UNIVERSITY OF CALIFORNIA SAN DIEGO

Targeted detection of secretion pathway protein interactions for therapeutic production and cellular engineering

A Thesis submitted in partial satisfaction of the requirements for the degree Master of Science

in

Bioengineering

by

Caressa Mystique Robinson

Committee in charge:

Professor Nathan Lewis, Chair
Professor Adam Engler, Co-chair
Professor Prashant Mali
Professor Samara Reck-Peterson

2020

The Thesis of Caressa Mystique Robinson is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____
Co-chair

_____
Chair

University of California San Diego

2020

EPIGRAPH

All cannot be lost
when there is still so much being found.

Lemony Snicket

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT.)

LIST OF FIGURES

LIST OF FIGURES (CONT.)

LIST OF TABLES

ACKNOWLEDGEMENTS

This project could not have been completed without the mentorship of Dr. Mojtaba Samoudi and my chair Dr. Nathan Lewis. Thank you, Moji for mentoring me in everything from cell culture, to making slides, to helping me perform western blot, after western blot. Thank you for sharing your abundance of knowledge with me and having the patience to guide me through the lab. Thank you, Dr. Lewis for bringing me into the lab and always providing a listening ear and good advice. This work would also not have been possible without the analytical guidance and expertise of Chih-Chung Kuo. Finally, thank you also to Dr. Phillip Sphan, Jamie Lee, and Dr. Ben Kellerman for keeping the wet lab interesting. I am forever grateful for the support of the entire Lewis Lab at UCSD.

On a personal note, thank you to my parents, Donna and David Robinson, for getting me where I am today. I hope this paper clears up what I have been doing for two years. However, despite your collective confusion, you have never failed to be supportive and understanding of my long nights on campus or single-minded studying. Thank you for fostering a wonder for science in me from the beginning and starting me on this path.

The Introduction and Chapter 1 contain some material from a manuscript currently in review for publication “In situ detection of protein interactions for recombinant therapeutic enzymes”. Samoudi, Mojtaba; Kuo, Chih-Chung; Robinson, Caressa M.; Shams-Ud-Doha, Km; Schinn, Song-Min; Kol, Stefan; Weiss, Linus; Bjorn, Sara Petersen; Voldborg, Bjorn G.; Campos, Alexandre Rosa; Lewis, Nathan E.. (2020). *In review.* The thesis author is a primary author of this material.

Chapter 2 is based on material currently being prepared for publication. “Identifying the essential proteins for supporting monoclonal antibody secretion by proximity-labeling mass spectrometry.” Samoudi, Mojtaba; Robinson, Caressa M.; Kuo, Chih-Chung; Pristovšek, Nuša; Shams-Ud-Doha, Km; Hansen, Henning G.; Kildegaard Helene F.; Lee, Gyun M.; Campos, Alexandre Rosa; Lewis, Nathan E.. The thesis author is expected to be a primary author of this material.

ABSTRACT OF THE THESIS

Targeted detection of secretion pathway protein interactions for therapeutic production and cellular engineering

by

Caressa Mystique Robinson

Master of Science in Bioengineering

University of California San Diego, 2020

Professor Nathan Lewis, Chair
Professor Adam Engler, Co-chair

Mammalian cells synthesize and secrete thousands of versatile proteins to interact with their environment. Over the past decades, this ability has been harnessed to produce a variety of lifesaving biotherapeutics to treat complex diseases, provide vaccines, or replace enzyme and hormone deficiencies—making the mammalian secretory pathway essential to protein biomanufacturing. In 2018, eleven of the top fifteen drugs by sales were recombinant proteins, produced almost exclusively in mammalian cells, mostly Chinese hamster ovary (CHO) cells [1]. These therapeutics may use the mammalian secretory pathway as an assembly line; however, the detailed pathway mechanisms remain poorly characterized. As such, there continues to exist bottlenecks in the secretory pathway such as machinery deficiencies that throttle protein secretion. First, we must better understand the protein interactions responsible for structure-specific post-

translational modifications, and second, determine the interactions that contribute to high recombinant protein yields. Employing systems-wide labeling of protein interactions within the mammalian secretory pathway using proximity-based biotinylation and mass spectrometry analysis, we validate the ability to identify interactions that occur within the endomembrane system and their potential as targets for cellular engineering for improved recombinant protein secretion.

Introduction

Recombinant proteins and other biotherapeutics have taken the pharmaceutical industry by storm, providing new solutions to a variety of complex and rare diseases. These products require equally involved host systems that have been engineered to produce these proteins. However, among differing products and expression systems there remains discrepancies in achievable protein yields and quality [2,3]. This may be attributed to post-translational regulatory pathways, transport, and other biological processing, but which of these pathways are responsible is largely unknown [4]. System-wide proteomics can provide context for the post-translational life and function of a protein, which are beyond the scope of information provided at the transcription level. This has made proteomics an increasingly valuable aid everything from therapeutic discovery to protein network mapping [5,6].

The protein interactions that reside within the intercellular secretory pathway are of particular interest here. During and immediately after translation, secreted proteins (SecPs) are subjected to a multitude of modifying enzymes and regulatory secretory pathway machinery components (SecMs) that reside primarily within the cell's endoplasm reticulum (ER), Golgi body, and trafficking vesicles [7]. Secreted proteins include cytokines, hormones, plasma proteins, extracellular components such as collagens, digestive enzymes, antibodies, and many more—any of which may have value as recombinant biotherapeutics or play an important role in a disease [8]. Furthermore, among the interacting SecMs that dictate the secretion of these protein there also exist potential therapeutic candidates and targets for cellular engineering.

While there are many SecMs known to reside within the mammalian secretory pathway, there is poor understanding of the distinct networks that control individual SecP processing and where the overlap exists across protein families. Using proximity dependent labeling coupled to LC-MS/MS, we hope to identify protein-protein interactions (PPIs) of various recombinant proteins of therapeutic value as they

move through the mammalian secretory pathway (Fig.1). We will answer questions of which PPIs control the success of protein secretion and influence overall protein titer.

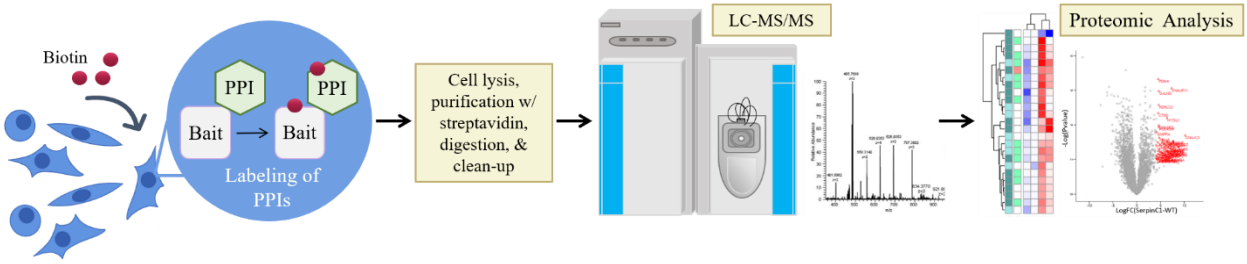

Figure 1. Workflow to detect in situ protein-protein interactions for proteins secretion.

The secretory pathway houses machinery pertinent to protein modification and secretion

The secretory pathway contains a complex repertoire of cellular processes for synthesizing and secreting host cell proteins, which the biotherapeutic industry has taken advantage of to produce recombinant proteins. Successful secretion is controlled by protein networks that perform the proper protein folding, post-translational modification (PTMs), and trafficking of SecPs. The precision and efficiency of the mammalian secretory pathway ultimately results from coordination among SecMs, such as chaperones, modifying enzymes such as protein disulfide-isomerases and glycosyltransferases, and transporters [8,9]. Overall, secretory pathway machinery components includes over 575 gene products that are tasked with the synthesis, quality control, and transporting of over 8000 secreted and membrane proteins [8–11].

At each step in a SecP's processing, SecMs assist with precise PPIs [7,12,13]. For example, SecMs include a verifiable army of enzymes that add a variety of post-translational modifications—anywhere from site-specific cleavages, to disulfide bond formation, to glycan attachments. While many SecMs serve a multitude of client SecPs, such as highly conserved chaperones like HSPA5, *a.k.a.* BiP, other SecMs assist only a few SecPs in a product-specific manner, either meeting their specific needs alone or with co-chaperones [14–18]. The diversity of proteins mammalian cells can secrete is accomplished by the similarly diverse yet versatile protein machinery that resides in the secretory pathway.

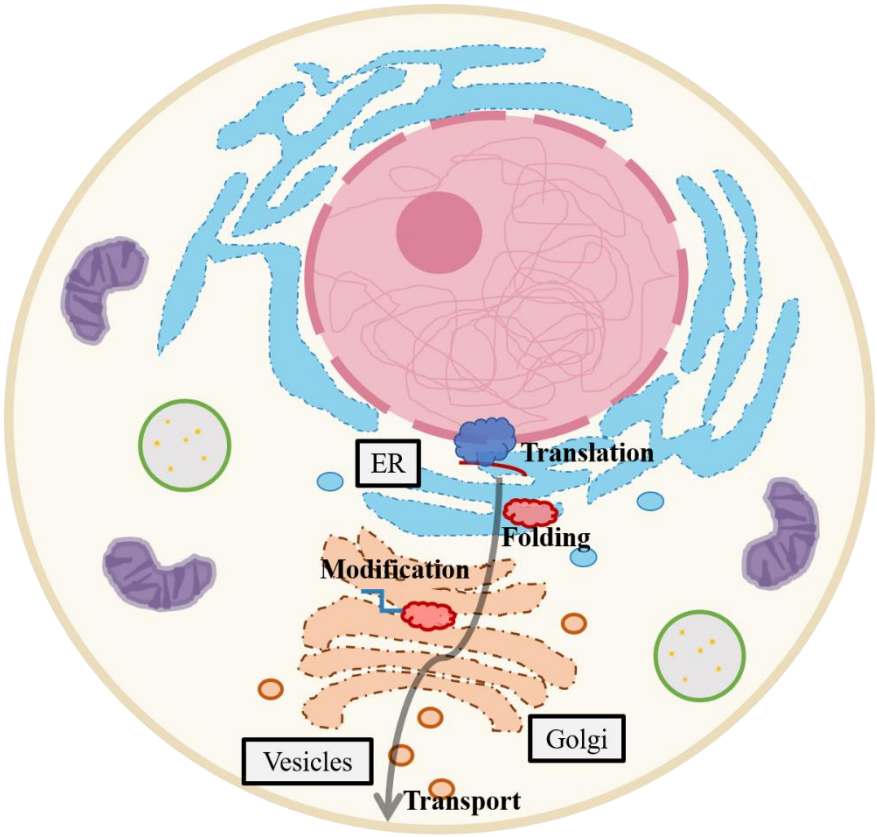

Figure 2. Model of the mammalian secretory pathway.

The secretory pathway is a common bottleneck in heterologous protein expression

The correlation between the human transcriptome and proteome is an ongoing area of scientific research. While transcription levels are largely consistent with native protein levels and are an invaluable determinant for steady-state protein levels, studies have found that a vast number of their MS captured proteins have uncorrelated levels of messenger RNA (mRNA) for yet to be understood reasons [2,4,19,20]. Transcriptomic assays such as RNA-seq are historically cheaper and, thanks to amplification, have greater sensitivity to lowly abundant molecules compared to protein assays; in many cases they remain standards for therapeutic/biomarker discovery and clone selection. However, to understand more about the discrepancies between transcription and observed protein levels, we must look at the PPIs that control secretion. Furthermore, recombinant proteins in a foregin host secretory pathway are subjected to more obstacles to proper secretion than native proteins. For many recombinant clones there may be adequate mRNA, but something within the secretory pathway is likely the rate limiting step to production [21].

As mentioned, major cell processes impacting recombinant protein secretion include transcription, protein folding, post-translational processing, and trafficking [22]. These are complex processes involving a multitude of components and feedback regulatory pathways, making maintaining homeostasis a priority. Recombinant protein expression incurs secretory pathway issues for several reasons. First, engineered or heterologous proteins have not evolved to optimize their contacts with the host-cell machinery, nor to eliminate problematic interactions. Second, heterologous protein expression is induced to levels far higher than physiological secretion levels [23,24]. Thus, even for popular host lines such as CHO cells, the secretory pathway is not always prepared to support vast amounts of recombinant proteins, leading to protein aggregation and cell stress that severely impairs protein quality and cell growth. Studies of difficult to express proteins have often shown an accumulation of the proteins in the ER, increasing stress response pathways [21,25]. To remedy these, researchers have targeted overwhelmed secretory pathway machinery, including vesicular transport and molecular chaperones, leading to increased titers of protein therapeutics [26–28]. However, hundreds of SecMs equip the secretory pathway, so it is often unclear which ones are deficient. Efficient synthesis of a recombinant protein drug requires one to meet its unique needs of SecM support; however, the necessary SecMs for any given recombinant protein drug are elusive. Therefore, we need to better understand the host cell machinery supporting high expression of diverse heterologous proteins.

Proximity biotinylation to quantify protein interactions in-situ

Each protein has its unique set of necessary SecMs, so identifying and balancing their expression is key to optimizing SecP secretion. Because many SecM-SecP interactions are transient and sensitive to the endomembrane system microenvironment, it is crucial to identify PPIs in the cell's native state, with intact subcellular organization and compartmentalization. However, many high-throughput PPI assays (e.g. yeast 2-hybrid or protein microarrays) are often done in heterologous systems and non-native compartments [29]. Other standard methods that maintain the native protein form such as co-immunoprecipitation assays are not sensitive to transient PPIs; Or, they disrupt PPIs and result in data loss as is common with some affinity purification methods [29,30]. Instead, a promising new class of methods chemically labels proteins proximal

to a target protein and quantifies the PPIs by mass spectrometry (MS). These methods overcome the aforementioned challenges by allowing the isolation of stable and transient interactions from the intact cell. Thus, these techniques could be used map and quantify SecM-SecP PPIs in the endomembrane system [31–33].

New proximity dependent labeling methods such as Proximity-dependent Biotin Identification (BioID) can identify weak, transient, and stable interactions in living cells [34,35]. BioID offers a high-throughput approach for systematic detection of intracellular PPIs occurring in various cellular compartments and has been used to characterize PPI networks and subcellular organization [30]. BioID relies on expressing a protein of interest fused to a promiscuous biotin ligase (BirA) that can biotinylate the proximal interactors in nanometer-scale labeling radius (Fig.3). For example, this approach has mapped protein interactions at human centrosomes and cilia, focal adhesions, and ER membrane-bound ribosomes [36–39].

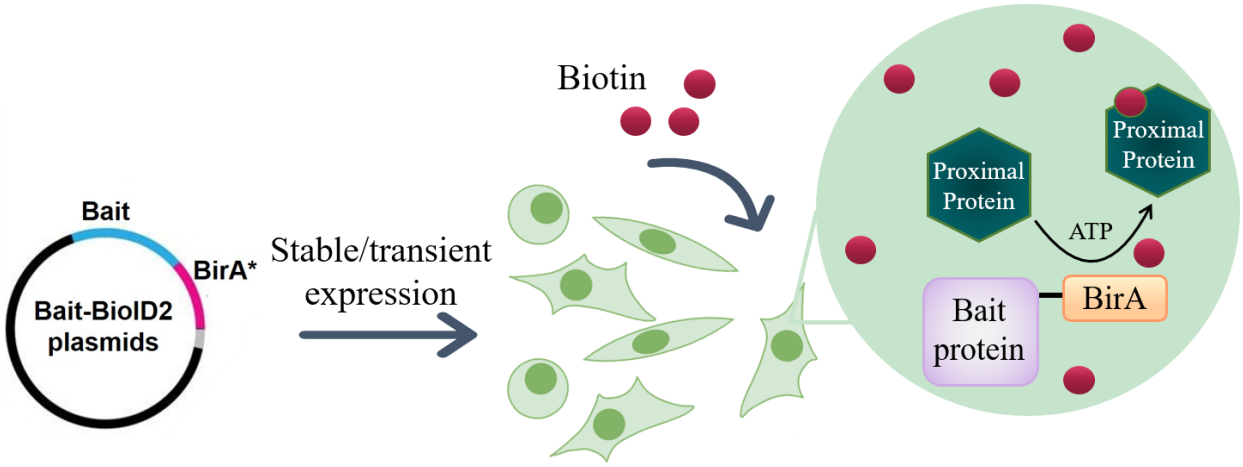

Figure 3. Schematic of Proximity-dependent Biotin Identification (BioID). Stable or transient expression of the fusion protein consisting of a chosen bait connected to a biotin ligase facilitates biotinylation of proximal proteins in the presence of excess biotin.

Another proximity-labeling method, biotinylation by antibody recognition (BAR), uses an antibody to target a protein of interest in fixed cells and removes the need for insertion of a fusion gene [40]. Labeling of proximal proteins with tyramide-biotin is driven by a secondary antibody conjugated with horseradish peroxidase (HRP) (Fig.4). Incubation with hydrogen peroxide creates an oxidizing environment that catalyzes the reaction between tyramide-biotin and HRP, producing an active form of tyramide that can

attach to the lysine residues of proximal proteins, the same reaction that is used in standard Tyramide Signal Amplification (TSA) systems. Antibody recognition allows PPI identification for almost any protein in most cell or tissue types. Thus, this technique could comprehensively identify the SecM-SecP PPIs and quantify how they change under different conditions.

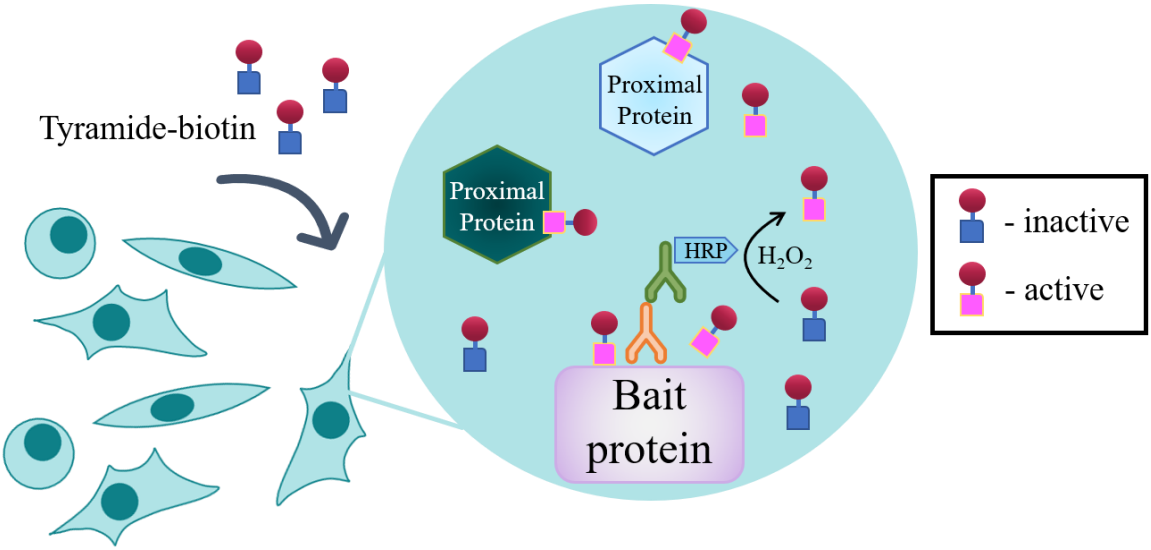

Figure 4. Schematic of Biotinylation by Antibody Recognition (BAR). Using an HRP conjugated antibody to target proteins of interest, we can biotinylate proximal proteins with the addition of tyramide-biotin and hydrogen peroxide. The active form of the tyramide-biotin will covalently attach to the lysine residues of nearby proteins. The reaction is quench by adding a reducing agent (i.e. sodium ascorbate).

