UNIVERSITY OF TURKU

Optimizing methods to study host cell signaling responses upon receptor docking of *Clostridioides difficile* binary toxin

Amanda Barsk

Physiology and Genetics
Department of Biology, Faculty of Science
Master's Thesis
Credits: 30op

Supervisors:
Arto Pulliainen
Tiina Henttinen

September 2025
Turku

Master's Thesis

Subject: Biology, Physiology and Genetics
Author: Amanda Barsk
Title: Optimizing methods to study host cell signaling responses upon receptor docking of *Clostridioides difficile* binary toxin
Supervisors: Arto Pulliainen, Tiina Henttinen
Number of pages: 64 pages + 7 appendix pages
Date: x.x.2025

Clostridioides difficile infection (CDI), an urgent threat to human health, is caused by the *Clostridioides difficile* bacterium that produces TcdA and TcdB toxins. Some strains additionally produce *C. difficile* binary toxin (CDT), which is thought to increase the bacterium's virulence. The CDT toxin has two subunits, CDTa (enzymatic subunit) and CDTb (cell binding subunit). CDTb docks to the lipolysis-stimulated lipoprotein receptor (LSR), enabling CDTa endocytosis, but it also has CDTa-independent functions, that are poorly known on the molecular level. Identification of LSR interacting proteins upon CDTb binding is expected to advance the molecular understanding of CDI.

Protein interactions can be studied with proximity dependent biotinylation methods combined with mass spectrometry (proximity labeling-MS). This thesis aimed to optimize methods for proximity labeling-MS to identify proteins interacting with LSR after CDTb docking. Experimental objectives included the following: Elucidation of protein characteristics of LSR and CDTb utilizing *in silico* tools. Transfection of cells to achieve MAC-tagged LSR expression. Simulation of canonical oligomerization of CDTb via trypsinization. Verification of interaction between LSR-MAC and CDTb on the plasma membrane of cells.

Experiments were conducted on Human Embryonic Kidney cells (HEK293T). The semi-stable expression of LSR-MAC partly on the plasma membrane was confirmed by cell fractionation, western blot, and confocal microscopy. Oligomerization of CDTb via trypsinization was confirmed with Size Exclusion Chromatography-Multi Angle light scattering (SEC-MALS). CDTb and LSR-MAC interaction was confirmed with flow cytometry. These optimized methods will enable utilization of proximity labeling-MS to study intracellular protein interactions following CDTb toxin docking to LSR.

Keywords: *C. difficile*, CDT, LSR, proximity labeling-MS

Pro gradu -tutkielma

Tutkinto-ohjelma, oppiaine: Biologia, Fysiologian ja genetiikan linja
Tekijä: Amanda Barsk
Otsikko: Optimizing methods to study host cell signaling responses upon receptor docking of *Clostridioides difficile* binary toxin
Ohjaajat: Arto Pulliainen, Tiina Henttinen
Sivumäärä: 64 sivua + 7 liitesivua
Päivämäärä: x.x.2025

Clostridioides difficile infektiota (CDI), joka on uhka ihmisterveydelle, aiheuttaa *Clostridioides difficile* bakteeri tuottamillaan TcdA ja TcdB toksiineilla. Jotkut kannat tuottavat lisäksi *C. difficile* binaaritoksiinia (CDT), jonka uskotaan lisäävän bakteerin virulenssia. CDT toksiinissa on kaksi alayksikköä, CDTa (entsymaattinen yksikkö) ja CDTb (soluun sitoutuva yksikkö). CDTb sitoutuu lipolyysis-stimuloituun lipoproteiini reseptoriin (LSR) mahdollistaen CDTa:n endosytoosin, mutta CDTb:llä on myös itsenäisiä funktioita, jotka tunnetaan heikosti molekyylitasolla. CDTb:n sitoutuessa LSR:ään, reseptorin kanssa interaktoivien proteiinien tunnistamisen uskotaan edistävän molekyylistä ymmärrystä CDI:stä.

Proteiini interaktioita voidaan tutkia proksimiteetti riippuvaisilla biotinylaatio metodeilla yhdistettynä massaspektrometriaan (proksimitetti merkitseminen-MS). Tämä turkielma pyrki optimoimaan metodeja proksimiteetti merkitseminen-MS:ää varten, jotta LSR:n kanssa interaktoivat proteiinit CDTb:n sitoutumisen yhteydessä voitaisiin tunnistaa. Kokeelliset tavoitteet olivat seuraavat: 1) LSR ja CDTb proteiinien ominaisuuksien selvittäminen hyödyntäen *in silico* työkaluja, 2) solujen transfektointi saavuttaaksemme MAC-tägätyn LSR:n ekspression, 3) kanonisen CDTb oligomerisaation simulointi trypsinisaatiolla, ja 4) LSR-MAC:in ja CDTb:n välisen interaktion varmistaminen solujen plasma membraanilla.

