte trafficking. In Microcirculation (Elsevier), pp. 449-482.

Koopal, C., Marais, A.D., and Visseren, F.L. (2017). Familial dysbetalipoproteinemia: an underdiagnosed lipid disorder. Current Opinion in Endocrinology, Diabetes and Obesity *24*, 133-139.

Kraemer, F.B., Khor, V.K., Shen, W.-J., and Azhar, S. (2013). Cholesterol ester droplets and steroidogenesis. Molecular and cellular endocrinology *371*, 15-19.

Lenoir, T., and Giannella, E. (2006). The emergence and diffusion of DNA microarray technology. Journal of biomedical discovery and collaboration *1*, 11.

Leung, L.L., Nishimura, T., and Myles, T. (2008). Regulation of tissue inflammation by thrombin-activatable carboxypeptidase B (or TAFI). In Current Topics in Complement II (Springer), pp. 59-67.

Luchoomun, J., Zhou, Z., Bakillah, A., Jamil, H., and Hussain, M.M. (1997). Assembly and secretion of VLDL in nondifferentiated Caco-2 cells stably transfected with human recombinant ApoB48 cDNA. Arteriosclerosis, thrombosis, and vascular biology *17*, 2955-2963.

Lyons, P.J., and Fricker, L.D. (2011). Carboxypeptidase O is a glycosylphosphatidylinositol-anchored intestinal peptidase with acidic amino acid specificity. J Biol Chem *286*, 39023-39032.

McGettigan, P.A. (2013). Transcriptomics in the RNA-seq era. Current opinion in chemical biology *17*, 4-11.

Meisenberg, G., and Simmons, W.H. (2016). Principles of Medical Biochemistry E-Book (Elsevier Health Sciences).

Meyers, A., Weiskittel, T.M., and Dalhaimer, P. (2017). Lipid Droplets: Formation to Breakdown. Lipids *52*, 465-475.

Min, B., Prout, M., Hu-Li, J., Zhu, J., Jankovic, D., Morgan, E.S., Urban Jr, J.F., Dvorak, A.M., Finkelman, F.D., and LeGros, G. (2004). Basophils produce IL-4 and accumulate in tissues after infection with a Th2-inducing parasite. The Journal of experimental medicine *200*, 507-517.

Mooradian, A.D. (2009). Dyslipidemia in type 2 diabetes mellitus. Nat Clin Pract Endocrinol Metab *5*, 150-159.

Nakajima, K., Nagamine, T., Fujita, M.Q., Ai, M., Tanaka, A., and Schaefer, E. (2014). Apolipoprotein B-48: a unique marker of chylomicron metabolism. In Advances in clinical chemistry (Elsevier), pp. 117-177.

Natoli, M., Leoni, B.D., D'Agnano, I., Zucco, F., and Felsani, A. (2012). Good Caco-2 cell culture practices. Toxicology in vitro *26*, 1243-1246.

Nauli, A.M., Sun, Y., Whittimore, J.D., Atyia, S., Krishnaswamy, G., and Nauli, S.M. (2014). Chylomicrons produced by Caco-2 cells contained ApoB-48 with diameter of 80–200 nm. Physiological reports *2*, e12018.

Novakovic, D., Kuo, A.C., Lin, J.H., Koschinsky, M.L., and Boffa, M.B. (2012). Identification of tristetraprolin as a factor that modulates the stability of the TAFI transcript through binding to the 3'-untranslated region. J Thromb Haemost *10*, 887-894.

Oelkers, P., Tinkelenberg, A., Erdeniz, N., Cromley, D., Billheimer, J.T., and Sturley, S.L. (2000). A lecithin cholesterol acyltransferase-like gene mediates diacylglycerol esterification in yeast. J Biol Chem *275*, 15609-15612.

Park, S.-J., Yoon, B.-H., Kim, S.-K., and Kim, S.-Y. (2019). GENT2: an updated gene expression database for normal and tumor tissues. BMC medical genomics *12*, 101.

Peck, B.C., Weiser, M., Lee, S.E., Gipson, G.R., Iyer, V.B., Sartor, R.B., Herfarth, H.H., Long, M.D., Hansen, J.J., and Isaacs, K.L. (2015). MicroRNAs classify different disease behavior phenotypes of Crohn's disease and may have prognostic utility. Inflammatory bowel diseases *21*, 2178-2187.

Pontén, F., Schwenk, J.M., Asplund, A., and Edqvist, P.H. (2011). The Human Protein Atlas as a proteomic resource for biomarker discovery. Journal of internal medicine *270*, 428-446.

Randolph, G.J., and Miller, N.E. (2014). Lymphatic transport of high-density lipoproteins and chylomicrons. The Journal of clinical investigation *124*, 929-935.

Reeves, P.J., Kim, J.-M., and Khorana, H.G. (2002). Structure and function in rhodopsin: a tetracycline-inducible system in stable mammalian cell lines for high-level expression of opsin mutants. Proceedings of the National Academy of Sciences *99*, 13413-13418.

Reisher, S.R., Hughes, T.E., Ordovas, J.M., Schaefer, E.J., and Feinstein, S.I. (1993). Increased expression of apolipoprotein genes accompanies differentiation in the intestinal cell line Caco-2. Proceedings of the National Academy of Sciences *90*, 5757-5761.

Roux, A., and Loewith, R. (2017). Tensing Up for Lipid Droplet Formation. Dev Cell *41*, 571-572.

Saka, H.A., and Valdivia, R. (2012). Emerging roles for lipid droplets in immunity and host-pathogen interactions. Annual review of cell and developmental biology *28*, 411-437.