In both the case of BioID and BAR, proximity-based labeling is achieved using biotin. Biotin has been shown to be ideal for protein purification due to its high affinity for avidin/streptavidin and its column extraction not disrupting PPIs [30]. This high affinity and limited endogenous intercellular biotin-binding complexes also allows for specificity and low background. As such, we investigate here the ability of these two systems to aid our exploration of the mammalian secretory pathway.

Acknowledgments

The Introduction contains some material from a manuscript currently in review for publication “In situ detection of protein interactions for recombinant therapeutic enzymes”. Samoudi, Mojtaba; Kuo, Chih-Chung; Robinson, Caressa M.; Shams-Ud-Doha, Km; Schinn, Song-Min; Kol, Stefan; Weiss, Linus; Bjorn, Sara Petersen; Voldborg, Bjorn G.; Campos, Alexandre Rosa; Lewis, Nathan E.. (2020). *In review*. The thesis author is a primary author of this material.

Chapter 1

BioID facilitated identification of the recombinant enzyme PPIs required for secretion in HEK cells

Abstract

Despite their therapeutic potential, many protein drugs remain inaccessible to patients because they are difficult to secrete from current cell hosts. Each recombinant protein has unique physicochemical properties and requires different machinery for proper folding, assembly, and post-translational modifications (PTMs). Targeted engineering of these components for better protein secretion requires specific knowledge of secretory pathway interactions for each recombinant protein. Here, we aimed to identify the protein-protein interaction (PPI) networks of four different recombinant proteins (SERPINA1, SERPINC1, SERPING1 and SEAP) with various PTMs and structural motifs using proximity-dependent biotin identification (BioID). This method allowed us to target protein interactions of each secreted protein as they are secreted from HEK293 cells. We found proteins involved in protein folding, disulfide bond formation and N-glycosylation were positively correlated with the corresponding features of the four model proteins. Knock-down of captured interacting PDIs PDIA4 and PDIA6 led to the decreased secretion of SERPINC1, which relies on its extensive disulfide bonds, compared to SERPINA1, which does not. Proximity-dependent labeling successfully identified the transient interactions supporting synthesis of secreted recombinant proteins and refined our understanding of key molecular mechanisms of the secretory pathway during recombinant protein production.

1.1 Introduction

Serine protease inhibitors, or SERPINs, are rising as a new class of biotherapeutics to treat autoimmune and inflammatory diseases, including diseases defined by the disruption of SERPIN regulation [41]. SERPINs are relatively large, complex proteins compared to other protease inhibitors that undergo an extensive functional conformational change [42]. While structurally related, the differing PTMs and functional conformations among members of the SERPIN superfamily give rise to their distinct functions and substrate interactions. Heterologous systems unable to replicate these distinct structural features result in functionally inactive SERPIN products [41]. As complexities associated with the mammalian secretory machinery remain a bottleneck in recombinant protein production and overexpression of heterologous proteins can result in adaptive responses that can impair both protein quantity and quality, optimized manufacturability of SERPINs and other recombinant enzymes may be helped by customizing secretion machinery for these products.

A previous study showed that human protein secretory pathway genes are expressed in a tissue-specific pattern as needed to support the diversity of modifications needed, suggesting that expression of some SecMs is regulated to support specific client SecPs [8]. Unfortunately, which SecMs are needed to support any specific secreted protein remains unclear. Thus, there is a need to elucidate the SecMs that support the expression of different recombinant proteins with specific features. This can guide mammalian synthetic biology efforts to engineer enhanced cells capable of expressing proteins of different kinds in a client-specific manner.

Here, BioID2, an improved smaller biotin ligase for BioID, was used to explore how SecM interactions vary for different secreted therapeutic proteins [30,33]. Specifically, BioID2 was employed to identify SecMs that interact with four recombinant enzymes : three SERPIN-family proteins—SERPINA1, used for treatment for Alpha-1-antitrypsin deficiency; SERPINC1, used for treatment for hereditary antithrombin deficiency; and SERPING1, used for treatment for acute attacks of hereditary angioedema—and secreted embryonic alkaline phosphatase (SEAP), which is a truncated form of Alkaline Phosphatase, Placental Type (ALPP) [42]. These proteins vary in their PTMs (e.g., glycosylation, disulfide bond formation,

and residue modifications) and have different amino acid sequences that consequently form different local motifs. Comparing the interactions captured by streptavidin pull-down and subsequent MS analysis against WT control, we will identify SecMs pertinent for each specific proteins' secretion. Identification of these PPIs will improve our understanding of how the secretory pathway functions during the expression of the recombinant proteins and introduce novel targets for secretory pathway engineering in a client specific manner.

1.2 Methods

1.2.1 Molecular cloning and generation of stable cell lines

All plasmids used in this study were constructed by PCR and Gibson isothermal assembly. The expression ORFs, hereafter named bait-BirA, were constructed by fusing BioID2 to the C-terminal of each model protein (with a glycine-serine linker added between) and a 3XFLAG tag at C-terminal to simplify the immune-detection. ORFs were inserted into pcDNA5/FRT (Invitrogen), which allows targeted integration of the transgenes into the host genome. Gibson assembly primers were designed by SnapGene software and used to amplify the corresponding fragments and vectors with long overlapping overhangs, which were then assembled using Gibson Assembly Master Mix (NEB). To obtain secretable BioID2 (without any bait protein), Gibson assembly was employed to fuse the signal peptide of SERPINC1 gene to the N-terminal of BirA (hereafter referred to as Signal-BirA). Assembly products were transformed to the chemically competent *E. coli*, and recombinant plasmids were verified by restriction digestion and sequencing. For all experiments, Flp-In 293 cells (Invitrogen) were cultured in DMEM media supplemented with fetal bovine serum (10 %) and antibiotics (penicillin, 100 U mL−1 and streptomycin, 100 μg mL−1) and maintained at 37 °C under 5 % CO_2. For generating stable cell lines, Flp-In 293 cells were seeded in 6 well plates at a density of 0.5×10^6 cells per well the day before transfection. Cells were then co-transfected with each pcDNA5/FRT vector containing expression cassette and pOG44 plasmid using Lipofectamine® 2000 according to the manufacturer's directions. After recovery from transfection, cells were grown in DMEM containing 10% FBS, 1% PenStrep, and 150 μg/mL Hygromycin B to select hygromycin-resistant

cells. Individual resistant colonies were isolated, pooled, and seeded in 24-well plates for further scaling up and screened for expression of the fusion proteins by Western Blotting.

1.2.2 Immunofluorescence

Recombinant HEK293 cells expressing BioID2 fusions were grown in complete medium supplemented with 50 μM biotin on coverslips until 70% confluent. Cells were then fixed in PBS containing 4% PFA for 10 min at room temperature. Blocking was performed by incubating fixed cells with 1% BSA and 5% normal goat serum in PBST. Anti-flag mouse monoclonal antibody-DyLight 650 conjugate (Thermofisher), targeting the bait-BirA, and streptavidin-DyLight 594 conjugate (Thermofisher), targeting the biotinylated proteins, were diluted at 1:300 and 1:1000 in blocking buffer, respectively and incubated with fixed cells for 30 minutes at room temperature. Cells were then washed, counterstained with DAPI, mounted on the slide using antifade vectashield mountant, and imaged using Leica SP8 Confocal with Lightning Deconvolution. Colocalization quantification was performed for the deconvolved images using Fiji's (ImageJ 1.52p) Coloc_2 analysis tool between the 650 (anti-flag) and 594 (Streptavidin) channels [43]. This tool generates a report for evaluating pixel intensity colocalization of two channels by various methods such as Pearson's Coefficient (range: -1.0 to 1.0), Manders' Colocalization Coefficients (MCC, range: 0 to 1.0), and Li's Intensity Correlation Quotient (IQC, range: -0.5 to 0.5) [44,45]. Background pixel intensity was subtracted using Fiji's rolling ball algorithm and a region of interest (ROI). Thresholds were determined using Coloc_2's bisection method, which is further used to adjust for background. Above threshold metrics were reported.

1.2.3 RNAi knockdown experiment

esiRNA targeting PDIA4 and PDIA6 were ordered from Sigma Aldrich. HEK293 cell expressing SERPINC1-BirA and SERPINA1-BirA were seeded at $0.6X10^5$ cells/well in 24-well plates with complete medium and reverse transfected with 72 ng of the appropriate PDI specific esiRNA or Luciferase esiRNA as a negative control or KIF11 esiRNA as positive control (for optimizing) using Lipofectamine RNAiMAX (Invitrogen). All transfections were performed according to the manufacturer's guidelines. Each experiment was done in triplicates and targeted gene knockdown by esiRNA was done for 96 hrs.. Culture supernatants

and cell pellets were harvested, clarified by low-speed centrifugation, then aliquoted and stored at −80°C for further experiments.

1.2.4 Western blotting

To validate the secretion of bait-BirA proteins, supernatants of cultures expressing fusion proteins were collected, centrifuged to remove cell debris, and protein content was concentrated using Amicon Ultra 4 mL 10 kDa filter unit (MilliporeSigma) and quantified using the Bradford assay (Expedeon). 20 μg of total protein was loaded on SDS-PAGE gel for electrophoresis and resolved proteins were transblotted to nitrocellulose membrane using Trans-Blot Turbo Transfer System from Bio-RAD. The membrane was blocked with 5% skim milk in TBST and probed with HRP-conjugated anti-flag mouse monoclonal antibody (Thermofisher) diluted at 1:10000 in the blocking buffer. The membrane was washed, and Clarity Western ECL Substrate was added. Protein bands were visualized using G:Box Gel Image Analysis Systems (SYNGENE). For staining of intracellular biotinylated proteins, cells were grown in complete medium supplemented with 50 μM biotin, lysed by RIPA buffer, and protein content was quantified using the Bradford assay. 20 μg of total protein was loaded and resolved and transblotted as described earlier. The membrane was blocked by 3% BSA in TBST and probed with HRP-conjugated streptavidin diluted in blocking buffer at 1:2000 ratio. For visualizing the proteins' bands, the same Clarity Western ECL Substrate was used.

Semi-quantitative western-blots were used to determine knock-down (KD) efficiency as well as impact on SERPIN secretion. For KD efficiency, cell pellets from each KD experiment were lysed with RIPA buffer and approximately 25μg of protein lysate was loaded onto the appropriate SDS-PAGE gel for PDIA4 and PDAI6 experiments. Resolved proteins for each gel were transblotted onto separate nitrocellulose membranes using the Trans-Blot Turbo Transfer System from Bio-Rad. Each membrane was blocked with Intercept (TBS) Blocking Buffer from LI-COR. Gels were probed with Rabbit αPDIA4 (1:2000) or αPDIA6 (1:2000) monoclonal antibody, respectively. We also targeted a housekeeping protein in each lysate for normalization with either a Mouse αAlpha-tubulin (1:20,000) or Mouse αBeta-actin (1:10,000) monoclonal antibody. For visualization, IR spectra was utilized by staining with LI-COR Goat

αMouse conjugated to 680 (1:20,000) and Goat αRabbit conjugated to 800 (1:20,000). Images were taken and analyzed with Image Studio Lite Version 5.2 for relative quantification. Impact on SERPIN secretion was determined from the aliquots of clarified supernatants from esiRNA transfected cultures using the same blotting method as above. Supernatants were loaded at a 2X dilution with sample buffer. The transblotted membrane was probed using Rabbit αFlag (Proteintech, 1:800) and then with Goat αRabbit 800 (1:20,000). KD efficiency and the effect of PDI knockdown on secretion of the model proteins was measured in comparison with the negative control of each cell line transfected with Luciferase esiRNA as negative control (see above).

1.2.5 Mass Spectrometry

Cells were grown in 245 mm plates (one plate per biological replicate in triplicate) to approximately 70% confluence in complete media and then incubated for 24 hrs. with 50 μM biotin. Cells were harvested and washed twice in cold PBS, lysed with vigorous shaking (20 Hz) in 8M urea, 50mM ammonium bicarbonate lysis buffer, extracted proteins were centrifuged at 14,000 x g to remove cellular debris and quantified by BCA assay (Thermo Scientific) as per manufacturer recommendations. Affinity purification of biotinylated proteins was carried out in a Bravo AssayMap platform (Agilent) using AssayMap streptavidin cartridges (Agilent), and the bound proteins were subjected to on-cartridge digestion with mass spec grade Trypsin/Lys-C Rapid digestion enzyme (Promega, Madison, WI) at 70ºC for 2h. Digested peptides were then desalted in the Bravo platform using AssayMap C18 cartridges and the organic solvent was removed in a SpeedVac concentrator prior to LC-MS/MS analysis. Dried peptides were reconstituted with 2% acetonitrile, 0.1% formic acid, and analyzed by LC-MS/MS using a Proxeon EASY nanoLC system (Thermo Fisher Scientific) coupled to a Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific). Peptides were separated using an analytical C18 Acclaim PepMap column 0.075 x 500 mm, 2μm particles (Thermo Scientific) in a 93-min linear gradient of 2-28% solvent B at a flow rate of 300nL/min. The mass spectrometer was operated in positive data-dependent acquisition mode. MS1 spectra were measured with a resolution of 70,000, an AGC target of 1e6 and a mass range from 350 to 1700 m/z. Up to 12 MS2 spectra per duty cycle were triggered, fragmented by HCD, and acquired with a resolution

of 17,500 and an AGC target of 5e4, an isolation window of 1.6 m/z and a normalized collision energy of 25. Dynamic exclusion was enabled with a duration of 20 sec.

1.2.6 MS Data Analysis

All mass spectra were analyzed with MaxQuant software version 1.5.5.1 [46]. MS/MS spectra were searched against the Homo sapiens Uniprot protein sequence database (version January 2018) and GPM cRAP sequences (commonly known protein contaminants). Precursor mass tolerance was set to 20ppm and 4.5ppm for the first search where initial mass recalibration was completed and for the main search, respectively. Product ions were searched with a mass tolerance 0.5 Da. The maximum precursor ion charge state used for searching was 7. Carbamidomethylation of cysteines was searched as a fixed modification, while oxidation of methionines and acetylation of protein N-terminal were searched as variable modifications. Enzyme digestion was set to trypsin in a specific mode and a maximum of two missed cleavages was allowed for searching. The target-decoy-based false discovery rate (FDR) filter for spectrum and protein identification was set to 1%. Enrichment of proteins in streptavidin affinity purifications were calculated as the ratio of intensity.

1.2.7 Detection of significant interactions and pathway analysis

To remove the systematic biases introduced during various steps of sample processing and data generation, dataset were normalized using the LOESS method integrated into Normalyzer [47,48]. Perseus software was employed for data preparation, filtering, and computation of differential protein abundance [49]. The DEP package was used to explore whether missing values in the dataset are biased to lower intense proteins. Left-censored imputation was performed using random draws from a shifted distribution in Perseus [50,51]. A student's t-test with a multi-sample permutation-based correction for an FDR of 0.05 was employed to identify differentially expressed proteins using log2 transformed data. Lastly, only interactors with positive fold changes with the WT were considered, as this optimizes the enrichment of known secretory pathway components among the significant interactors. Finally, biological process and cellular component analysis was done using Gene Ontology (GO) Enrichment Analysis tool release date “2019-01-01” and version number “10.5281/zenodo.2529950” [52–54]. After DE analysis, a Bayesian model framework

developed by Chih-Chung Kuo of our lab was applied to quantify associations of each interacting protein with protein features. This framework is not discussed in detail here, but the results of this analysis were the selection of the highlighted proteins presented and gave rise to the conclusions drawn here. For this thesis, however, we will only be discussing results as they are given in context of the DE analysis.

1.3 Results

1.3.1 Successful biotinylation of bait-specific proteins can be achieved with BioID

The BioID method was applied to the four chosen SERPRINs and SEAP bait-BirA proteins to validate if intracellular PPIs between each SecP and their supporting SecMs can be labeled and subsequently analyzed. First, we observed successful secretion of bait-BirA proteins into culture supernatant, evaluated by Western blot (Fig. 1.1a). This indicates that the addition of the BirA fusion did not prevent secretion of the model proteins, and therefore their entry into the secretory pathway. Following biotin incubation by way of the BioID methodology, we also verified the biotinylation profile by western blot for each cell line in the presence and absence of biotin. The biotinylation profile of the bait-BirA cells with biotin incubation show a substantial increase in biotinylation of various proteins, while no obvious change is observed for WT (Fig. 1.1b), suggesting that BioID2 successfully tagged proteins in a bait-specific manner within the cells. However, endogenously biotinylated proteins are also identified in all samples, WT included. These include proteins that bind biotin as a cofactor such as carboxylases.

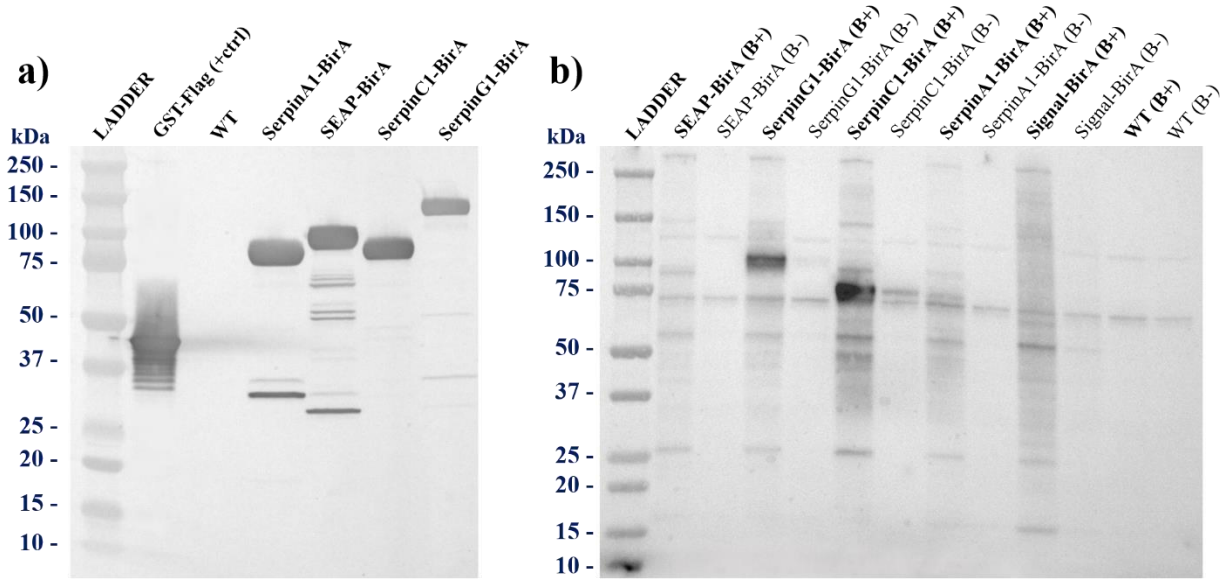

Figure 1.1. Expression of bait-BirA proteins results in a substantial increase in biotinylated proteins. a) Successful secretion of the bait-BirA proteins into the culture supernatant was evaluated by western blot using HRP-anti-flag antibody. b) Biotinylation profiling of WT control and HEK293 cells expressing the model proteins lysates with HRP-streptavidin. BirA domain fused to the model proteins and biotin addition (B+) led to the biotinylation of a subset of proteins which are not seen in WT or absence of biotin (B-). A few endogenously biotinylated proteins appear in all lanes at around 74kDa and 125 kDa. These bands are consistent with known carboxylases that bind biotin as a cofactor.