Kokeet suoritettiin Human Embryonic Kidney –soluilla (HEK293T). LSR-MAC:in semi-stabiili ekspressio osaksi solujen plasma membraanilla varmistettiin solufraktioinilla, western blotilla, ja konfokaali mikroskopialla. CDTb:n oligomerisaatio trypsinoinilla varmistettiin size exclusion chromatography-multi angle light scattering (SEC-MALS) –metodilla. CDTb ja LSR-MAC interaktio varmistettiin flow cytometrialla. Nämä optimoidut metodit tulevat mahdollistamaan proksimiteetti merkisteminen-MS:n hyödyntämisen solunsisäisten proteiini interaktioiden tutkimiseksi CDTb toksiinin sitoutuessa LSR reseptoriin.

Keywords: *C. difficile*, CDT, LSR, proksimiteetti merkitseminen-MS

List of abbreviations

ADP	Adenosine diphosphate
AIM	Autoinduction medium
APD	Autoprotease domain
AP-MS	affinity purification and mass spectrometry
ATP	adenosine triphosphate
BAP	biotin acceptor peptide
BirA	biotin ligase found in *E. Coli*
BSA	Bovine serum albumin
Caco-2	human colorectal adenocarcinoma cells
CDC	Centers for Disease Control and Prevention
CDI	*Clostridioides difficile* infection
CDT	*Clostridioides difficile* binary toxin
CdtLoc	*Clostridioides difficile* binary toxin locus
cdtR	response regulator gene
CD44	cluster of differentiation 44
CoA	Coenzyme A
CROP	Combined repetitive oligopeptides
CSPG4	chondroitin sulfate proteoglycan 4
DABCO	1,4-diazabicyclo[2.2.2]octane
DAPI	4′,6-diamidino-2-phenylindole
DCA	deoxycholic acid
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
DRBD	Delivery/Receptor binding domain
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
FFA	free fatty acid
FZD	Frizzled
GTD	Glycosyl transferase domain
GFP	Green Fluorescent protein
GTP	guanosine triphosphate
HA	hemagglutinin
HCT116	human colon cancer cell line
HEK	Human Embryonic Kidney cells
HeLa	Henrietta Lacks
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horseradish Peroxidase
iFBS	inactivated fetal bovine serum
IHC	Immunohistochemisrty
$InsP_6$	inositol hexakisphosphate
K	Clone
LCA	lithocholic acid

LCT	Large clostridial toxin
LDL	low-density lipoprotein
LSR	Lipolysis-stimulated lipoprotein receptor
ltk	leukocyte tyrosine kinase
MAC	Multiple Approaches Combined
MEK1	Mitogen-activated protein kinase kinase
MALS	Multi-angle light scattering
MS	Mass spectrometry
NaKATPase	Sodium Potassium Pump
NFDM	Nonfat dry milk
Ni-NTA	Nickel-Nitriloacetic acid
NP-40	Nonidet P-40
PaLoc	pathogenicity locus
PBS	Phosphate-buffered saline
PMSF	Phenylmethylsulfonyl fluoride
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulfate
SEC	Size Exclusion chromatography
sGAG	sulfated glycosaminoglycans
SH2	src homology 2
Strep	Streptavidin
SV40	Simian virus 40
TBST	Tris-buffered saline with Tween-20
TcdA	*Clostridioides difficile* toxin A
TcdB	*Clostridioides difficile* toxin B
tcdC	anti-sigma factor
tcdE	holing
tcdL	endolysin
tcdR	alternate sigma factor
TcsH	hemorrhagic toxin
TcsL	lethal toxin
VLDL	very-low-density-lipoprotein
WGA	Wheat Germ Agglutinin

CONTENTS

1 INTRODUCTION

1.1 Clostridioides difficile infection (CDI)

The gram-positive *Clostridioides difficile* (previously known as *Clostridium difficile*) is an obligate anaerobic bacterium first identified in 1935, that can exist in a spore (metabolically inactive) or vegetative (metabolically active) form. The highly resistant spores can infect the host by transmission via the fecal-oral route, after which germination of the spores into vegetative cells happens in the intestine. The vegetative form is the actively growing and dividing stage of the bacterium, and it is responsible for toxin production and thus causing the disease. The survival of the vegetative cells is dependent on the host's anaerobic environment. In the intestine, proliferation and production of virulence factors TcdA and TcdB, disrupts the intestinal epithelium integrity. Additionally, *C. difficile* binary toxin (CDT) produced by some strains is thought to enhance the toxicity of TcdA and TcdB by augmenting adherence of the bacteria to cells (Lourenço et al., 2025; Smits et al., 2016; Martínez-Meléndez et al., 2022). Disruptions of the intestinal epithelium by the toxins lead to tissue damage, which causes inflammation (Porliotopoulou et al., 2024).

Patients with *Clostridium difficile* infection (CDI) may be asymptomatic or show clinical manifestations ranging from mild diarrhea to fulminant colitis or colonic perforation (Czepiel et al., 2019). Increased risk of CDI is associated with antibiotics, especially broad-spectrum antimicrobials, as well as long-term hospitalization. Old age and medications that reduce stomach acid production are also risk factors of CDI. Treatment of CDI usually requires conventional and alternative treatment approaches, as recurrence rates post treatment are high (20–25%). In addition to the high rate of recurrence, there are several other challenges associated with the treatment of CDI, such as the easy transmission of *C. difficile* and the bacterium's resistance to multiple antibiotics (Spigaglia et al. 2018). *C. difficile* is classified as one of the top five urgent threats to human health by the Centers for Disease Control and Prevention (CDC) for its high transmission rates, difficult treatment and manageability, lethality and concurrent economic burden (Porliotopoulou et al.,2024).