Sanders, M.A., and Basson, M.D. (2000). Collagen IV-dependent ERK activation in human Caco-2 intestinal epithelial cells requires focal adhesion kinase. Journal of Biological Chemistry *275*, 38040-38047.

Sartor, R.B. (2006). Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. Nature clinical practice Gastroenterology & hepatology *3*, 390-407.

Satsangi, J., Silverberg, M., Vermeire, S., and Colombel, J. (2006). The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. Gut *55*, 749-753.

Shah, T.R., and Misra, A. (2011). 8-Proteomics. Challenges in Delivery of Therapeutic Genomics and Proteomics, 387-427.

Shanahan, F. (2002). Crohn's disease. The Lancet *359*, 62-69.

Soares, P.A., Vaz, A.F., Correia, M.T., Pessoa Jr, A., and Carneiro-da-Cunha, M.G. (2012). Purification of bromelain from pineapple wastes by ethanol precipitation. Separation and purification technology *98*, 389-395.

Song, N., Kim, N., Xiao, R., Choi, H., Chun, H.I., Kang, M.H., Kim, J.H., Seo, K., Soundrarajan, N., Do, J.T., *et al.* (2015). Lack of Cytosolic Carboxypeptidase 1 Leads to Subfertility due to the Reduced Number of Antral Follicles in pcd3J-/- Females. PLoS One *10*, e0139557.

Thia, K.T., Sandborn, W.J., Harmsen, W.S., Zinsmeister, A.R., and Loftus Jr, E.V. (2010). Risk factors associated with progression to intestinal complications of Crohn's disease in a population-based cohort. Gastroenterology *139*, 1147-1155.

Tobey, N.A., Lyn-Cook, L.E., Ulshen, M.H., and Heizer, W.D. (1986). Intestinal brush border peptidases: activities in normal and abnormal peroral intestinal biopsy specimens. The Journal of laboratory and clinical medicine *107*, 221-227.

Tortora, G.D., Bryan (2009). Principles of anatomy and physiology, 12 edn (John Wiley & Sons).

Tsujii, M., Kawano, S., and DuBois, R.N. (1997). Cyclooxygenase-2 expression in human colon cancer cells increases metastatic potential. Proceedings of the National Academy of Sciences *94*, 3336-3340.

van Greevenbroek, M.M., van Meer, G., Erkelens, D.W., and de Bruin, T.W. (1996). Effects of saturated, mono-, and polyunsaturated fatty acids on the secretion of apo B containing lipoproteins by Caco-2 cells. Atherosclerosis *121*, 139-150.

Vevea, J.D., Garcia, E.J., Chan, R.B., Zhou, B., Schultz, M., Di Paolo, G., McCaffery, J.M., and Pon, L.A. (2015). Role for Lipid Droplet Biogenesis and Microlipophagy in Adaptation to Lipid Imbalance in Yeast. Dev Cell *35*, 584-599.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nature reviews genetics *10*, 57-63.

Wedemeyer, J., Tsai, M., and Galli, S.J. (2000). Roles of mast cells and basophils in innate and acquired immunity. Current opinion in immunology *12*, 624-631.

Welte, M.A. (2015). Expanding roles for lipid droplets. Current biology *25*, R470-R481.

Welte, M.A., and Gould, A.P. (2017). Lipid droplet functions beyond energy storage. Biochim Biophys Acta Mol Cell Biol Lipids *1862*, 1260-1272.

Wilfling, F., Haas, J.T., Walther, T.C., and Farese, R.V., Jr. (2014). Lipid droplet biogenesis. Curr Opin Cell Biol *29*, 39-45.

Xiao, C., Hsieh, J., Adeli, K., and Lewis, G.F. (2011). Gut-liver interaction in triglyceride-rich lipoprotein metabolism. Am J Physiol Endocrinol Metab *301*, E429-446.

Xiao, C., Stahel, P., and Lewis, G.F. (2018). Regulation of Chylomicron Secretion: Focus on Post-Assembly Mechanisms. Cell Mol Gastroenterol Hepatol *7*, 487-501.

Zeissig, S., Bürgel, N., Günzel, D., Richter, J., Mankertz, J., Wahnschaffe, U., Kroesen, A.J., Zeitz, M., Fromm, M., and Schulzke, J.D. (2007). Changes in expression and distribution of claudin 2, 5 and 8 lead to discontinuous tight junctions and barrier dysfunction in active Crohn's disease. Gut *56*, 61-72.

Zhao, S., Fung-Leung, W.-P., Bittner, A., Ngo, K., and Liu, X. (2014). Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PloS one *9*.

Zheng, M., Streck, R., Scott, R., Seidah, N., and Pintar, J. (1994). The developmental expression in rat of proteases furin, PC1, PC2, and carboxypeptidase E: implications for early maturation of proteolytic processing capacity. Journal of Neuroscience *14*, 4656-4673.

Zheng, S., Cherniack, A.D., Dewal, N., Moffitt, R.A., Danilova, L., Murray, B.A., Lerario, A.M., Else, T., Knijnenburg, T.A., and Ciriello, G. (2016). Comprehensive pan-genomic characterization of adrenocortical carcinoma. Cancer cell *29*, 723-736.

Zimmermann, P., Hennig, L., and Gruissem, W. (2005). Gene-expression analysis and network discovery using Genevestigator. Trends in plant science *10*, 407-409.

Zurzolo, C., and Simons, K. (2016). Glycosylphosphatidylinositol-anchored proteins: Membrane organization and transport. Biochimica et Biophysica Acta (BBA)-Biomembranes *1858*, 632-639.