Colocalization of the bait-BirA proteins and the biotinylated proteins was then visualized by multicolor co-immunofluorescence microscopy to investigate whether biotinylated proteins were indeed proximal-partners of the model proteins. The results demonstrated successful labeling of the interactors by BirA through colocalization of the biotin-labeled proteins and bait-BirA, while WT did not show increased biotinylation under the same experimental conditions (Fig. 1.2.). To quantify the colocalization between biotin and model proteins, we calculated various colocalization metrics from each image and compared to the WT (Table 1.1 and Fig. 1.3.). The Pearson's and Manders' colocalization coefficients are used to determine the degree of overlap between images and are mathematically similar; however, the Pearson's coefficient utilizes deviation from the average intensity of each channel, whereas the MCC are each relative to the total fluorescence in each channel [45]. The MCC are said to allow for strong differences between channel signal intensities due to labeling or settings during/after image capture, however, it is more susceptible to offset and background where the Person's value is not. Li's IQC is also derived from Pearson's coefficient, but further aids in determining the dependency of staining intensities, i.e. pixel

channel intensities vary around mean channel intensities together when there is synchrony of two proteins in a complex [44]. The results corroborated the specificity of the BioID labeling system—i.e. Li's ICQ value closer to 0.5 and Pearson's R value and Manders' Colocalization Coefficients closer to one, demonstrating a dependent protein staining pattern between the red and green channels.

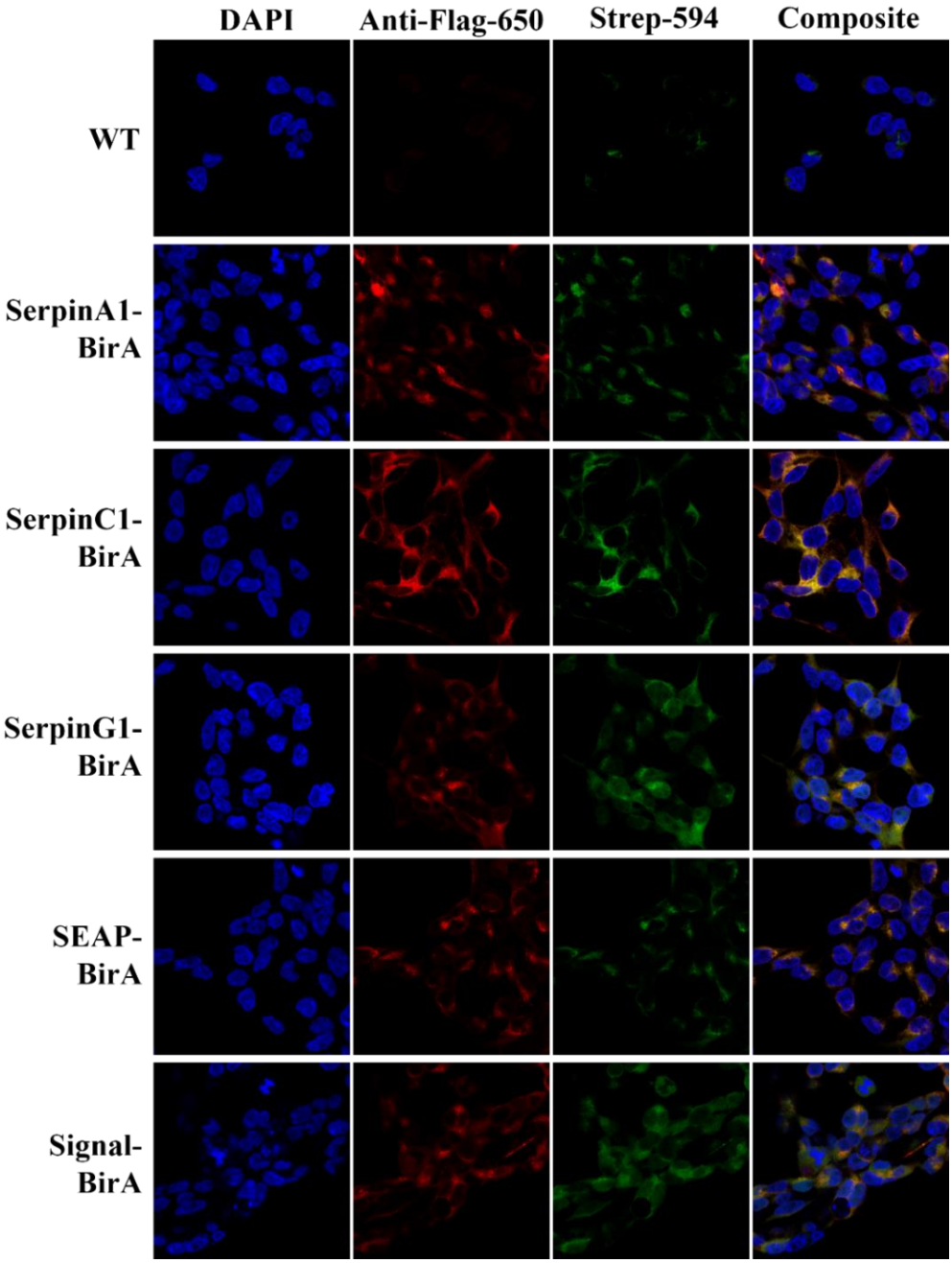

Figure 1.2. Bait-BirA fusion proteins are colocalized with biotin-staining. Co-Immunofluorescence demonstrated the intracellular colocalization of the biotin-labeled proteins (stained with Streptavidin-Dylight 594 and illustrated in green color) and bait-BirA (stained with anti-flag monoclonal antibody-Dylight 650 and illustrated in red color), while WT did not show increased biotinylation under the same experimental conditions.

Table 1.1. Colocalization metrics determined for each clone between the 650 (anti-flag) and 594 (Streptavidin) channels in Fig. 1.1. There is little correlation reported for the WT, whereas there is much higher correlation for clones containing BirA.

Cells	Pearson's R value	Li's ICQ value	Manders' Coeff. (tM_g,tM_r)
WT	0.22	0.219	0.812, 0.518
SERPINA1-BirA	0.74	0.318	0.962, 0.888
SERPING1-BirA	0.73	0.260	0.999, 0.914
SERPINC1-BirA	0.88	0.392	0.991, 0.982
SEAP-BirA	0.93	0.392	0.952, 0.935
Signal-BirA	0.71	0.308	0.841, 0.791

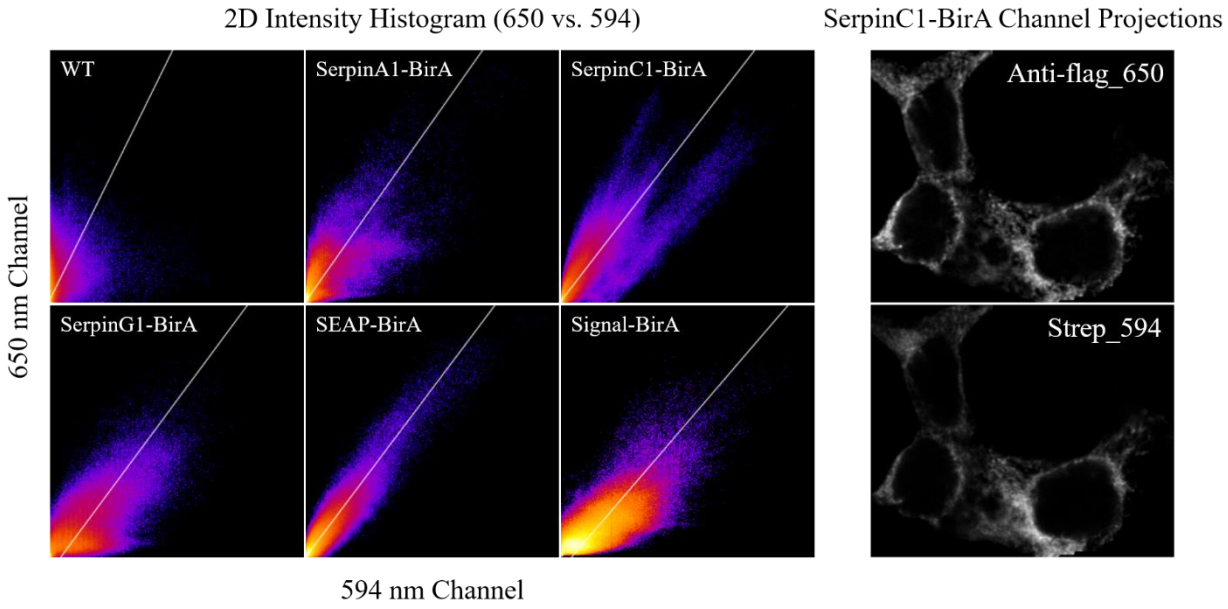

Figure 1.3. Colocalization between the 650 (Anti-flag) and 594 (Streptavidin) channels. Max projections of anti-flag_650 and streptavidin_594 fluorescent channels, such as the example one shown above for SERPINC1-BirA, were used for quantification. Corresponding intensity histogram for one of various regions tested for each clone are shown. These intensity plots were created using the same ImageJ Coloc2 tools as the colocalization metrics in Table 1. The relative linearity of the intensity graph demonstrates colocalization for the clones expressing the BirA fusions and biotin labeling for each recombinant clone compared to the WT.

1.3.2 Data analysis shows enrichment for secretory pathway machinery the chosen SecPs

Analysis of MS data with MaxQuant found 5931 total proteins captured. After data processing in Perseus, principal component analysis (PCA) was performed using R package *pcaMedhods*' probabilistic PCA (PPCA) method. PCA revealed good separation of conditions and clustering of biologic and technical replicates (Fig. 1.4.).

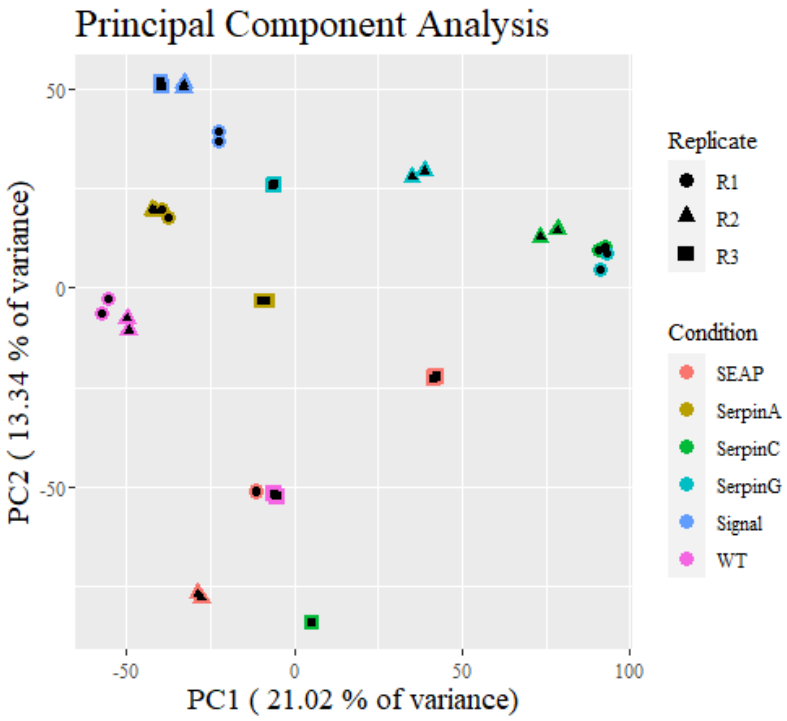

Figure 1.4. PCA for all captured proteins and samples. Points are colored by bait protein and shape correlates to technical replicate. There are three biological replicates for each condition.

In addition, Gene Ontology (GO) term enrichment was performed on the top 150 loadings in the first principal component, using an FDR <0.05 to identify significant terms. This indicated that the major biological processes accounting for variance among samples belong primarily to processes within the secretory pathway, i.e. the unfolded protein response (UPR), ERAD, protein folding, and protein modification (Fig. 1.5.).

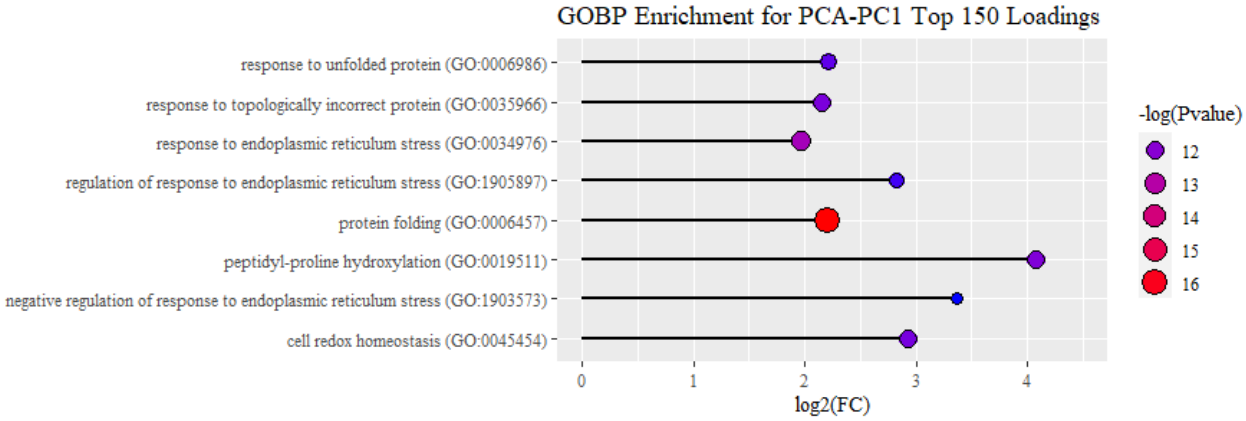

Figure 1.5. GO term enrichment results for the top 150 loadings in the first principal component. Biological processes contributing to variance appear to reside within the secretory pathway.

Differential expression of proteins was determined using a student's t-test for each bait protein against the WT (Fig.1.6.). Significant proteins were selected using an FDR < 0.05 and an s0 of 2. The corresponding bait protein for each construct is easily distinguishable due to high positive fold change in each sample, as expected since each bait protein is most susceptible to large amounts of biotin deposition.

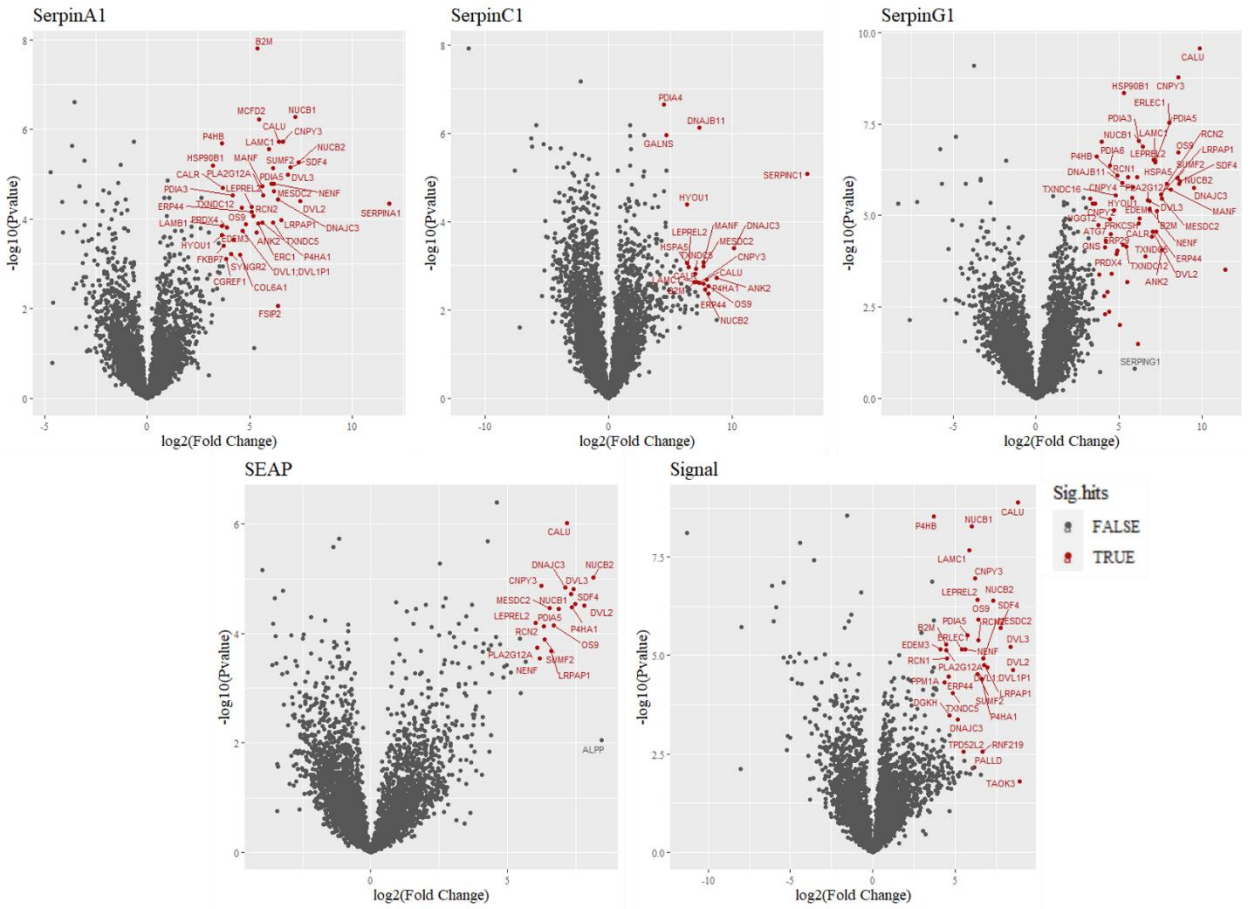

Figure 1.6 Volcano Plots showing the significant proteins (red) identified between each bait protein and WT enrichment. Proteins with particularly low p-value and/or large fold change are also labeled with their gene ID. As expected, SERPINA1, SERPINC1, SERPING1, and SEAP (ALPP) were all identified to have a small p-value and/or have a large FC in each of their corresponding samples.

Gene Ontology term enrichment for cellular compartments (CC) and biological pathways (BP) using all significant proteins identified among SERPINs and SEAP, except for those also identified as significant in Signal-BirA vs. WT, shows GO terms associated with ER, Golgi, and vesicle compartments and processes are significantly overrepresented (Fig. 1.7). Significant proteins found in the Signal-BirA vs. WT analysis were excluded to avoid looking at interaction effects that could be due only to the addition of

the biotin ligase and linker. There was a final total of 42 significant proteins pulled from all SERPINs and SEAP that did not have overlap in Signal significant hits.