Changes in the gut microbiome have a substantial role in the effective germination of *C. difficile* spores in the human gut. While primary bile acids that are conjugated with taurine and glycine can promote the germination of *C. difficile* spores, secondary bile acids like deoxycholic acid (DCA) and lithocholic acid (LCA) produced by a healthy gut microbiome interfere with the

germination of *C. difficile* and prevent its growth, thus also preventing CDI. A dysbiotic microbiome on the other hand is not as effective in inhibiting *C. difficile*. For this reason, treatment with antibiotics is a predisposing factor in *C. difficile* infection. As antibiotics alter the gut microbiome, oral uptake of *C. difficile* spores leads to a higher risk of germination and proliferation. Successful germination and toxin production by the bacteria causes damage to the epithelium of the colon, inflammation, and necrosis, the typical CDI symptoms. (Aktories 2022; Wahlström et al., 2024.)

1.1.1 *Clostridioides difficile* toxins A and B

C. difficile produces TcdA and TcdB toxins as its main virulence factors. The toxins belong to the large clostridial toxin (LCT) family, which also encompasses hemorrhagic toxin (TcsH) and lethal toxin (TcsL) of *Clostridium sordellii*. The toxins are glucosyltransferases that act by irreversibly modifying guanosine triphosphate (GTP) -associated Ras and Rho enzymes (GTPases) inside the host cell. This inactivation of regulatory GTPases eventually leads to cell detachment from its environment due to actin cytoskeleton destruction and cell rounding. (Porliotopoulou et al., 2024.)

Toxins TcdA and TcdB are encoded by *tcdA* and *tcdB* genes, respectively. These genes are located in the 19.6-kb DNA region called the pathogenicity locus (PaLoc), which also includes four other genes: *tcdR, tcdE, tcdC* and *tcdL* (Figure 1A). Proteins tcdR and tcdC encoded by *tcdR* and *tcdC* genes respectively, function as transcriptional regulators of *tcdA* and *tcdB.* The TcdR protein is an alternative sigma factor, and a positive regulator of *tcdA* and *tcdB*. It also functions as an activator of its own transcription, leading to an even higher expression of large clostridial toxins. Environmental changes, such as temperature fluctuations, and availability of nutrients, modulate TcdR activity via influencing the protein's activators and repressors. In contrast, the TcdC protein, is a membrane-bound anti-sigma factor and negatively regulates *tcdA* and *tcd*B. Proteins encoded by *tcdE* and *tcdL* have functions that are thought to be associated with TcdA and TcdB toxin secretion. TcdE protein functions as a holin, while TcdL protein might function as an endolysin. Both of these proteins may influence the virulence of CDI via their possible roles in toxin secretion. (Kordus et al., 2022; Porliotopoulou et al., 2024.)

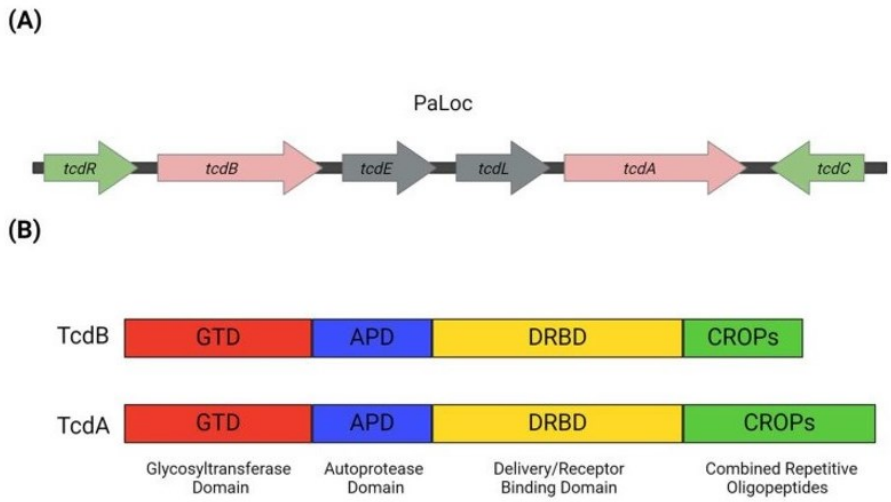


Figure 1 TcdA and TcdB toxin genetic and structural information. (A) Pathogenicity locus (PaLoc) of *Clostridioides difficile*. The proteins encoded by regulatory genes *tcdR* and *tcdC*, shown here in green arrows, influence the transcription of *tcdA* and *tcdB*; TcdR is a positive regulator, while tcdC is a negative one. The genes *tcdA* and *tcdB* encode TcdA– and TcdB toxins and are shown in pink arrows. Represented here in grey arrows, genes *tcdE* and *tcdL* encode proteins involved in toxin secretion: *tcdE* encodes a holin, while *tcdL* encodes an endolycin. The direction of arrows represents directionality of transcription. **(B)** TcdA– and TcdB toxins are divided into the glycosyltransferase domain (GTD; red) that inactivates Rho GTPase family members via glucose molecule transfer; the autoprotease domain (APD; blue) with functions in autoproteolytic cleavage as well as toxin processing; the delivery/receptor-binding domain (DRBD; yellow), responsible for interaction with the cell surface receptors as well as the release of GTD into the cytosol of host cells; and the combined repetitive oligopeptides (CROPs; green) partaking in toxin binding to the receptor. (Porliotopoulou et al., 2024.)