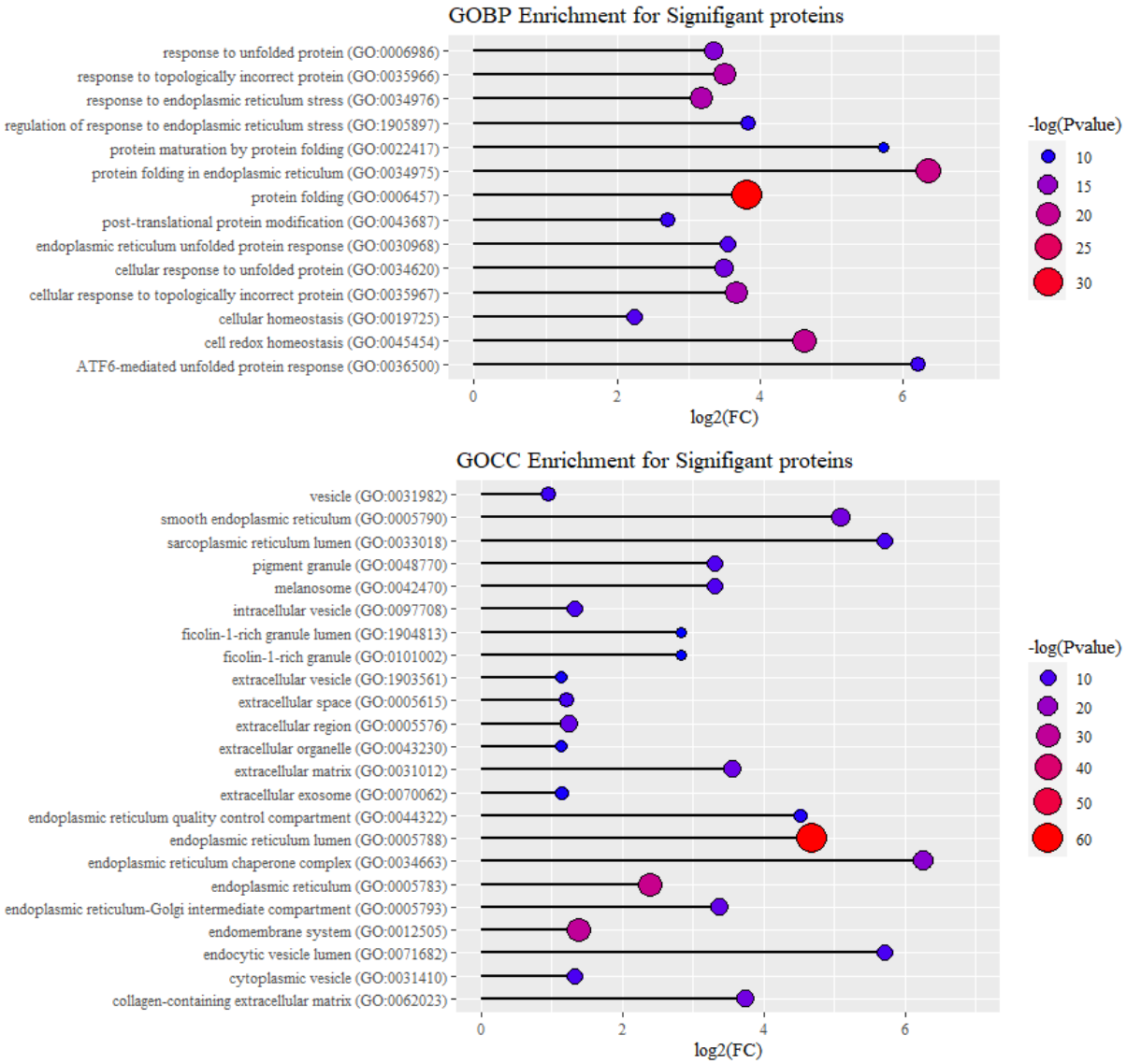

Figure 1.7. GO term enrichment results for all significant proteins identified for both biological function (BP) and cellular components (CC) annotations. Significant hits from SERPINS and SEAP associated with GO terms that indicate secretory pathway localization and SecM processes.

1.3.3 Significant interactors reflect post-translational and structural features of model proteins

The significant interactions captured between the bait-BirA proteins link secreted proteins to the responsible SecMs. As the bait-BirA proteins differ in structural composition and post-translational modifications (Fig. 1.8.), we explore if bait-BirA proteins with shared features have higher affinity for specific interactors. More specifically, we hypothesize that proteins requiring a specific PTM would

preferentially interact with the secretory machinery components responsible for the PTM synthesis. Disulfide bond and N-linked glycosylation features noticeably differ among bait proteins; Furthermore, a multitude of PDIs and glycosylation associated SecMs were captured. As such, we subsequently chose to analyze these interactions between the bait-BirA..

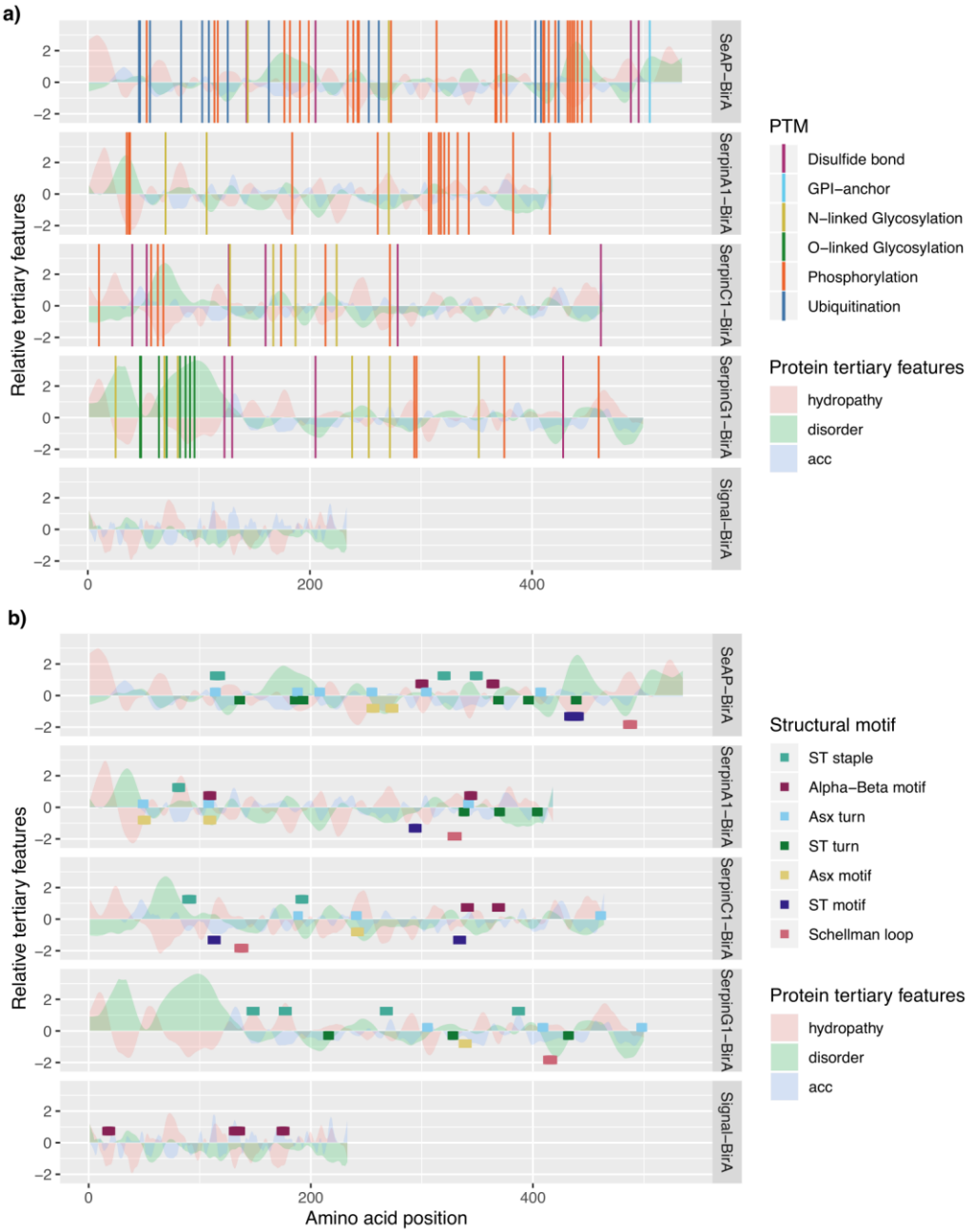

Figure 1.8 The bait proteins show diversity in their PTM and structural content. Lines and blocks represent known PTM sites and structural motifs in panel (a) and (b) respectively. The hills and valleys indicate protein tertiary features. Note that solvent accessibility and structural motifs are only available for regions covered by the PDB structure, whereas predicted features such as protein hydrophobicity and disorder are available for the entire protein.

From Figure 1.8., we see that SERPING1 and SERPINC1 have more disulfide bond and N-linked glycosylation features than SERPINA1 and SEAP. Among significant interactors, we see high expression of members of the Calnexin/Calreticulin cycle and related processes for more heavily glycosylated proteins (Fig. 1.9.). For SERPING1, SERPINC1, and SERPINA1, we see high interaction with calreticulin (CALR), a calcium-binding chaperone that promotes folding, oligomeric assembly, and quality control of N-linked glycoproteins in the ER [55,56]. For SERPING1, we also see significance for UGGT1, which recognizes glycoproteins with minor folding defects and re-glucosylates single N-glycans near the misfolded part of the protein. Re-glucosylated proteins are then recognized by CALR for recycling to the ER and refolding or degradation [57]. In addition, two members of the PDI family, PDIA3 and ERp29, which form a complex with calreticulin/calnexin, showed association with SERPINA1 and/or SERPING1. Calnexin/Calreticulin-PDIA3 complexes promote the oxidative folding of nascent polypeptides and ERp29 promotes isomerization of peptidyl-prolyl bonds to attain the native polypeptide structure [58,59].

SERPING1 and SERPINC1 also significantly interacted with chaperone HSPA5, a component of the glycoprotein quality-control (GQC) system (Fig. 1.9). GQC recognizes glycoproteins with amino acid substitutions, and targets them for ER-associated degradation (ERAD) [60]. There is also high interaction with SIL1, which helps ATP facilitated HSPA5 by removing ADP to allow for further activity, and/or DNAJB11, a co-chaperone with HSPA5, for SERPINC1 and SERPING1. Given that most of these molecular chaperones and enzymes are involved in ERAD mediated degradation of the misfolded glycoproteins, these findings suggest the quality control pathways are critical for synthesizing and secreting proteins with N-linked glycans.

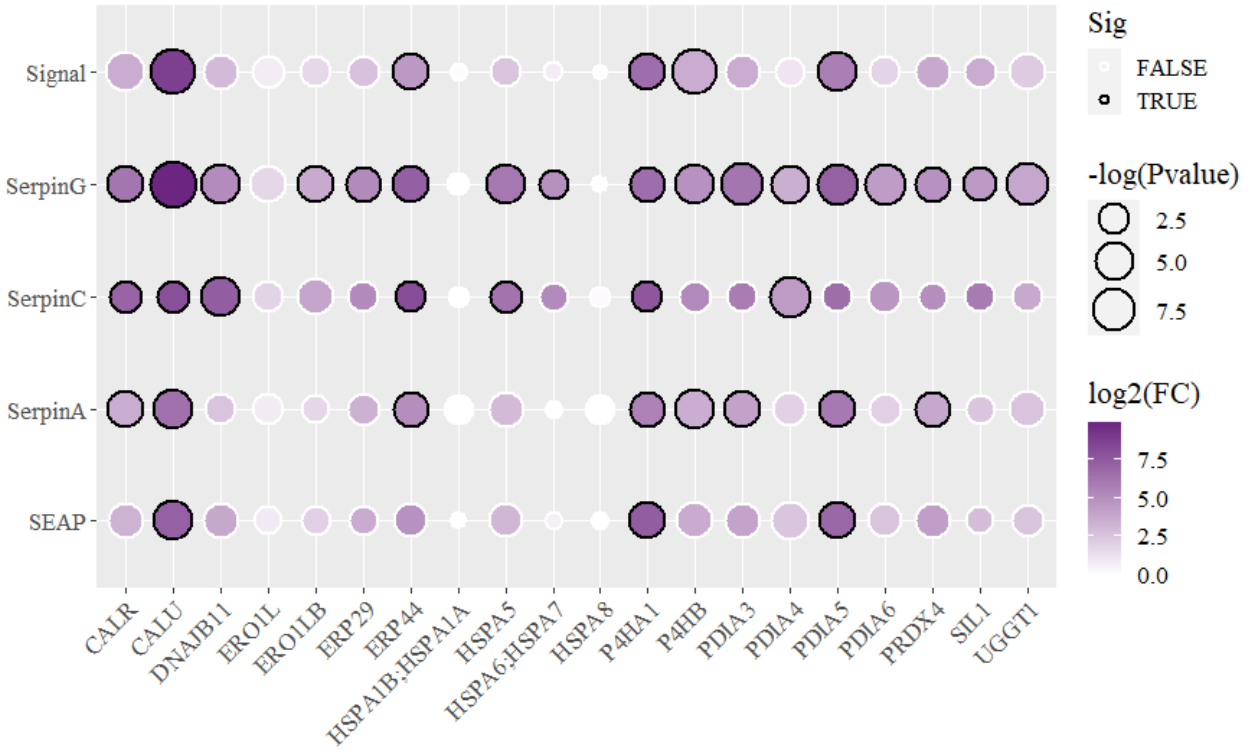

Figure 1.9. Visualization of select DE results. Several proteins from various PDI, molecular chaperones, and other protein families were selected from the significant interactors to demonstrate possible public versus more private interactions. Motifs previously selected from a Bayesian model analysis done by our lab (see methods).

Several members of the PDI family—including PDIA3, PDIA4 and PDIA6—significantly interacted with model-BirAs (Fig. 1.9). These enzymes catalyze the formation, breakage and rearrangement of disulfide bonds through thioredoxin-like domains [61]. The identification of various PDIs highlights the importance of the oxidative folding enzymes in protein folding and maintaining stability that can limit the efficiency of protein secretion. We see that SERPING1 has a consistently lower P-value and large fold-change for each of the PDIs selected, and looking only at PDIs that are not significantly interacting with Signal, we see that SERPINC1 and SERPING1 interact more with PDIA4 and PDIA6 than the other bait proteins.

Furthermore, the PDI ERp44 showed the strongest association (LFC > 8) with disulfide bond enriched proteins i.e. SERPINC1 and SERPING1. ERp44 mediates the ER localization of the oxidoreductase Ero1α (an oxidoreductin that reoxidizes P4HB to enable additional rounds of disulfide formation) through the formation of reversible mixed disulfides [62]. Hence, the strong association of ERp44

suggests the importance of the thiol-mediated retention in disulfide bond formation, particularly when secretory is loaded with the proteins with dominant disulfide bond. In addition, ERO1LB is an ER-localized enzyme associated with disulfide bond formation. ERO1LB efficiently reoxidizes P4HB, another protein which facilitates the formation and breakage of disulfide bonds [63]. Present as well is PRDX4, which couples hydrogen peroxide catabolism with oxidative protein folding by reducing hydrogen peroxide [64]. This oxidoreductase enzymes may highlight the importance of ER redox homeostasis in disulfide bond formation and protecting cells from the consequences of misfolded proteins.

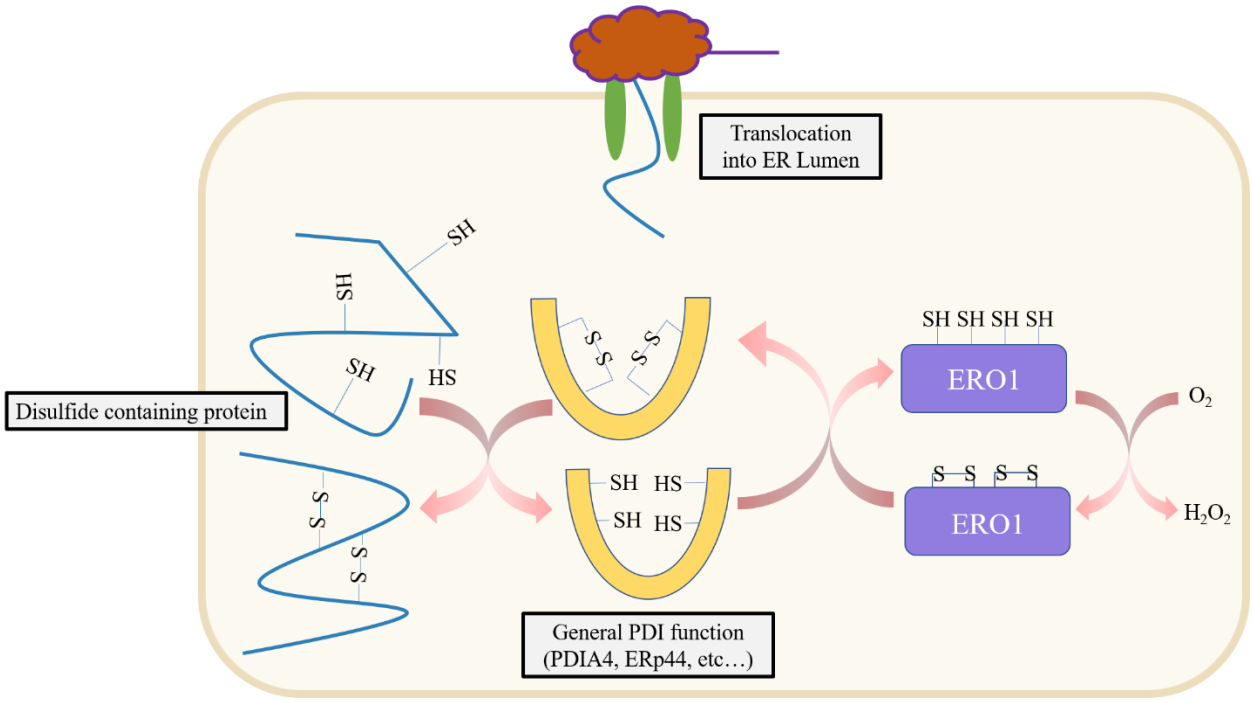

Figure 1.10. Modification of bait proteins requiring isomerase-facilitated disulfide-bond formation[65].

1.3.4 Knockdowns of significant SecMs show impact on SecP secretion

We performed knockdown of select isomerases that showed significant interaction with our bait proteins, then quantified relative SERPIN secretion to analyze knockdown effects of various SecMs on protein production. SERPINC1 has many disulfide modifications, while SERPINA1 does not. Fig. 1.9 indicates that the isomerase PDIA4 significantly interacts more with SERPINC1 than SERPINA1. We therefore compare SERPINA1 and SERPINC1 secretion changes when select PDIs are reduced, as SERPINC1 secretion is predicted to be impacted more than SREPRINA1 for PDIA4. PDIA6 which did not

show significant interaction with SERPINA1 nor SERPINC1 was used to compare. With PDIA4 KD, there is greater reduction in SERPINC1 secretion and a lesser impact on SERPINA1 secretion, while KD of PDIA6 did not produce the same SERPIN reduction (Fig. 1.11). To account for differences in KD efficacy, assuming greater KD efficiency might result in a greater SERPIN reduction, we can normalize the effects of PDI KD on SERPIN secretion using the KD efficiencies calculated. If we do this, then SERPINC1 reduction would increase by 12% and a new p-value of 0.026 is calculated for PDIA4KD effects between SERPINA1 and SERPINC1.

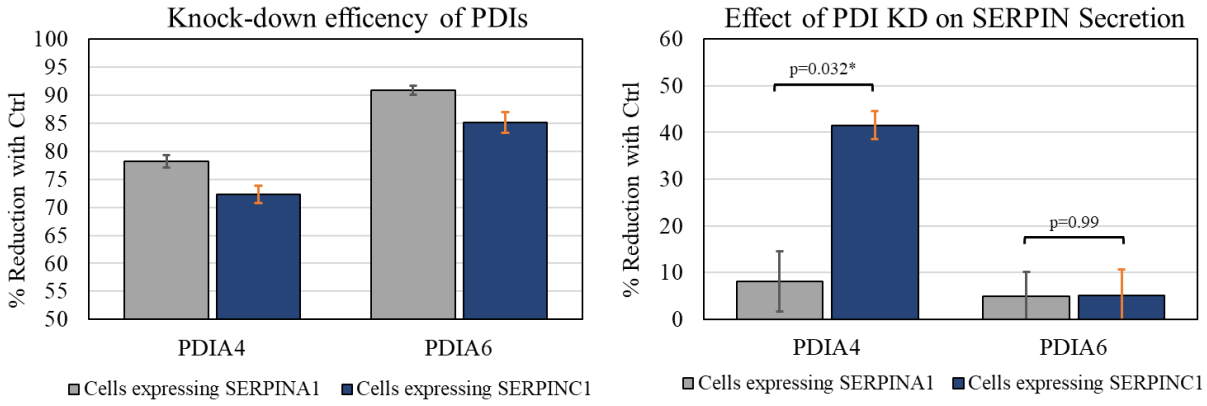

Figure 1.11. Effects of esiRNA mediated knockdown of isomerases PDIA4 and PDIA6 on SERPIN secretion. Semi-quantitative western blots were used to find the relative abundance of each PDI and SERPIN after esiRNA transfection in both cells expressing SERPINA1 and SERPINC1. We have reported these results as percent reduction compared to the negative control for PDIs (left) and SERPIN secretion (right). SERPINC1, which has various disulfide PTMs, shows a greater reduction in SERPIN production in PDIA4 experiments compared to SERPINA1, which does not have disulfide bridges in its structure. This is consistent with expected interactions found with BioID analysis.

1.4 Discussion

BioID has been shown to successfully profile interactomes for proteins-protein interactions that are transient and/or within difficult to reach cellular structures [30]. However, this is the first time that BioID has been used to identify proteomes of the secretory pathway during recombinant protein expression. We found that disulfide-bridge forming enzymes showed the strongest association with bait proteins enriched in disulfide bonds, supporting their critical roles in protein secretion and maintaining ER stability. Isomerases are critical for disulfide bond formation and thus proper protein structure, without which has been shown to initiate intracellular protein degradation by via ERAD mediated pathways [66]. Thus, insufficient

interaction between the secreted proteins with critical disulfide bonds and PDIs can limit secretion efficiency and serve as a rate-limiting step for protein production [8]. PDI expression, along with other SecMs, have also been shown to be tissue specific; in one study, positive correlation was found between P4HB/PDIA4 and liver tissue, where many disulfide-bond rich products are secreted by hepatocytes [67]. These observations suggest there is tissue-specific fine-tuning of the PDI family expression level in response to the enrichment of the disulfide sites. Cell hosts would similarity benefit from restoring ER stability by matching SecMs to recombinant proteins secretion.

While we show that BioID works well for studying the synthesis of secreted proteins, we acknowledge that biotin-based methods have limitations as well. Biotin is actively imported into the cytoplasm of cells and can freely diffuse to the nucleus, but it may not be as accessible in the secretory pathway, thus reducing labeling efficacy in that compartment [33]. Here we showed this challenge is not an insurmountable issue, i.e. the BioID2 construct successfully identified many expected luminal interactions. BioID2 requires less biotin supplementation, and exhibits enhanced labeling of proximate protein allowing for BioID to be introduced to new systems where biotinylation supplementation may not be easily accomplished [33,68]. More recently, two promiscuous mutants of biotin ligase, TurboID and miniTurbo, have been developed to catalyze proximity labeling even with much greater efficiency and therefore can be considered as an effective method when proximal labeling of the endomembrane organelles is desired [69].

Furthermore, the BirA fusion did not prevent identification of unique SecMs for the model proteins, and we utilized a Signal-BirA control to account for interference. The model proteins are assumed to enter the secretory pathway where they are processed and packaged for secretion with limited BirA interference. However, the coverage of proteins captured may be obscured by three factors: 1) the presence of endogenous biotinylation, 2) the maturation process of the biotin ligase, and 3) general variation in cell-stress responses caused by recombinant production. For example, although the extent to which the general endogenous biotinylation has not been systematically quantified, the biotinylated proteins isolated from the WT sample showed considerable overlap with interacting proteins detected in other model protein samples, suggesting endogenous biotinylation may be more pervasive than previously believed. In addition,

biotinylation occurs once the biotin ligase has been fully transcribed and functional. Therefore, early secretory pathway machinery may not be captured until the biotin ligase is fully active. Lastly, recombinant expression can cause various stress responses due to high protein secretion and/or the expression of non-native proteins. Therefore, we see much evidence of ERAD and UPR activity in our results. Variations in mRNA level caused by random integration can also trigger adaptive response such as the unfolded protein response in some cell lines which reciprocally alters the active PPIs network involved in the secretion. In this case, we use the Flp-In™ system (see Materials and Methods) for targeted integration of the transgenes into the same genomic locus to ensure comparable transcription rates of each transgene. Nonetheless, while we see evidence that ERAD pathways have increased expression that correlates to quality control of our structurally complex bait-proteins, ERAD and UPR responses can also to contributed to effects of recombinant expression. Continued improvement of the sensitivity/specificity of proteomic profiling of the secretory pathway will increase confidence and illuminate further PPIs of note.

In summary, we demonstrate here an approach to identify the protein interactions that synthesize and support secreted proteins, and thus define the product-specific secretory pathway. The identification of such machinery opens avenues for mammalian synthetic biology, wherein biotherapeutic production hosts can be rationally engineered to improve the titer and quality of diverse proteins in a client specific manner.

Acknowledgments

Chapter 1 contains some material from a manuscript currently in review for publication “In situ detection of protein interactions for recombinant therapeutic enzymes.” Samoudi, Mojtaba*; Kuo, Chih-Chung*; Robinson, Caressa M*.; Shams-Ud-Doha, Km; Schinn, Song-Min; Kol, Stefan; Weiss, Linus; Bjorn, Sara Petersen; Voldborg, Bjorn G.; Campos, Alexandre Rosa; Lewis, Nathan E.. (2020). *In review.* The thesis author is a primary author of this material.

Chapter 2

BAR facilitated identification of PPIs impacting protein production of Rituximab in CHO cells

Abstract

Chinese hamster ovary (CHO) cells are currently the industrial mammalian cell line of choice for producing recombinant therapeutic proteins, e.g., monoclonal antibodies, cytokines and vaccines. But despite CHO cells' usual versatility and suitability for human therapeutics, many proteins still fail to secrete well from them. Even with adequate mRNA levels, protein expression can be hindered due to bottlenecks in the secretory pathway, i.e. posttranslational events during protein production [1,2]. To enable efficient manufacturing of high-value proteins, we must decipher molecular mechanisms controlling the secretory pathway during protein expression. This knowledge would inform rational design for secretory pathway engineering, increasing recombinant protein production and ultimately expanding the global therapeutic portfolio. This study seeks to establish methods to unravel the host cell secretory pathway machinery that directly regulates secretion of specific proteins, and we apply it here to identify the key machinery modulating secretion of a valuable monoclonal antibodies (mAbs). Fc-mediated biotinylation by antibody recognition (FcBAR) was performed to quantify protein-protein interaction (PPI) networks regulating secretion of Rituximab from a panel of CHO-S clones with varying productivity levels [3,4]. Preliminary proteomic analysis reveals the importance of various vesicle transport, glycosylation (e.g. Mgat2, St3gal4, and St3gal1), and protein homeostasis proteins facilitating antibody secretion in high producing clones. Further analysis of this data will help guide host cell engineering for optimized biomanufacturing and provide insights into the functions of the mammalian secretory pathway during recombinant production.

2.1 Introduction

Increased demand for biotherapeutics parallels the improvements made in cell-line engineering over the years. Our ability to supply large amounts of a wide-range of molecule types is dependent on cell-lines that have been directly modified or screened to provide high titers, prolonged half-lifes, and stable therapeutic production [70]. Much of this has been vastly accelerated by the ability to perform targeted gene engineering with zinc finger nucleases, TALENS, and CRISPR-Cas based techniques [70]. Of increasing interest is the ability to engineer host cells with the proper machinery to perform necessary post-translational modifications (PTMs) [22,24]. While CHO cells are superior to bacterial and fungal host systems when it comes to performing PTMs, there is both a limit to the similarities between the CHO and human secretory pathways, as well as ample opportunities for leaner, more efficient CHO secretory pathways. Of efforts made to address these opportunities, there have already been several critical discoveries to improve the quantity and/or quality of biotherapeutics. For example, knocking out the fucosyltransferase FUT8 increases mAb ability to perform antibody-dependent, cell-mediated cytotoxicity, ideal for CD20 targeting antibodies; in addition, overexpressing N-acetylglucosaminyltransferase-III in CHO has been shown to increase protein glycosylation accuracy, while overexpressing chaperones in CHO cells has been shown to improve protein folding and secretion [70].

System-wide proteomics offers a way to identify a much larger pool of potential protein-specific targets for cell-line engineering. These potential targets can also be computationally compared by mass analyzed for their likely impacts on cell function and protein production. Targets can be identified for their impact on specific biotherapeutics, allowing for potentially highly specialized cell-lines. Lastly, illuminating deficiencies in the CHO secretion pathway will open-up the biopharmaceutical industry for previously difficult to express, or even un-manufacturable proteins [26].

Rituximab is a long-standing biotherapeutic approved by the FDA in 1997 and was the first monoclonal antibody approved to treat cancer. It is expressed in CHO cells as a murine (mouse)/human chimeric mAb and is used to target the CD20 antigen of B cells [71,72]. It is approved to treat a variety of

autoimmune diseases, such as rheumatoid arthritis, as well as cancers, such as non-Hodgkin's lymphoma. The Rituximab brand name, Rituxan under Roche and Biogen, expired in Europe in 2013 and 2018 in the U.S. In 2019, amongst competition from Pfizer and Amgen, the first FDA approved Rituximab biosimilar Truxima (Teva and Celltrion) arrived followed closely by Pfizer's Ruxience [73]. However, despite anticipated price cuts, Rituxan remains approximately $940/100mg with Truxima offered only 10% less [74]. Biosimilars will maintain momentum in a highly competitive market through cell line and manufacturing improvements that will facilitate quicker market entry and more drastic price reductions. Since the first FDA approved mAb Muromonab-CD3 in 1986, mAbs have been one of the fastest growing classes of biotherapeutics [75]. This, as well as Rituximab's well-characterized structure and function make Rituximab an ideal candidate therapeutic to study the impacts of secretory pathway regulation on titer.

Here, a system-wide labeling of secretory protein interactions coupled to MS was used to map interaction networks regulating the recombinant expression of the high value therapeutic Rituximab. For ten Rituximab-expressing CHO clones with varying protein yields and an appropriate CHO-S control, we used a Fc-mediated BAR (FcBAR) method to label PPIs followed by LC-MS/MS and label-free quantification (LFQ). In this case, we do not require a primary antibody to target Rituximab, hence the FC-mediated BAR method. For the panel of clones chosen, we ensured that there was little to no correlation between productivity and light nor heavy chain mRNA level, thus controlling for effects due to transcription levels. Our findings will elucidate cell processes (e.g., protein folding, modification, and transport) managing heterologous protein expression and identify critical PPIs that correlate positively with higher mAb productivity. Furthermore, it will guide efforts to engineer the CHO secretory pathway for enhanced protein expression in a product-specific manner.

2.2 Methods

2.2.1 Cell culture and clone expansion

A panel of twenty cell CHO-S clones expressing Rituximab were obtained from collaborators at DTU. A Laminin-alpha1 knock-out CHO-S clone line was also obtained to use as the wild-type (WT) as

laminin is an extracellular constituent and for our purposes will not significantly impact the secretory pathway. CHO-S cells were grown in 250ml shaking flasks at 30 rpm and 37ºC with CD CHO Media supplemented with 8mM L-Glutamine and Gibco anti-clumping agent. The Rituximab producing clones were grown in media also containing 250ug/ml G418 for selection.

2.2.2 Relative mRNA expression level using RT-qPCR

Total RNA was extracted from 1-2 x 106 cells using the RNeasy Plus Mini Kit (Qiagen) according to the manufacturer's instructions. RNA concentration was measured with Qubit fluorometric analysis (Life technologies). cDNA was synthesized from 1.5-3 µg of total RNA using the Maxima First Strand cDNA Synthesis Kit for RT-qPCR with dsDNAse treatment (ThermoFisher Scientific). RT-qPCR analyses were performed on the QuantStudio 5 Real-Time PCR System using TaqManTM Multiplex Master Mix (Thermo Fisher Scientific) in triplexes (gene of interest and two normalization genes) using the following amplification conditions: 50°C for 2 min, 95°C for 10 min; 40x: 95°C for 15 s, 60°C for 1 min. Custom-made Taqman assays were used for Rituximab light chain and heavy chain, as well as normalization genes Gnb1 and Fkbp1a. All the primers and probes were validated by melting curve analysis and primer efficiency test. Using the ΔΔCt method, the relative expression levels of Rituximab heavy and light chain were calculated by normalization to the geometric mean of expression levels of the two normalization genes. Each experiment included controls with no template and was performed using technical triplicates.

2.2.3 PPI labeling in CHO-S cells using the FcBAR method

Sample preparation began by harvesting approximately twenty million cells of each CHO-S clone expressing Rituximab and WT at the mid-exponential phase and then washing cells with warm PBS. Cells were fixed with 4.0% paraformaldehyde in PBST for ten minutes and then permeabilized with 0.4% PBST for ten minutes. Following this, peroxidases were inactivated with 0.2% H_2O_2 in PBS for ten minutes and then blocked for two hours with 5.0% normal goat serum in PBST, before a two-hour incubation with goat anti-human HRP-conjugated antibody (1:1000, ab97175). Proximal protein biotinylation occurs with treatment of H_2O_2 and tyramide-biotin (TSA Biotin Tyramide Reagent Pack, Perkin Elmer SAT700001EA), resulting in tyramide-biotin radicalization and covalent deposition onto proximal proteins. Tyramide biotin

stock solution was added to each sample and incubated for 5 minutes before addition of the H_2O_2 Amplification Diluent at the manufacturers' recommended ratio of 1:50, tyramide to diluent. The reaction proceeded for either one, three, or five minutes before being stopped with 500 mM Sodium Ascorbate solution in PBS.

2.2.3.1 Optimization of FcBAR reaction time

Three samples each of a rituximab producing clone (A4_1) and WT were grown and prepared as described above. Each sample was biotinylated using either a one, a three, or a five-minute reaction. Two biological replicates were prepared of each experiment before being sent for mass spectrometry. Enrichment was done to determine which reaction time provided enough coverage without sacrificing specificity of the biotinylation. This was done by removing background using the WT biotinylation profile, and then profiling enrichment for secretory pathway-related gene sets.

Because BAR labeling is dependent on proximity of targets and the time given for the active biotin-tyramide to diffuse from the HRP source to those targets, too little time given could result in low interaction detection, but too long of a reaction time could result in extraneous labeling outside of those that are real interactions. To determine which reaction time resulted in sufficient significant interactions with secretory pathway-related gene sets, we performed an iterative enrichment analysis in which we included the most significant interactions first and iteratively included less significant interactors with lower fold changes. The cumulative coverage of 3 secretory pathway related gene sets was reported as a function of the number of interactors included at each iteration.

2.2.3.2 Comparison of low and high producing CHO-S clones for detection of cell engineering targets

Ten rituximab producing CHO-S clones (A4_1, A6_3, B2_3, B5_1, E6_3, E7_3, E9_2, E10_3, F1_3, G12_4) and a non-producing CHO-S wild-type (from here on out referred to as WT) were grown and prepared as described above. After optimization, each sample was biotinylated using a three-minute reaction. Three biological replicates were prepared of each experiment before being sent for mass spectrometry.

2.2.4 Validation of labeling by immunofluorescence

A subsample of each experiment prepared was stained with streptavidin-DyLight 594 conjugate (1:1000, ThermoFisher #21842), targeting the biotinylated proteins, and anti-human goat monoclonal antibody-DyLight 650 conjugate (1:1000, ab98593), targeting Rituximab in blocking buffer for thirty minutes at room temperature. Cells were then washed, mounted on a slide using antifade Vectashield mountant containing DAPI, and imaged using Leica SP8 Confocal with Lightning Deconvolution.

2.2.5 Biotinylation profile by western blot

For staining of intracellular biotinylated proteins, cells were lysed by RIPA buffer. Protein was loaded for a x3.5 dilution onto SDS-PAGE gel and resolved proteins were transblotted to nitrocellulose membrane using Trans-Blot Turbo Transfer System from Bio-RAD. The membrane was blocked by 5% BSA in TBST and probed with HRP-conjugated streptavidin diluted in blocking buffer at a 1:2000 ratio. Membranes were washed and ProSignal Pico ECL Substrate was added (Prometheus Protein Biology). Protein bonds were visualized using G:Box Gel Image Analysis Systems (SYNGENE).

2.2.6 Mass spectrometry

MS quantification of the biotinylated proteins was done similar to that described above in Chapter 1. Briefly, biotinylated proteins were enriched using streptavidin beads, followed by trypsin digestion. Digested peptides were subjected to LC/MS-MS and identified by MS peptide mapping in MaxQuant against the UniProt *Cricetulus griseus* data-base [46,76–78] . All LC-MS/MS was done by Dr. Alexandre Rosa Campos and his Proteomics Core at the SBP Medical Discovery Institute, where he has been an active collaborator on all our proximity labeling studies.

2.2.7 Data Analysis

Data analysis was performed using Perseus version 1.6.14.0 to quantify protein interactions between Rituximab-producing clones relative to WT. Data clean-up, filtering, and left-centered imputation for were performed before a Welch t-test was performed using and s0 of three to calculate fold-change threshold and an FDR of < 0.05. Analysis of significant proteins pathways was performed using DAVID V.6.8 [79,80].

2.3 Results

2.3.1 mRNA quantification of Rituximab for CHO-S clones

The panel of CHO-S clones obtained from our collaborator Dr. Nuša Pristovšek at DTU covers >2-fold range of productivities (i.e., picograms rituximab produced per cell per day), but since limiting transgene mRNA levels can reduce CHO cell productivity, the Rituximab mRNA levels in the CHO-S cell clones were quantified by qPCR to ensure productivity of selected clones could not be explained by low levels of either heavy or light chain mRNA. Productivity of protein production were seemingly not influenced by low mRNA levels (Figure 2.1). Thus, the variation in protein titer is anticipated to be due to limitations in protein synthesis, folding and/or trafficking. These clones will be further studied in this proposed work.

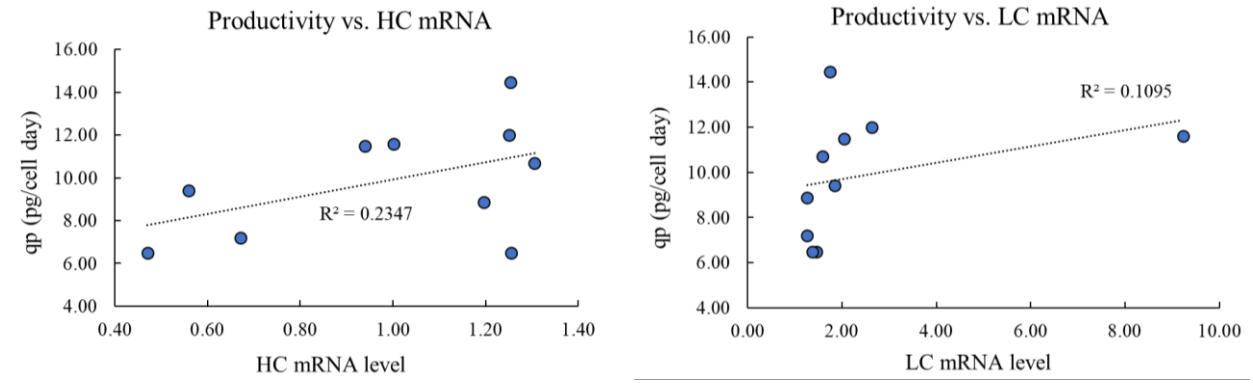

Figure 2.1. Cell specific productivity is poorly predicted by transgene mRNA level. For 10 CHO-S clones expressing Rituximab, we measured mAb production (qp: pg/cell day) and mRNA level by qPCR, for both the heavy chain (HC) (left) and the light chain (LC) (right). For all clones, variation in mRNA level showed little to no correlation with productivity.

2.3.2 Implementation of proximity biotinylation assays for CHO-S clones

First, we perform the BAR methodology for each CHO-S clone and validate if there is evidence of specific antibody facilitated biotinylation of intracellular PPIs between Rituximab and its supporting SecMs. Western blots verified successful biotinylation of various proteins occurred with the BAR method performed (Fig. 2.2). Specifically, blots of the cell lysates stained with streptavidin-HRP against the biotinylated proteins revealed the biotinylation profile for both the biotin labeled and unlabeled samples of the Rituximab-producing clones and WT. The biotinylation profiles of the negative controls (unlabeled clones and labeled/unlabeled WT) and Rituximab-producing biotin labeled clones differed substantially,

with many specific biotinylated proteins in the labeled clones excluding the WT, suggesting that proximity biotin labeling occurs on specific interacting partners.