The large clostridial toxins, TcdA and TcdB, have a sequence identity of 47%, and both contain four functional domains: the glucosyltransferase domain (GTD), an autoprotease domain (APD), the delivery/receptor-binding domain (DRBD), and the combined repetitive oligopeptides domain (CROPs) (Figure 1B). Toxin binding to its respective receptor is thought to be mediated via CROPs through glycan binding interactions. The N-terminal GTD of the toxins is responsible for inactivating host Rho GTPases via the transfer of glucose. The GTD is bound to the APD until inositol hexakisphosphate ($InsP_6$) activates APD to be able to autocleave itself, releasing GTD. The DRBD is responsible for interaction with the cell surface receptors as well as the release of GTD-APD structure from the endosome into the cytosol of host cells. (Porliotopoulou et. al., 2024.)

The mechanism of action for the TcdA and TcdB toxins includes multiple steps (Figure 2). Firstly, the toxins bind to their cell surface receptors with the help of the mentioned DRBD as well as the CROPs. TcdA is thought to bind to sulfated glycosaminoglycans (sGAGs), glycans, and possibly to low-density lipoprotein receptor family members. TcdB also binds to glycans in addition to other receptors, for example the chondroitin sulfate proteoglycan 4 (CSPG4), Nectin 3, and a few of the Frizzled receptors (FZD1, FZD2, FZD7). The binding induces

receptor-mediated endocytosis, which leads to internalization of TcdA and TcdB into host cells. The internalization process of TcdA involves PACSIN2, a clathrin-independent entry mechanism, while TcdB utilizes a clathrin-mediated endocytosis mechanism. In the endosome, the low pH caused by proton influx leads to the pore formation of the DRBD of the toxins. Through this pore, the GTD-APD structure can translocate into the cytosol. Next, APD is activated by $InsP_6$ leading to autoproteolytic cleavage of the structure and finally to the release of GTD. The free GTD can then inactivate RhoGTPases via glycosylation. This inactivation leads to changes in cell functions, resulting in the collapse of tight junctions, and even apoptosis of the cell. (Kordus et al., 2022; Porliotopoulou et al., 2024.)