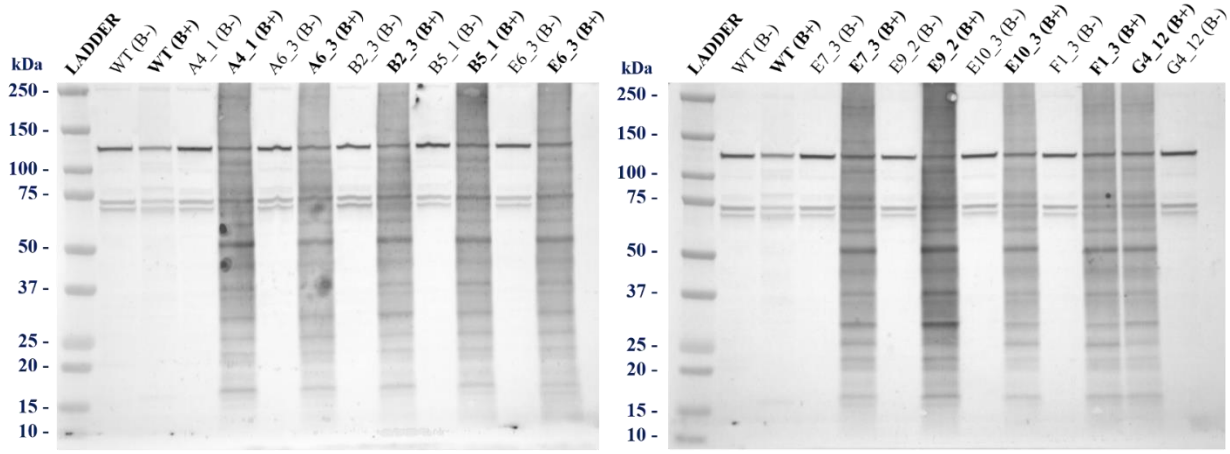

Figure 2.2. Expression of Rituximab results in a substantial increase in biotinylated proteins after BAR implementation. Immunoblotting biotinylation profiling of labeled proteins in WT control and CHO cells with HRP-streptavidin. With BAR implementation, Rituximab secreting cells experienced biotinylation of a subset of proteins (B+) which are not seen in WT or absence of the BAR reaction (B-). A few endogenously biotinylated proteins appear in the absence of BAR and in the WT.

The subcellular localization of Rituximab and biotinylated interactors were determined by multicolor co-immunofluorescence microscopy using a Leica SP8 Confocal microscope (Fig. 2.3. & Fig. 2.4.). We see evidence of biotinylation in the strepdavidin-594 channel for clones expressing Rituximab, but little to no labeling for WT. Thus, indicating the FcBAR system successfully labels Rituximab-proximal proteins.

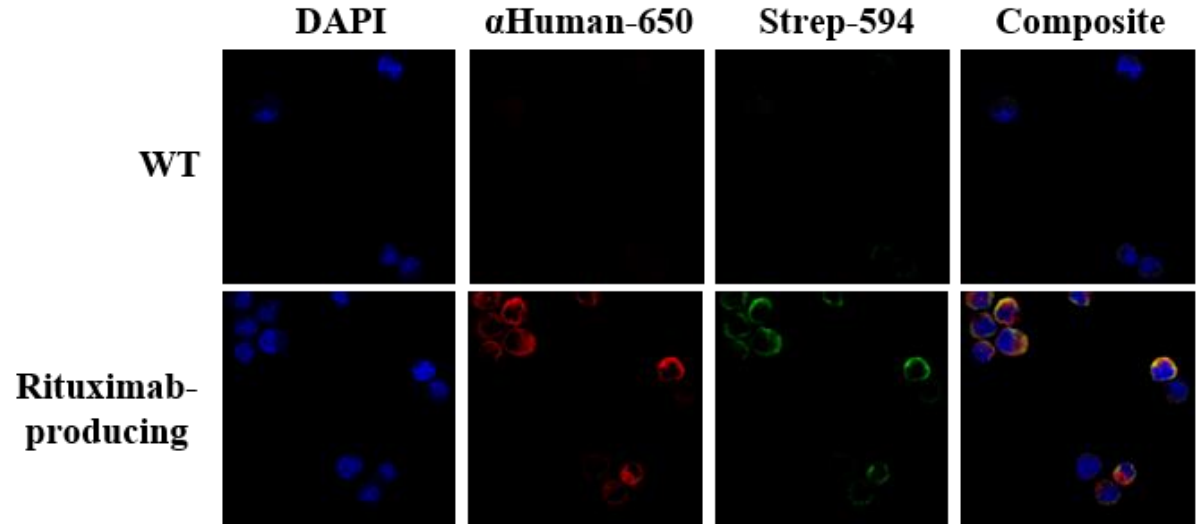

Figure 2.3. Up-close view of rituximab producing clone (B2_3) versus the WT.

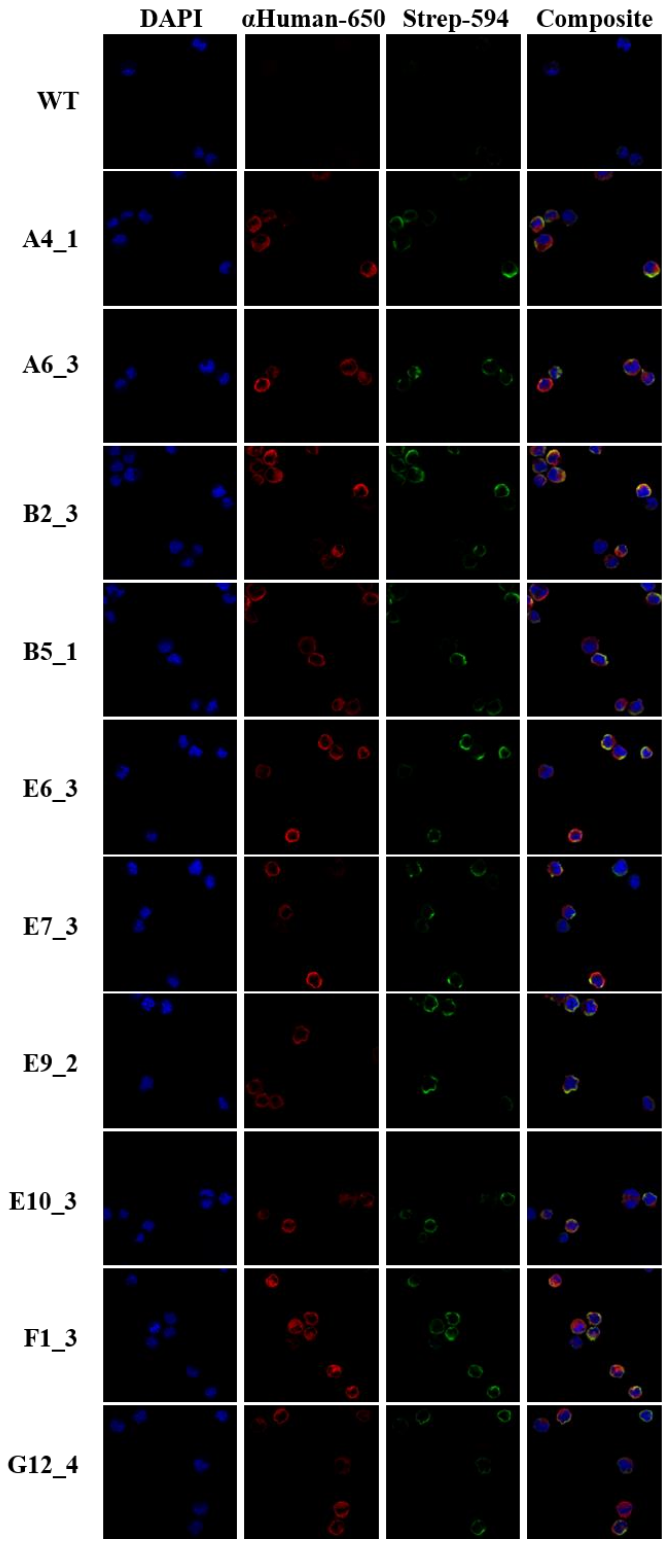

Figure 2.4. Rituximab colocalized with biotin-staining. Co-Immunofluorescence demonstrated the intracellular colocalization of the biotin-labeled proteins (stained with Streptavidin-Dylight 594 and illustrated in green color) and bait-BirA (stained with anti-human monoclonal antibody-Dylight 650 and illustrated in red color), while WT did not show increased biotinylation under the same experimental conditions.

2.3.3 Preliminary data analysis shows enrichment of SecMs

2.3.3.1 Optimization of FcBAR reaction time

We deployed FcBAR to identify SecMs interacting with Rituximab in one drug-producing clone and contrasted it with a non-producing WT control clone (WT). We then tested if significant interactions were enriched in three independent secretory pathway gene sets (i.e. known SecMs, secretory pathway resident proteins, and co-secreted proteins). A reaction time of 3-minutes ultimately resulted in the best enrichment for known secretory pathway components (Fig. 2.5)

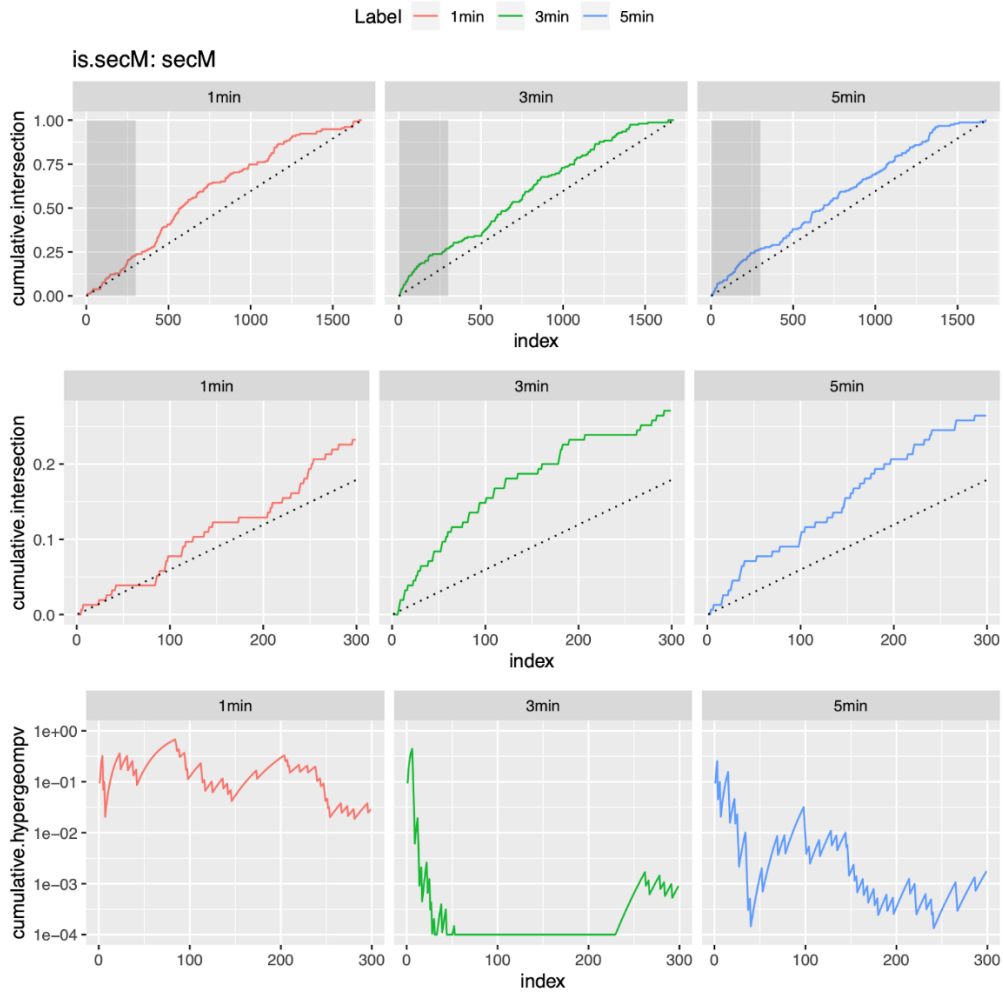

Figure 2.5. Significant interactions enrich for secretory pathway-related secretory machinery genes best for a 3-minute reaction. ROC curve, where the X-axis represents the number of interactors included, and the Y-axis the coverage of each secretory pathway gene set. The coverage of the secretory machinery gene set (top) and shadowed-window zoom-in (middle), along with their corresponding hypergeometric enrichment p-values (bottom) are shown. The 1-minute reaction resulted in less coverage, while the 5-minute reaction contained more background.

2.3.3.2 Comparison of low and high producing CHO-S clones for detection of cell engineering targets

Analysis of MS data with MaxQuant found 4305 total proteins captured. After data processing in Perseus, principal component analysis (PCA) was performed using R package pcaMedhods' probabilistic PCA (PPCA) method. PCA revealed good separation of WT and Rituximab-producing clones (Fig 2.6).

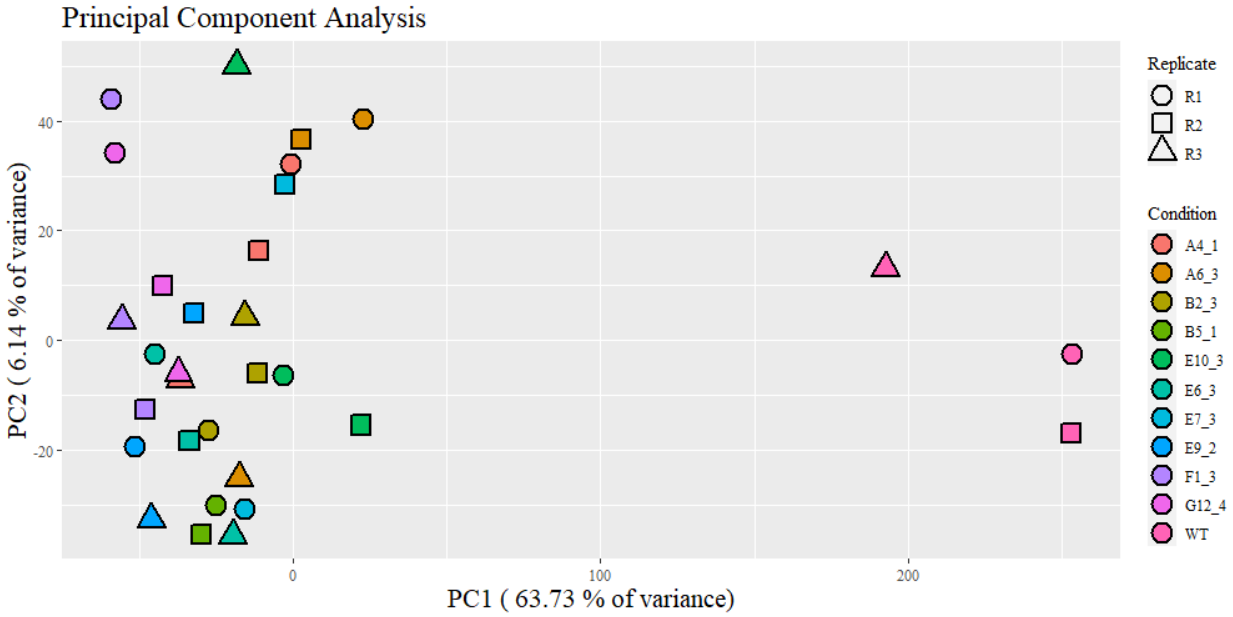

Figure 2.6. PCA for all captured proteins and samples.

DAVID Gene Ontology (GO) term enrichment was performed on the top 100 loadings in the first principal component, using an FDR <0.05 to identify significant terms. This indicated that the major cellular components and biological processes accounting for variance among samples belong primarily to processes within the secretory pathway, i.e. protein folding and protein transportation (Fig. 2.7).

Differential expression of proteins was determined using a Welch's t-test for each clone protein against the WT (Fig.2.8). Significant proteins were selected using an FDR < 0.05 and a fold-change threshold determine by an s0 of 3. There was little background seen in the WT, which demonstrates the specificity of our anti-body and lack of excess label-leakage, but also causes an abundance of significant proteins with a positive fold-change.

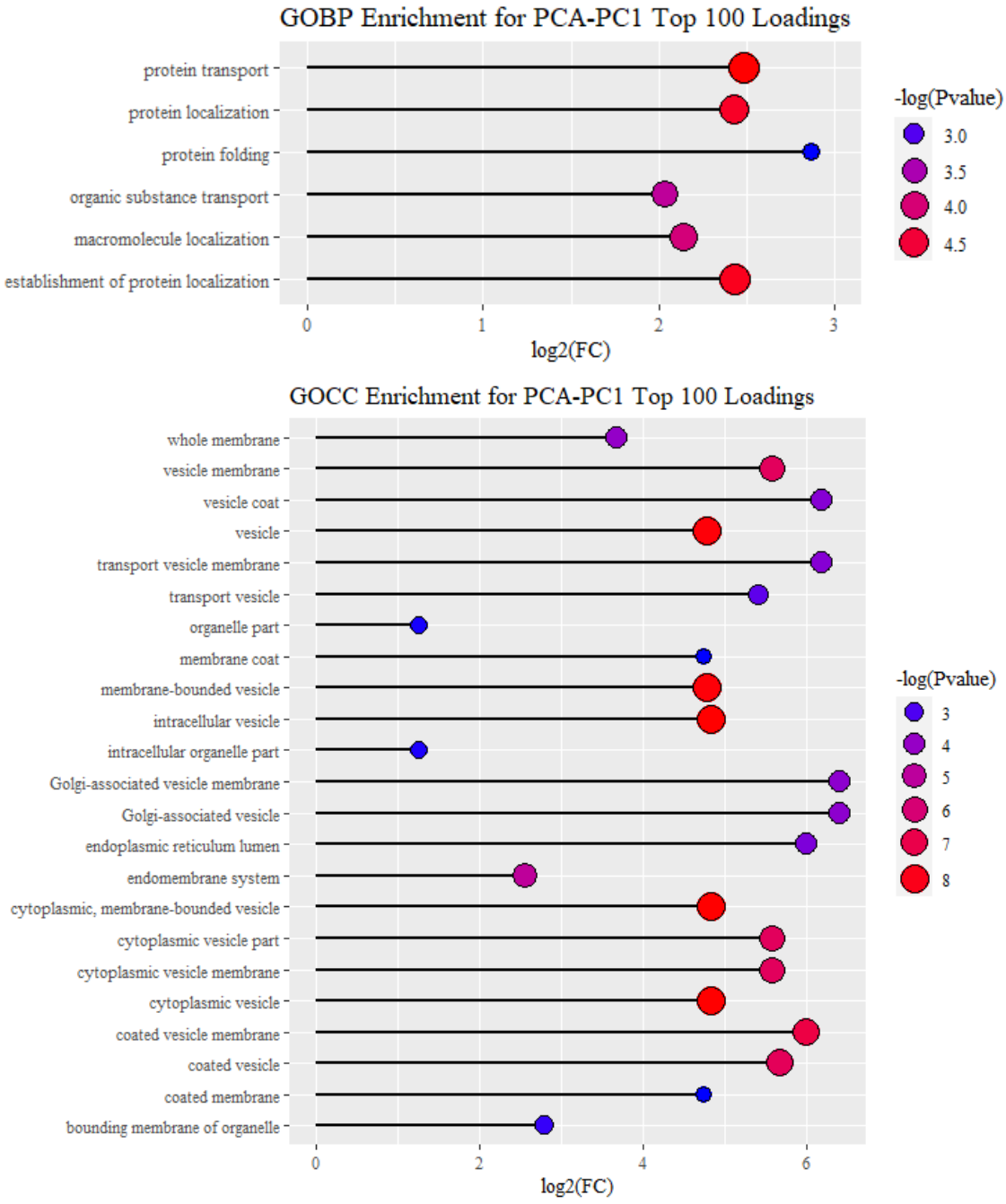

Figure 2.7. DAVID pathway results for GO biological function (BP) and cellular components (CC) annotations using PCA PC1 top 100 loadings. These terms that indicate secretory pathway localization and SecM processes dominate among the proteins accounting most for variation between samples.

Figure 2.8. Differential expression results for each clone versus the WT. Significant proteins are labeled in red and those points that have a -log(Pvalue) > 4 and a log2(Fold change) >6 are also labeled by Protein ID. The lack of background noise found in the WT causes an abundance of positive significant hits.

Visualization of differential expression (DE) results was done with heatmaps for binary significance (True or False), fold change, and FDR adjusted p-value (Fig. 2.9). Proteins were visualized from largest to smallest productivity, but any clustering of results is not obvious. As all clones are of the same cell-line and successfully express Rituximab, it is not unusual that there are more similarities between clones than differences.

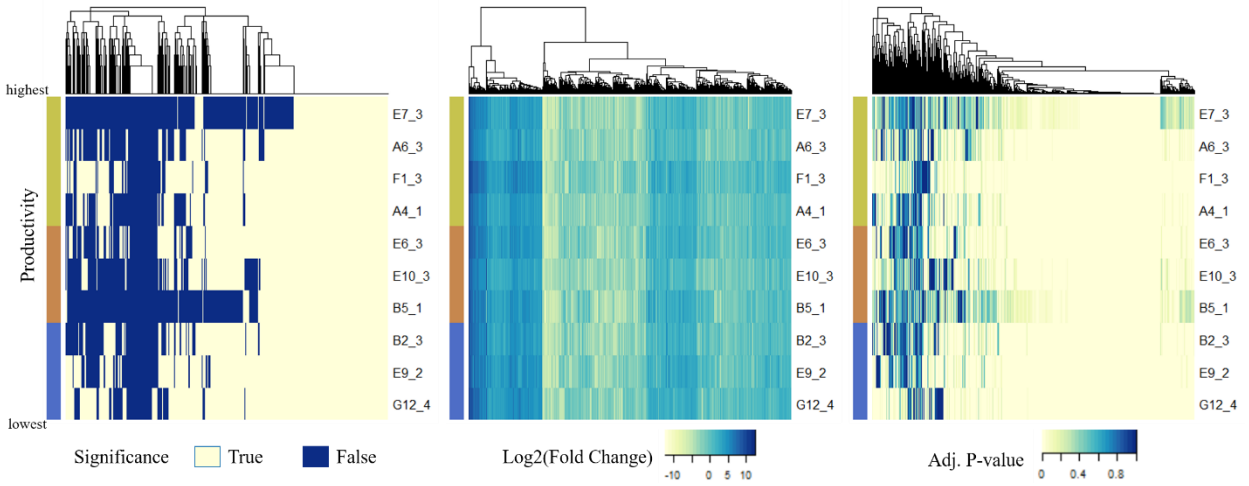

Figure 2.9. Heat maps showing significance (left), fold change (middle), and FDR adjusted P-value results for DE of each clone with the WT. The clones are ordered largest from smallest productivity (pg/day protein production).

Using the PCA results and strongest significance hits to narrow the list of possible SecM interactors, we visualize DE results from each clone for a subset of proteins (Fig.2.10). We see a strong interaction among all clones for Dnajc10 (a member of the ERAD pathway). Like many other IgG1 mAbs, Rituximab has 16 disulfide bonds that connects its heavy and light chains and requires proper quality control [81]. We also see evidence of necessary regulation with high fold-change observed for granulin precursor and cathepsin Z (Table 2.1). Both progranulin and cathepsin localize within and regulate the function of lysosomes [82,83]. In addition, we also see a large fold-change for dynactin subunit 2 (Dctn2). Dynactin is a cytosolic structural component; however, it also facilitates ER-to-Golgi transport and movement of vesicles, such as lysosomes [84]. Thus, maintenance of quality control pathways and vesicle function are of importance during recombinant secretion.

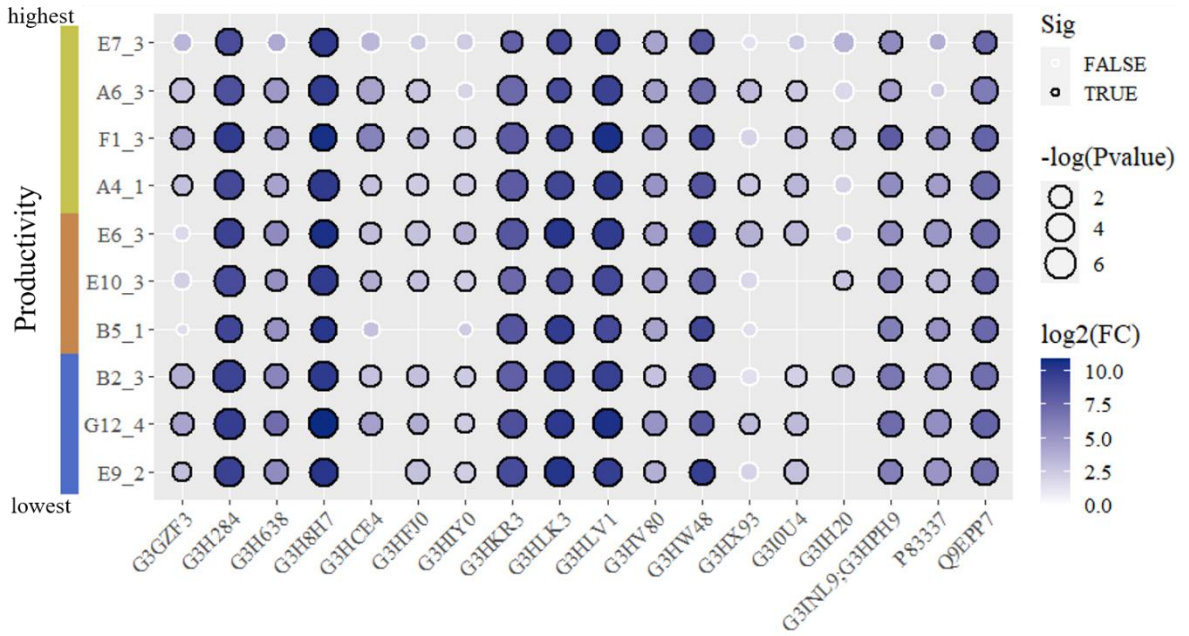

Figure 2.10. Visualization of select potential PPIs for Rituximab. The clones are ordered largest from smallest productivity (pg/day protein production).

Table 2.1. Functional information for the subset of proteins visualized in Figure 2.10.

Protein ID	Species	David Gene Name	Functional Information
G3GZF3_CRIGR	Cricetulus griseus	mannosyl (alpha-1,6-)-glycoprotein beta-1,6-N-acetyl-glucosaminyltransferase(Mgat5)	GOBP: protein N-linked glycosylation via asparagine
G3H284_CRIGR	Cricetulus griseus	endoplasmic reticulum protein 29 (Erp29)	GOBP: protein secretion, activation of MAPK activity
G3H638_CRIGR	Cricetulus griseus	activator of Hsp90 ATPase activity 1(Ahsa1)	GOMF: Hsp90 protein binding, chaperone binding
G3H8H7_CRIGR	Cricetulus griseus	DnaJ heat shock protein family (Hsp40) member C10 (Dnajc10)	GOBP: cell redox homeostasis, protein folding in ER, ubiquitin-dependent ERAD
G3HCE4_CRIGR	Cricetulus griseus	fucosyltransferase 8(Fut8)	GOBP: N-glycan fucosylation
(G3HFJ0) Q80WL0_CRIGR	Cricetulus griseus	ST3 beta-galactoside alpha-2,3-sialyltransferase 4(St3gal4)	GOBP: protein glycosylation, sialylation
G3HIY0_CRIGR	Cricetulus griseus	transmembrane protein 167A (Tmem167a)	GOBP: constitutive secretory pathway
G3HKR3_CRIGR	Cricetulus griseus	protein tyrosine phosphatase, non-receptor type 23 (Ptpn23)	GOBP: positive regulation of early endosome to late endosome transport, protein dephosphorylation
G3HLK3_CRIGR	Cricetulus griseus	granulin precursor (Grn)	GoBP: Protein homeostasis
G3HLV1_CRIGR	Cricetulus griseus	FK506 binding protein 10(Fkbp10) (Peptidylprolyl isomerase)	GOMF: peptidyl-prolyl cis-trans isomerase activity
G3HV80_CRIGR	Cricetulus griseus	vesicle associated membrane protein 7 (Vamp7)	GOBP: Golgi to plasma membrane protein transport, SNARE complex assembly, endosome to lysosome transport
G3HW48_CRIGR	Cricetulus griseus	dynactin subunit 2 (Dctn2)	GOBP: Vesicle, endoplasmic reticulum to Golgi vesicle-mediated transport
G3HX93_CRIGR	Cricetulus griseus	CCAAT/enhancer binding protein zeta (Cebpz)	GOMF: RNA polymerase II cis-regulatory region sequence-specific DNA binding GOBP: ribosome biogenesis
G3I0U4_CRIGR	Cricetulus griseus	glycogenin 1(Gyg1)	GOMF: glycogenin glucosyltransferase activity
G3IH20_CRIGR	Cricetulus griseus	mannosyl (alpha-1,6-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase(Mgat2)	GOBP: protein N-linked glycosylation via asparagine
G3INL9_CRIGR; G3HPH9_CRIGR	Cricetulus griseus	ST3 beta-galactoside alpha-2,3-sialyltransferase 1 (St3gal1)	GOBP: protein glycosylation, O-glycan processing, sialylation
(P83337) OFUT1_CRIGR	Cricetulus griseus	protein O-fucosyltransferase 1 (Pofut1)	GOBP: protein glycosylation
Q9EPP7_CRIGR	Cricetulus griseus	cathepsin Z (Ctsz)	GOMF: carboxypeptidase activity GOBP: proteolysis, cytoplasmic vesicle

In addition, we see evidence of interaction with N-acetylglucosaminyltransferases Mgat2 and Mgat5, as well as sialyltransferases St3gal1 and St3gal4 (Fig. 2.10.). Mgat2 and Mgat5 are both important for N-linked glycan branching within the Golgi; however, Mgat2 is responsible for forming biantennary, complex N-glycans, while Mgat5 is responsible for adding GlcNAc to generate branched, tri/tetraantennary N-glycans [85,86]. Rituximab glycan heterogeneity is typically reserved for biantennary structures, mostly G0F and G1F [87]. Sialytransferases transfer sialic acid to N-glycans and have effects on protein stability and increased *in vivo* and serum half-life [88,89]. Previous profiling of glycosylation in CHO have also showed St3gal4 to have a dominant role in CHO sialylation [85]. N-linked glycosylation is pivotal for protein function, but has also been implicated as necessary for protein stability and secretion [90,91]. We see that Mgat2 is captured in more high-producing clones, however not with high significance in all clones. The capturing of Mgat5 could be Rituximab proximity to co-secreted proteins that require tri/tetraantennary N-glycans. Lastly, core fucosylation is catalyzed almost exclusively by Fut8, thus its presence in our samples in expected as Rituximab typically contains this modification, regardless of advancements in afucosylated mAbs with antibody-dependent cellular cytotoxicity activity [92]. Here, we see higher fold-change and log p-value of Fut8 in two clones with highest titer.

2.4 Discussion

BAR, as an alternative to using fusion-protein dependent BioID, was successfully used to profile interactions during recombinant mAb secretion in CHO-S clones. Unlike BioID from Chapter 1 that saw a promiscuity of the biotin deposition in the WT, the specificity of the antibody used and the ability to block endogenous peroxidases limited the amount of background captured in the WT sample. With access to an appropriate antibody, BAR can therefore be applied in a variety of systems with good specificity of biotinylation. On the other hand, the opposite can be true and availability or lack of a good antibody to target the protein the interest can act as an obstacle to performing the BAR method. We also showed ability to tailor the BAR reaction by profiling the extent of labeling with different reaction incubation times. Optimization of reaction time can prevent under/excessive biotin labeling. However, the high prevalence

of Rituximab throughout cells made determination of the most optimal reaction time by immunofluorescence not possible. Performing the optimization thus required its own LC-MS/MS analysis, requiring materials and time.

Here, we have applied the BAR method to profile the secretory pathway interactions of Rituximab-producing clones that have differing experimental titers. We have used a parental-like CHO-S clone to remove background from endogenous biotinylation and secondary-antibody non-specificity. Strong interactions found significant among Rituximab-producing samples included SecMs required for protein homeostasis and protein transport, as well as modifying enzymes such as isomerizes and glucosyltransferases. Heterogeneity of glycosylation from various Mgat, ST3gal family, and Fut8 interactions may impact secretion and protein stability [90,91]. Previous efforts have primarily focused on the engineering of glucotransferases for controlling glycan heterogeneity [85,93]. Overexpression of Mgat1 and Mgat4 for example have been used to increase sialyation of recombinant human erythropoietin in CHO cells as they promote complex glycan branching [89]. Further data analysis is needed to explore correlation, if it can be seen here, between productivity and SecM interactors among clones. The preliminary data analysis done here will facilitate a deeper look at each clone's interactome and correlation with titer.

We acknowledge that this study is limited by the clones available and would be improved by a larger population of Rituximab-producing clones. That said, strong interactors among clones indicate potential targets for engineering as they are shown to facilitate Rituximab secretion. Additional analysis will likely reveal additional SecM's required for high protein production. Future steps may also include the addition of experimental controls for biotin label-leakage. This can be achieved by profiling interactions for a house-keeping protein that resides outside of the secretory pathway. While we partially control biotin-labeling with reaction time, label-leakage of biotin into non-relevant compartments can also occur due to excessive permeabilization or the overwhelming presence of the target recombinant protein. Immunofluorescence revealed a ubiquity of intracellular Rituximab, this may cause masking of interesting interactions and may explain the high number of total significant interactions captured.

Acknowledgments

Chapter 2 is based on material currently being prepared for publication. “Identifying the essential proteins for supporting monoclonal antibody secretion by proximity-labeling mass spectrometry.” Samoudi, Mojtaba; Robinson, Caressa M.; Kuo, Chih-Chung; Pristovšek, Nuša; Shams-Ud-Doha, Km; Hansen, Henning G.; Kildegaard Helene F.; Lee, Gyun M.; Campos, Alexandre Rosa; Lewis, Nathan E.. The thesis author is expected to be a primary author of this material.

References

1. Urquhart L. Top drugs and companies by sales in 2018. *Nat Rev Drug Discov*. Published online March 12, 2019. doi:10.1038/d41573-019-00049-0

2. Barnes LM, Bentley CM, Dickson AJ. Molecular definition of predictive indicators of stable protein expression in recombinant NS0 myeloma cells. *Biotechnol Bioeng*. 2004;85(2):115-121. doi:10.1002/bit.10893

3. Uhlen M, Tegel H, Sivertsson Å, Kuo C-C, Gutierrez JM, Lewis NE, Forsström B, Dannemeyer M, Fagerberg L, Malm M, Vunk H, Edfors F, Hober A, Sjöstedt E, Kotol D, Mulder J, Mardinoglu A, Schwenk JM, Nilsson P, Zwahlen M, Takanen JO, Feilitzen K von, Stadler C, Lindskog C, Ponten F, Nielsen J, Palsson BO, Volk A-L, Lundqvist M, Berling A, Svensson A-S, Kanje S, Enstedt H, Afshari D, Ekblad S, Scheffel J, Katona B, Vuu J, Lindström E, Xu L, Mihai R, Bremer L, Westin M, Muse M, Mayr LM, Knight S, Göpel S, Davies R, Varley P, Hatton D, Fields R, Voldborg BG, Rockberg J, Schiavone LH, Hober S. The human secretome – the proteins secreted from human cells. *bioRxiv*. Published online November 27, 2018:465815. doi:10.1101/465815

4. Jin L, Li L, Hu C, Paez-Cortez J, Bi Y, Macoritto M, Cao S, Tian Y. Integrative analysis of transcriptomic and proteomic profiling in inflammatory bowel disease colon biopsies. *Inflamm Bowel Dis*. 2019;25(12):1906-1918. doi:10.1093/ibd/izz111

5. Yugandhar K, Gupta S, Yu H. Inferring Protein-Protein Interaction Networks From Mass Spectrometry-Based Proteomic Approaches: A Mini-Review. *Comput Struct Biotechnol J*. 2019;17:805-811. doi:10.1016/j.csbj.2019.05.007

6. Guo S, Zou J, Wang G. Advances in the proteomic discovery of novel therapeutic targets in cancer. *Drug Des Devel Ther*. 2013;7:1259-1271. doi:10.2147/DDDT.S52216

7. Anelli T, Sitia R. Protein quality control in the early secretory pathway. *EMBO J*. 2008;27(2):315-327. doi:10.1038/sj.emboj.7601974

8. Feizi A, Gatto F, Uhlen M, Nielsen J. Human protein secretory pathway genes are expressed in a tissue-specific pattern to match processing demands of the secretome. *npj Syst Biol Appl*. 2017;3(1):22. doi:10.1038/s41540-017-0021-4

9. Lund AM, Kaas CS, Brandl J, Pedersen LE, Kildegaard HF, Kristensen C, Andersen MR. Network reconstruction of the mouse secretory pathway applied on CHO cell transcriptome data. *BMC Syst Biol*. 2017;11(1):37. doi:10.1186/s12918-017-0414-4

10. Novick P, Ferro S, Schekman R. Order of events in the yeast secretory pathway. *Cell*. 1981;25(2):461-469. doi:10.1016/0092-8674(81)90064-7

11. EG R, JC S. Navigating the Secretory Pathway: Conference on Exocytosis Membrane Structure and Dynamics. *EMBO Rep*. 2002;3(9). doi:10.1093/EMBO-REPORTS/KVF185

12. Pearl LH, Prodromou C. Structure and Mechanism of the Hsp90 Molecular Chaperone Machinery. *Annu Rev Biochem*. 2006;75(1):271-294. doi:10.1146/annurev.biochem.75.103004.142738

13. Calakos N, Bennett MK, Peterson KE, Scheller RH. Protein-protein interactions contributing to the specificity of intracellular vesicular trafficking. *Science (80-)*. 1994;263(5150):1146-1149. doi:10.1126/science.8108733

14. Wang M, Wey S, Zhang Y, Ye R, Lee AS. Role of the unfolded protein response regulator GRP78/BiP in development, cancer, and neurological disorders. *Antioxidants Redox Signal*. 2009;11(9):2307-2316. doi:10.1089/ars.2009.2485

15. Taipale M, Tucker G, Peng J, Krykbaeva I, Lin ZY, Larsen B, Choi H, Berger B, Gingras AC, Lindquist S. A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell*. 2014;158(2):434-448. doi:10.1016/j.cell.2014.05.039

16. Taipale M, Krykbaeva I, Koeva M, Kayatekin C, Westover KD, Karras GI, Lindquist S. Quantitative analysis of Hsp90-client interactions reveals principles of substrate recognition. *Cell*. 2012;150(5):987-1001. doi:10.1016/j.cell.2012.06.047

17. Taipale M, Jarosz DF, Lindquist S. HSP90 at the hub of protein homeostasis: Emerging mechanistic insights. *Nat Rev Mol Cell Biol*. 2010;11(7):515-528. doi:10.1038/nrm2918

18. Fernández-Rivero N, Franco A, Velázquez-Campoy A, Alonso E, Muga A, Prado A. A Quantitative Characterization of Nucleoplasmin/Histone Complexes Reveals Chaperone Versatility. *Sci Rep*. 2016;6(1):1-14. doi:10.1038/srep32114