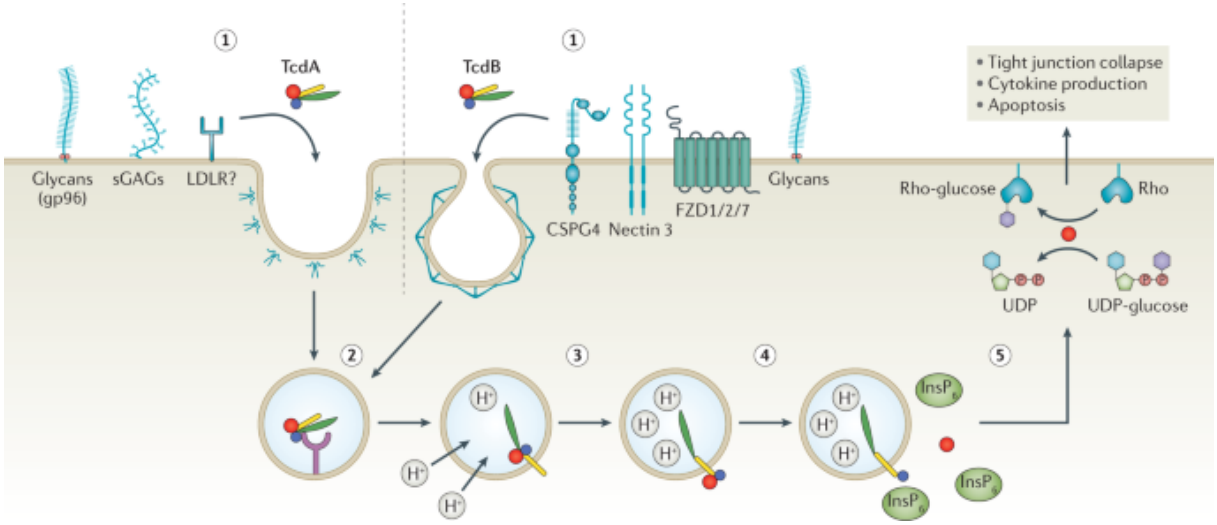


Figure 2. Mechanism of action for large clostridial toxins TcdA and TcdB. (1) TcdA and TcdB toxins bind to their respective cell surface receptors. Glycans, sulfated glycosaminoglycans (sGAGs) and possibly low-density lipoprotein receptors (LDLR) are thought to function as docking targets for TcdA. TcdB binds to glycans as well, but additionally to chondroitin sulfate proteoglycan 4 (CSPG4), Nectin 3, Frizzled 1 (FZD1), FZD2, and FZD7. (2) After toxins bind to their respective cell surface receptors, TcdA is endocytosed into the cell via PACSIN2, a clathrin-independent mechanism, and TcdB via clathrin-mediated endocytosis. (3) The low pH in the endosome results in pore formation of the toxins through the endosomal membrane. (4) Through this pore, the glucosyltransferase domain (GTD) and the autoprotease domain (APD) of the toxins are translocated to the cytosol, where inositol hexakisphosphate (InsP6) induces autoproteolysis of APD and subsequent release of GTD. (5) Released GTD inactivates RhoGTPases, which leads to changes in cell function, such as tight junction collapse, cytokine production, and apoptosis of the host cell. (Kordus et al., 2022.)

1.1.2 *Clostridioides difficile* binary toxin

In addition to the Rho-glycosylating toxins TcdA and TcdB, some *C. difficile* strains produce a third toxin, *Clostridioides difficile* binary toxin (CDT), which is thought to enhance the toxicity of the other two (Martínez-Meléndez et al. 2022). The production of CDT is considered a hallmark of hypervirulent *C. difficile* strains (Porliotopoulou et al., 2024).

CDT is a binary toxin that consists of two subunits: an enzymatic component CDTa and a binding component CDTb. CDTa has adenosine diphosphate (ADP) -ribosyltransferase activity, while CDTb interacts with the toxins cell surface receptor, the lipolysis stimulated lipoprotein receptor (LSR) (Aktories 2022; Papatheodorou et al., 2024). CDTa and CDTb are encoded in an operon by two genes: *cdtA* and *cdtB* (Figure 3A). These two genes as well as a response regulator gene, *cdtR*, are located on the binary toxin locus (CdtLoc). *CdtR* is a positive transcriptional regulator for CDT and is located upstream of the *cdtAB* operon (Smits et al., 2016).

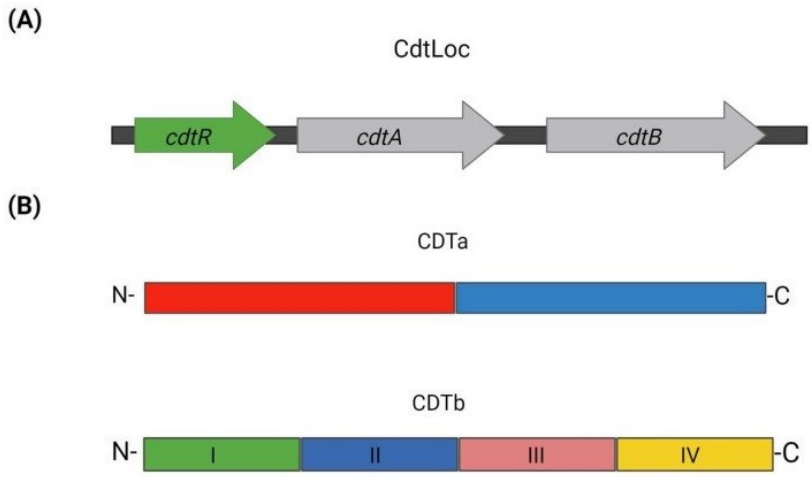


Figure 3. CDT toxin genetic and structural information. (A) CDT locus (CdtLoc) of *C. difficile*. Depicted here in grey arrows, the *cdtA* and *cdtB* genes encode the CDT toxin subunits CDTa and CDTb respectively. Depicted here in a green arrow, the regulatory gene, *cdtR,* is a positive regulator of *cdtA* and *cdtB*. The direction of arrows represents directionality of transcription. **(B)** CDTa and CDTb are divided into domains based on the function of these regions. CDTa has two domains: the N-terminal domain is responsible for interaction with CDTb, while the C-terminal domain has ADP-ribosyltransferase activity. CDTb is divided into four domains: domain I functions as the activation domain, domain II has roles in membrane insertion and pore formation, domain III is responsible for oligomerization and pore formation, and domain IV functions as the receptor binding domain of CDTb. (Porliotopoulou et al., 2024.)

The subunits of CDT are divided into domains with different functions (Figure 3B). The CDTa subunit is comprised of two domains, the N-terminal and the C-terminal domain. The N-terminal domain matures via proteolytic cleavage, after which it can interact with the CDTb subunit. The C-terminal domain of CDTa has ADP-ribosyltransferase activity, which enables CDTa to enact its functions in the cytosol of the host cell (Porliotopoulou et al. 2024). The CDTb subunit on the other hand is comprised of four domains (I – IV). At the N-terminus of the protein, domain I functions as the activation domain of CDTb. It is followed by domain II that has a role in membrane insertion as well as in the pore formation of the oligomerized CDTb.

Domain III enables the oligomerization and formation of the CDTb pore. At the C-terminus of CDTb, domain IV functions as the receptor binding domain (Papatheodorou et al., 2024).

The binding of the CDTb subunit to its cell surface receptor LSR, involves oligomerization (Figure 4). During the binding of CDTb monomer to LSR, CDTb is activated by proteolytic cleavage leading to heptamerization of the toxin. It is still unclear if this activation occurs before, during, or after the toxin binds to the receptor. After oligomerization of the CDTb subunit, the enzymatic component CDTa, can bind to the heptamer, followed by endocytosis of the toxin-receptor complex into the cell. The underlying intracellular protein interactions of this endocytosis are still unclear. The low pH environment of the endosome enables the oligomerized CDTb heptamer to form a pore, through which CDTa can translocate into the cytosol of the cell. The enzymatic CDTa component can then ADP-ribosylate a major cytoskeletal protein, the G-actin monomer. CDTa facilitates an ADP-ribose addition from NAD^+ onto G-actin at arginine-177. The addition leads to ADP-ribosylated G-actin not only being unable to polymerize but also functioning as a plus-end actin-capping protein to inhibit polymerization of unmodified actin as well. This leads to G-actin subunits being unable to polymerize correctly into F-actin polymers. Depolymerization of F-actin leads to the release of septins, which are GTP-binding proteins considered as the unofficial fourth components of the cytoskeleton (Aktories 2022). Septins have a key role in microtubule organization into networks, and consequently in guiding and modifying microtubules as well as motor-dependent transport (Mostowy & Cossart 2012). The release of septins from F-actin results in the formation of microtubule-based protrusions constructed from another cytoskeletal protein, tubulin. Through these protrusions, septins can affect the transport of recycling Rab-5 and Rab-11 associated vesicles containing fibronectin. Fibronectin is an extracellular matrix (ECM) molecule that promotes and facilitates cell-ECM adhesion (Sicari et al., 2023). The modified actin cytoskeleton and newly formed microtubule protrusions enable the trafficking of the fibronectin containing vesicles from the basolateral membrane of the cell to the apical membrane. The release of fibronectin from vesicles at the CDT-induced microtubule protrusions located at the apical membrane leads to enhanced attachment of *C. difficile* bacteria to cells (Aktories 2022). While the CDT subunits have known interrelated mechanisms of action, the CDTb subunit has also been shown to decrease cell viability and increase cell rounding without CDTa present (Landenberger et al., 2021). CDTb alone has cytotoxic effects and is known to lead to F-actin redistribution in the cell. However, the molecular mechanisms behind these effects are still unknown (Ernst et al., 2021).