19. Wang D, Eraslan B, Wieland T, Hallström B, Hopf T, Zolg DP, Zecha J, Asplund A, Li L, Meng C, Frejno M, Schmidt T, Schnatbaum K, Wilhelm M, Ponten F, Uhlen M, Gagneur J, Hahne H, Kuster B. A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. *Mol Syst Biol*. 2019;15(2). doi:10.15252/msb.20188503

20. Liu Y, Beyer A, Aebersold R. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell*. 2016;165(3):535-550. doi:10.1016/j.cell.2016.03.014

21. Mathias S, Fischer S, Handrick R, Fieder J, Schulz P, Bradl H, Gorr I, Gamer M, Otte K. Visualisation of intracellular production bottlenecks in suspension-adapted CHO cells producing complex biopharmaceuticals using fluorescence microscopy. *J Biotechnol*. 2018;271:47-55. doi:10.1016/j.jbiotec.2018.02.009

22. Alves CS, Dobrowsky TM. Strategies and Considerations for Improving Expression of "Difficult to Express" Proteins in CHO Cells. In: *Methods in Molecular Biology*. Vol 1603. Humana Press Inc.; 2017:1-23. doi:10.1007/978-1-4939-6972-2_1

23. Lee GW, Fecko JK, Yen A, Donaldson D, Wood C, Tobler S, Vunuum S, Luo Y, Leonard M. Improving the Expression of a Soluble Receptor:Fc fusion Protein in CHO Cells by Coexpression with the Receptor Ligand. In: *Cell Technology for Cell Products*. Springer Netherlands; 2007:29-39. doi:10.1007/978-1-4020-5476-1_4

24. Johari YB, Estes SD, Alves CS, Sinacore MS, James DC. Integrated cell and process engineering for improved transient production of a "difficult-to-express" fusion protein by CHO cells. *Biotechnol Bioeng*. 2015;112(12):2527-2542. doi:10.1002/bit.25687

25. Hasegawa H, Wendling J, He F, Trilisky E, Stevenson R, Franey H, Kinderman F, Li G, Piedmonte DM, Osslund T, Shen M, Ketchem RR. In vivo crystallization of human IgG in the endoplasmic reticulum of engineered chinese hamster ovary (CHO) cells. *J Biol Chem*. 2011;286(22):19917-19931. doi:10.1074/jbc.M110.204362

26. Peng RW, Guetg C, Tigges M, Fussenegger M. The vesicle-trafficking protein munc18b increases the secretory capacity of mammalian cells. *Metab Eng*. 2010;12(1):18-25. doi:10.1016/j.ymben.2009.08.007

27. Peng RW, Abellan E, Fussenegger M. Differential effect of exocytic SNAREs on the production of recombinant proteins in mammalian cells. *Biotechnol Bioeng*. 2011;108(3):611-620. doi:10.1002/bit.22986

28. Becker E, Florin L, Pfizenmaier K, Kaufmann H. An XBP-1 dependent bottle-neck in production of IgG subtype antibodies in chemically defined serum-free Chinese hamster ovary (CHO) fed-batch processes. *J Biotechnol*. 2008;135(2):217-223. doi:10.1016/j.jbiotec.2008.03.008

29. Rao VS, Srinivas K, Sujini GN, Sunand Kumar GN. Protein-Protein Interaction Detection: Methods and Analysis. Published online 2014. doi:10.1155/2014/147648

30. Varnaitė R, MacNeill SA. Meet the neighbors: Mapping local protein interactomes by proximity-dependent labeling with BioID. *Proteomics*. 2016;16(19):2503-2518. doi:10.1002/pmic.201600123

31. Chen CL, Perrimon N. Proximity-dependent labeling methods for proteomic profiling in living cells. *Wiley Interdiscip Rev Dev Biol*. 2017;6(4). doi:10.1002/wdev.272

32. Rees JS, Li XW, Perrett S, Lilley KS, Jackson AP. Selective proteomic proximity labeling assay using tyramide (SPPLAT): A quantitative method for the proteomic analysis of localized membrane-bound protein clusters. *Curr Protoc Protein Sci*. 2015;2015:19.27.1-19.27.18. doi:10.1002/0471140864.ps1927s80

33. Kim DI, Roux KJ. Filling the Void: Proximity-Based Labeling of Proteins in Living Cells. *Trends Cell Biol*. 2016;26(11):804-817. doi:10.1016/j.tcb.2016.09.004

34. Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol*. 2012;196(6):801-810. doi:10.1083/jcb.201112098

35. Kim DI, Birendra KC, Zhu W, Motamedchaboki K, Doye V, Roux KJ. Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc Natl Acad Sci U S A*. 2014;111(24). doi:10.1073/pnas.1406459111

36. Firat-Karalar EN, Stearnsx T. Probing mammalian centrosome structure using BioID proximity dependent biotinylation. *Methods Cell Biol*. 2015;129:153-170. doi:10.1016/bs.mcb.2015.03.016

37. Gupta GD, Coyaud É, Gonçalves J, Mojarad BA, Liu Y, Wu Q, Gheiratmand L, Comartin D, Tkach JM, Cheung SWT, Bashkurov M, Hasegan M, Knight JD, Lin ZY, Schueler M, Hildebrandt F, Moffat J, Gingras AC, Raught B, Pelletier L. A Dynamic Protein Interaction Landscape of the Human Centrosome-Cilium Interface. *Cell*. 2015;163(6):1484-1499. doi:10.1016/j.cell.2015.10.065

38. Dong JM, Tay FPL, Swa HLF, Gunaratne J, Leung T, Burke B, Manser E. Proximity biotinylation provides insight into the molecular composition of focal adhesions at the nanometer scale. *Sci Signal*. 2016;9(432):rs4-rs4. doi:10.1126/scisignal.aaf3572

39. Hoffman AM, Chen Q, Zheng T, Nicchitta C V. Heterogeneous translational landscape of the endoplasmic reticulum revealed by ribosome proximity labeling and transcriptome analysis. *J Biol Chem*. 2019;294(22):8942-8958. doi:10.1074/jbc.RA119.007996

40. Bar DZ, Atkatsh K, Tavarez U, Erdos MR, Gruenbaum Y, Collins FS. Biotinylation by antibody recognition - a method for proximity labeling. *Nat Methods*. 2018;15(2):127-133. doi:10.1038/nmeth.4533

41. Lucas A, Yaron JR, Zhang L, Macaulay C, McFadden G. Serpins: Development for therapeutic applications. In: *Methods in Molecular Biology*. Vol 1826. Humana Press Inc.; 2018:255-265. doi:10.1007/978-1-4939-8645-3_17

42. Law RHP, Zhang Q, McGowan S, Buckle AM, Silverman GA, Wong W, Rosado CJ, Langendorf CG, Pike RN, Bird PI, Whisstock JC. An overview of the serpin superfamily. *Genome Biol*. 2006;7(5):216. doi:10.1186/gb-2006-7-5-216

43. Schindelin J, Rueden CT, Hiner MC, Eliceiri KW. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol Reprod Dev*. 2015;82(7-8):518-529. doi:10.1002/mrd.22489

44. Li Q, Lau A, Morris TJ, Guo L, Fordyce CB, Stanley EF. A Syntaxin 1, Gαo, and N-Type Calcium Channel Complex at a Presynaptic Nerve Terminal: Analysis by Quantitative Immunocolocalization. *J Neurosci*. 2004;24(16):4070-4081. doi:10.1523/JNEUROSCI.0346-04.2004

45. MANDERS EMM, VERBEEK FJ, ATEN JA. Measurement of co-localization of objects in dual-colour confocal images. *J Microsc*. 1993;169(3):375-382. doi:10.1111/j.1365-2818.1993.tb03313.x

46. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*. 2016;11(12):2301-2319. doi:10.1038/nprot.2016.136

47. Smyth GK. limma: Linear Models for Microarray Data. In: *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer-Verlag; 2005:397-420. doi:10.1007/0-387-29362-0_23

48. Chawade A, Alexandersson E, Levander F. Normalyzer: A tool for rapid evaluation of normalization methods for omics data sets. *J Proteome Res*. 2014;13(6):3114-3120. doi:10.1021/pr401264n

49. Tyanova S, Cox J. Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. In: *Methods in Molecular Biology*. Vol 1711. Humana Press Inc.; 2018:133-148. doi:10.1007/978-1-4939-7493-1_7

50. Wei R, Wang J, Su M, Jia E, Chen S, Chen T, Ni Y. Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. *Sci Rep*. 2018;8(1):663. doi:10.1038/s41598-017-19120-0

51. Zhang X, Smits AH, Van Tilburg GBA, Ovaa H, Huber W, Vermeulen M. Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat Protoc*. 2018;13(3):530-550. doi:10.1038/nprot.2017.147

52. The Gene Ontology Resource: 20 years and still GOing strong. doi:10.1093/nar/gky1055

53. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: Tool for the unification of biology. *Nat Genet*. 2000;25(1):25-29. doi:10.1038/75556

54. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2018;47:419-426. doi:10.1093/nar/gky1038

55. Nauseef WM, McCormick SJ, Clark RA. Calreticulin functions as a molecular chaperone in the biosynthesis of myeloperoxidase. *J Biol Chem*. 1995;270(9):4741-4747. doi:10.1074/jbc.270.9.4741

56. Leach MR, Williams DB. Calnexin and Calreticulin, Molecular Chaperones of the Endoplasmic Reticulum. Published online 2013. Accessed July 29, 2020. https://www.ncbi.nlm.nih.gov/books/NBK6095/

57. Ferris SP, Jaber NS, Molinari M, Arvan P, Kaufman RJ. UDP-glucose:glycoprotein glucosyltransferase (UGGT1) promotes substrate solubility in the endoplasmic reticulum. *Mol Biol Cell*. 2013;24(17):2597-2608. doi:10.1091/mbc.E13-02-0101

58. Sakono M, Seko A, Takeda Y, Ito Y. PDI family protein ERp29 forms 1:1 complex with lectin chaperone calreticulin. *Biochem Biophys Res Commun*. 2014;452(1):27-31. doi:10.1016/j.bbrc.2014.08.041

59. Tannous A, Pisoni GB, Hebert DN, Molinari M. N-linked sugar-regulated protein folding and quality control in the ER. *Semin Cell Dev Biol*. 2015;41:79-89. doi:10.1016/j.semcdb.2014.12.001

60. Ferris SP, Kodali VK, Kaufman RJ. Glycoprotein folding and quality-control mechanisms in protein-folding diseases. *DMM Dis Model Mech*. 2014;7(3):331-341. doi:10.1242/dmm.014589

61. Kozlov G, Määttänen P, Thomas DY, Gehring K. A structural overview of the PDI family of proteins. *FEBS J*. 2010;277(19):3924-3936. doi:10.1111/j.1742-4658.2010.07793.x

62. Anelli T, Alessio M, Bachi A, Bergamelli L, Bertoli G, Camerini S, Mezghrani A, Ruffato E, Simmen T, Sitia R. Thiol-mediated protein retention in the endoplasmic reticulum: The role of ERp44. *EMBO J*. 2003;22(19):5015-5022. doi:10.1093/emboj/cdg491

63. Mezghrani A, Fassio A, Benham A, Simmen T, Braakman I, Sitia R. Manipulation of oxidative protein folding and PDI redox state in mammalian cells. *EMBO J*. 2001;20(22):6288-6296. doi:10.1093/emboj/20.22.6288

64. Zito E. PRDX4, an endoplasmic reticulum-localized peroxiredoxin at the crossroads between enzymatic oxidative protein folding and nonenzymatic protein oxidation. *Antioxidants Redox Signal*. 2013;18(13):1666-1674. doi:10.1089/ars.2012.4966

65. Xiong B, Jha V, Min JK, Cho J. Protein disulfide isomerase in cardiovascular disease. *Exp Mol Med*. 2020;52(3):390-399. doi:10.1038/s12276-020-0401-5

66. Mathias S, Wippermann A, Raab N, Zeh N, Handrick R, Gorr I, Schulz P, Fischer S, Gamer M, Otte K. Unraveling what makes a monoclonal antibody difficult-to-express: From intracellular accumulation to incomplete folding and degradation via ERAD. *Biotechnol Bioeng*. 2020;117(1):5-16. doi:10.1002/bit.27196

67. Feizi A, Österlund T, Petranovic D, Bordel S, Nielsen J. Genome-Scale Modeling of the Protein Secretory Machinery in Yeast. *PLoS One*. 2013;8(5). doi:10.1371/journal.pone.0063284

68. Sears RM, May DG, Roux KJ. BioID as a tool for protein-proximity labeling in living cells. In: *Methods in Molecular Biology*. Vol 2012. Humana Press Inc.; 2019:299-313. doi:10.1007/978-1-4939-9546-2_15

69. Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, Feldman JL, Perrimon N, Ting AY. Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol*. 2018;36(9):880-898. doi:10.1038/nbt.4201

70. Dangi AK, Sinha R, Dwivedi S, Gupta SK, Shukla P. Cell line techniques and gene editing tools for antibody production: A review. *Front Pharmacol*. 2018;9(JUN). doi:10.3389/fphar.2018.00630

71. Miranda-Hernández MP, López-Morales CA, Ramírez-Ibáñez ND, Piña-Lara N, Pérez NO, Molina-Pérez A, Revilla-Beltri J, Flores-Ortiz LF, Medina-Rivero E. Assessment of Physicochemical Properties of Rituximab Related to Its Immunomodulatory Activity. Published online 2015. doi:10.1155/2015/910763

72. Cerutti ML, Pesce A, Bès C, Seigelchifer M. Physicochemical and Biological Characterization of RTXM83, a New Rituximab Biosimilar. *BioDrugs*. 2019;33(3):307-319. doi:10.1007/s40259-019-00349-2

73. Davio K. FDA Approves Pfizer's Rituximab Biosimilar, Ruxience. AJMC | The Center for Biosimilars. Published February 23, 2019. Accessed July 17, 2020. https://www.centerforbiosimilars.com/news/fda-approves-pfizers-rituximab-biosimilar-ruxience

74. Davio K. First Rituximab Biosimilar, Truxima, Launches in the United States. AJMC | The Center for Biosimilars. Published November 7, 2016. Accessed July 17, 2020. https://www.centerforbiosimilars.com/news/first-rituximab-biosimilar-truxima-launches-in-the-united-states

75. Lu RM, Hwang YC, Liu IJ, Lee CC, Tsai HZ, Li HJ, Wu HC. Development of therapeutic antibodies for the treatment of diseases. *J Biomed Sci*. 2020;27(1):1-30. doi:10.1186/s12929-019-0592-z

76. Han X, Aslanian A, Yates JR. Mass spectrometry for proteomics. *Curr Opin Chem Biol*. 2008;12(5):483-490. doi:10.1016/j.cbpa.2008.07.024

77. Rupp O, MacDonald ML, Li S, Dhiman H, Polson S, Griep S, Heffner K, Hernandez I, Brinkrolf K, Jadhav V, Samoudi M, Hao H, Kingham B, Goesmann A, Betenbaugh MJ, Lewis NE, Borth N, Lee KH. A reference genome of the Chinese hamster based on a hybrid assembly strategy. *Biotechnol Bioeng*. 2018;115(8):2087-2100. doi:10.1002/bit.26722

78. Li S, Cha SW, Heffner K, Hizal DB, Bowen MA, Chaerkady R, Cole RN, Tejwani V, Kaushik P, Henry M, Meleady P, Sharfstein ST, Betenbaugh MJ, Bafna V, Lewis NE. Proteogenomic Annotation of Chinese Hamsters Reveals Extensive Novel Translation Events and Endogenous Retroviral Elements. *J Proteome Res*. 2019;18(6):2433-2445. doi:10.1021/acs.jproteome.8b00935

79. Wei Huang D, Sherman BT, Lempicki RA. SURVEY AND SUMMARY Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2008;37(1):1-13. doi:10.1093/nar/gkn923

80. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57. doi:10.1038/nprot.2008.211

81. Suzuki M, Yamanoi A, Machino Y, Kobayashi E, Fukuchi K, Tsukimoto M, Kojima S, Kohroki J, Akimoto K, Masuho Y. Cleavage of the interchain disulfide bonds in rituximab increases its affinity for FcγRIIIA. *Biochem Biophys Res Commun*. 2013;436(3):519-524. doi:10.1016/j.bbrc.2013.05.137

82. Lee CW, Stankowski JN, Chew J, Cook CN, Lam Y-W, Almeida S, Carlomagno Y, Lau K-F, Prudencio M, Gao F-B, Bogyo M, Dickson DW, Petrucelli L. The lysosomal protein cathepsin L is a progranulin protease. *Mol Neurodegener*. 2017;12(1):55. doi:10.1186/s13024-017-0196-6

83. Butler VJ, Cortopassi WA, Argouarch AR, Olivia Pierce M, Vohra M, Oses-Prieto JA, Gao F, Caballero B, Chand S, Seeley W, Miller BL, Coppola G, Burlingame AL, Ashrafi K, Cuervo M, Jacobson MP, Kao AW. Title: C. elegans granulins promote an age-associated decline in protein homeostasis via 1 lysosomal protease inhibition. doi:10.1101/472258

84. Jaarsma D, Hoogenraad CC. Cytoplasmic dynein and its regulatory proteins in Golgi pathology in nervous system disorders. *Front Neurosci*. 2015;9(OCT):397. doi:10.3389/fnins.2015.00397

85. Liang C, Chiang AWT, Hansen AH, Arnsdorf J, Schoffelen S, Sorrentino JT, Kellman BP, Bao B, Voldborg BG, Lewis NE. A Markov model of glycosylation elucidates isozyme specificity and glycosyltransferase interactions for glycoengineering. *Curr Res Biotechnol*. 2020;2:22-36. doi:10.1016/j.crbiot.2020.01.001

86. Miwa HE, Koba WR, Fine EJ, Giricz O, Kenny PA, Stanley P. Bisected, complex N-glycans and galectins in mouse mammary tumor progression and human breast cancer. *Glycobiology*. 2013;23(12):1477-1490. doi:10.1093/glycob/cwt075

87. Lee KH, Lee J, Bae JS, Kim YJ, Kang HA, Kim SH, Lee SJ, Lim KJ, Lee JW, Jung SK, Chang SJ. Analytical similarity assessment of rituximab biosimilar CT-P10 to reference medicinal product. *MAbs*. 2018;10(3):380-396. doi:10.1080/19420862.2018.1433976

88. Byrne B, Donohoe GG, O'kennedy R. Sialic acids: carbohydrate moieties that influence the biological and physical properties of biopharmaceutical proteins and living cells. *Drug Discov Today*. 2007;12. doi:10.1016/j.drudis.2007.02.010

89. Cha HM, Lim JH, Yeon JH, Hwang JM, Kim D Il. Co-overexpression of Mgat1 and Mgat4 in CHO cells for production of highly sialylated albumin-erythropoietin. *Enzyme Microb Technol*. 2017;103:53-58. doi:10.1016/j.enzmictec.2017.04.010

90. Goto Y, Niwa Y, Suzuki T, Uematsu S, Dohmae N, Simizu S. N-glycosylation is required for secretion and enzymatic activity of human hyaluronidase1. *FEBS Open Bio*. 2014;4:554-559. doi:10.1016/j.fob.2014.06.001

91. Goh JB, Ng SK. Impact of host cell line choice on glycan profile. *Crit Rev Biotechnol*. 2018;38(6):851-867. doi:10.1080/07388551.2017.1416577

92. Wang Z, Zhu J, Lu H. Antibody glycosylation: impact on antibody drug characteristics and quality control. *Appl Microbiol Biotechnol*. 2020;104(5):1905-1914. doi:10.1007/s00253-020-10368-7

93. Pierpont TM, Limper CB, Richards KL. Past, present, and future of Rituximab-The world's first oncology monoclonal antibody therapy. *Front Oncol*. 2018;8(JUN):163. doi:10.3389/fonc.2018.00163