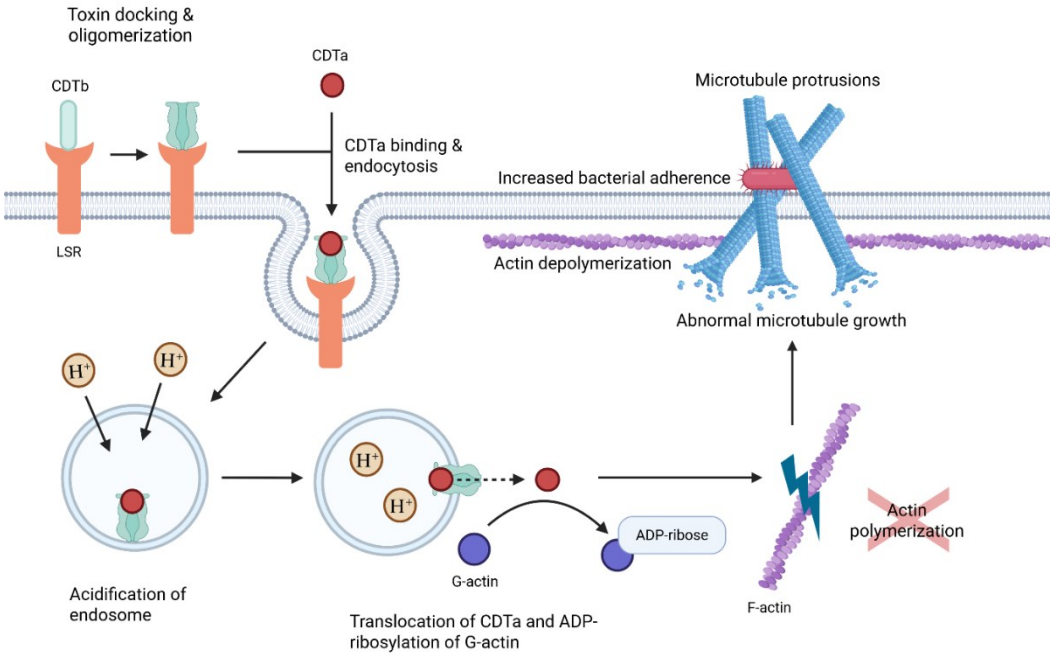


Figure 4. Mechanism of action for CDT. The *Clostridioides difficile* binary toxin subunit b (CDTb) binds to the lipolysis stimulated lipoprotein receptor (LSR) on the cell surface. Toxin docking results in LSR accumulation into lipid rafts, and oligomerization of the CDTb toxin into a heptameric prepore. The *C. difficile* binary toxin subunit a (CDTa) can bind to the prepore, after which the toxin-receptor complex is endocytosed into the cell. Due to proton influx into the endosome, the pH lowers and enables CDTb prepore to transition into a transmembranous pore, through which the CDTa subunit can translocate into the cytosol. CDTa then ADP-ribosylates monomeric G-actin inhibiting its polymerization into F-actin. Depolymerization of cytoskeletal F-actin leads to the release of septins and consequent growth of abnormal microtubule protrusions at the apical surface of the cell. Microtubule protrusions enhance the adherence of *C. difficile* onto host cells. In addition to enabling CDTa translocation into the cell, CDTb has been shown to bind to LSR and have cytotoxic effects on its own. However, the CDTa-independent functions of CDTb are still unclear on the molecular level. (Kordus et al., 2021; Created in https://BioRender.com, 02.06.2025; Ernst et al., 2021.)

1.2 Lipolysis-stimulated lipoprotein receptor

The lipolysis-stimulated lipoprotein receptor (LSR) is a member of the lipoprotein receptor family, which includes receptors involved in the transport of lipids across the plasma membrane of cells. Other roles associated to lipoprotein receptors include mediation of signaling between cells, modulation of cellular trafficking, as well as the removal of macromolecules, such as proteases, from the cell surface and the extracellular space of cells (Dieckmann et al., 2012). LSR binds low-density lipoprotein (LDL) in the presence of free fatty acids (FFAs), as well as apoB, apoE, and triglyceride-rich lipoprotein such as chylomicrons and very-low-density lipoproteins (VLDLs) (Yen et al., 1999).

LSR functions as a host-receptor for actin-ADP-ribosylating iota-like toxins, a subfamily of the clostridiceae family. It's a single-pass transmembrane protein expressed in the liver, intestines and multiple other tissues. It has an extracellular N-terminal part that is bound by CDTb, and an intracellular cysteine-rich region with a cytoplasmic tail. The extracellular part of the receptor has an uncharacterized region at the N terminus, in addition to an immunoglobulin-like domain of the V-type. These two extracellular subregions could interact with CDTb either co-operatively or independently. (Papatheodorou et al., 2024; Kanda et al., 2023.)

In the liver, LSR functions as a lipoprotein receptor that facilitates chylomicron remnant uptake into parenchymal cells from the blood (Gerding et al., 2013). In other tissues, LSR has a role in the proper formation of blood-brain barrier during embryogenesis, and in the formation of tricellular tight junctions in epithelial and endothelial cells (Higashi & Furuse, 2022; Fabien et al., 2015). The overexpression of LSR has also been reported in different cancer types. LSR has been linked to cancer cell proliferation and invasion, cancer progression, as well as metastasis (Papatheodorou et al., 2024; Kanda et al., 2023).

Even though the endocytosis of CDT is still not completely understood, it has been shown that the binding of CDTb to LSR leads to the accumulation of LSR in lipid rafts. This accumulation also occurs without the oligomerization of CDTb, as the LSR clustering has been shown to occur due to the binding of only the receptor binding domain of CDTb to LSR (Papatheodorou et al., 2013). These rafts might also include membrane proteins that assist or trigger the endocytosis of the ligand-receptor complex. For example, lipid-raft associated protein, CD44 (cluster of differentiation 44), has been shown to be a critical factor for CDT uptake into cells in mice, but specifics of the interaction with LSR have not been identified (Hemmasi et al. 2015; Wigelsworth et al. 2012).

1.3 Proximity labeling techniques in proteomics

Proximity labeling techniques combined with mass spectrometry (proximity labeling-MS) provide an approach for analyzing proteins and their interactions in a desired space. Proximity labeling is based on covalently tagging neighboring proteins via reactive radicals created by enzymes. Isolation of the tagged proteins can then be achieved with affinity purification (AP), and mass spectrometry (MS) can be utilized to further analyze and identify the proteins. In addition to strong protein-protein interactions, proximity labeling-MS is especially useful in

identifying weak or transient interactions, which are often missed with canonical affinity purification MS approaches. (Bosch et al. 2021.)

Several proximity labeling techniques are utilized for elucidation of proteins involved in cellular processes. One of these is called proximity-dependent biotin identification (BioID), which allows identification of proteins in close vicinity to a protein of interest, whether they directly interact or are simply in the same cellular environment (Figure 5). This is achieved by fusing the protein of interest with a mutant form of the *E. coli* biotin ligase enzyme BirA (BirA*) and introducing the fusion protein into cells. BirA functions as a catalyst in the conversion of unreactive biotin to reactive biotinyl-5'-AMP, and the mutation (R118G) of the enzyme's active site reduces its affinity towards biotinyl-5'-AMP. The release of the highly reactive compound into the vicinity leads to biotinyl-5'-AMP binding non-specifically to the lysine residues of nearby proteins in the radius of approximately 10nm. These biotinyl-5'-AMP tagged, i.e. biotinylated, proteins can then be isolated and identified by mass spectrometry to understand possible protein interactions with the protein fused with BirA*. (Bosch et al. 2021; Liu et al. 2018.)

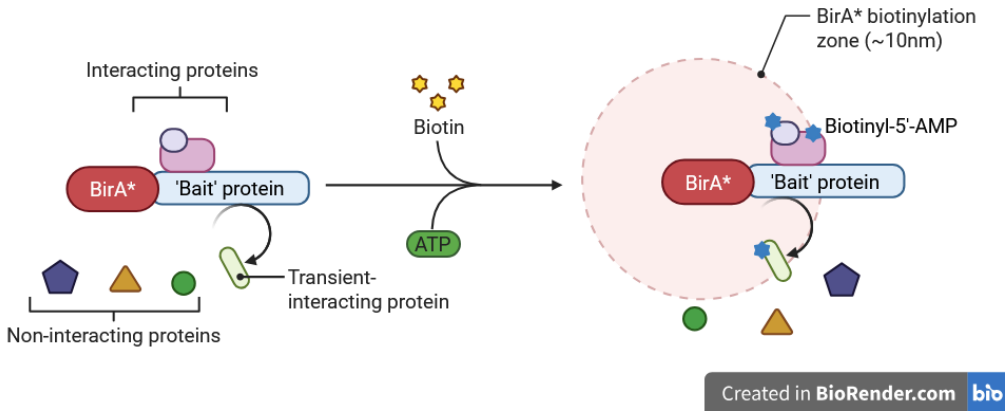


Figure 5. Illustration of the BioID method. The BirA enzyme attained from *E. Coli* has been mutated (BirA*) to enable promiscuous labeling of proximal proteins. When BirA* is fused to a protein of interest, i.e. 'bait', it can be utilized to study possible protein-protein interactions happening between the 'bait' and proximal proteins. In the presence of biotin and ATP, BirA* creates a highly reactive compound, biotinyl-5'-AMP (blue), which then binds to the lysine residues of nearby (~10nm) proteins, thus 'tagging' them with the compound. These 'tagged' proteins can later be purified and then identified to understand possible interactions happening between tagged proteins and the 'bait'. (Bosch et al., 2021; Created in biorender, https://app.biorender.com, 02.06.2025.)

While proximity dependent biotinylation with the BioID method combined with mass spectrometry enables the study of possible protein interactions, the BioID construct has its drawbacks. The construct is quite large (approximately 36kDa) and has slow labeling kinetics. The large size may interfere with the function or trafficking of the tagged protein, or even with

the interactions and correct localization of the fusion protein. The slow labeling kinetics necessitate hours-long biotin incubation of cells for sufficient protein biotinylation to occur. To avoid these, a compound known as ultraID has been engineered by Kubitz et al. (2022). The ultraID construct is much smaller (under 20kDa) than the BioID, so it should not interfere with protein localization or function. Additionally, the labeling kinetics are faster, and background activity low. This means that efficient labeling of proximal proteins can be achieved faster than with the BioID construct, while still having relatively low biotinylation before biotin incubation. (Kubitz et al., 2022.) In this thesis, the ultraID construct has been utilized for proximity dependent biotinylation.

For analyzing proteins and their interactions via proximity labeling-MS, isolation and identification of biotinylated proteins can be achieved with affinity purification and mass spectrometry (AP-MS). An additional tag, e.g. epitope tag or antibody, in BirA* -fused proteins enables the purification of the tag as well as the protein associated with it. Common epitope tags used in this method include HA, GFP, and Strep. The Strep tag enables native competitive elution utilizing biotin, and produces high protein purity in physiological purification conditions. For these reasons, it has become a staple in affinity purification proteomics and is used in this thesis as well, in its fused form of 2xStrep (StrepII). The single-step Strep AP-MS method can be used in combination with the mentioned BioID method by utilizing a single construct, the Gateway®-compatible Multiple Approaches Combined-tag (MAC-tag) (Liu et al., 2018). As the MAC-tag enables one step affinity purification of the construct and the tagged proteins (Liu et al. 2020), it has been utilized in this thesis as well.

1.4 Aims of the study

Clostridioides difficile infection (CDI) is currently one of the most significant hospital-acquired infections and the leading cause of health-care-associated infective diarrhea. The development of CDI is associated with the use of antibiotics, especially broad-spectrum penicillins. The virulence of *C. difficile* is thought to be increased by the *C. difficile* binary toxin (CDT), which makes this toxin an interesting topic of research. To possibly prevent hypervirulent CDI, it is important to understand the cellular mechanisms behind it. It has been shown that the lipolysis-stimulated lipoprotein receptor (LSR) functions as the host-receptor for CDT, but the specific downstream signals of this interaction are yet unknown.

This thesis aims to optimize methods to enable the investigation of downstream signals caused by CDTb docking to LSR. For this aim, characterization of LSR-MAC expressing cells and confirmation of CDTb and LSR-MAC interaction are essential goals, and will be pursued with the help of the following experimental objectives.

This thesis has the following experimental objectives:

1) *In silico* studies of expression, structure, and possible functions of LSR and CDTb

2) Transfection of HEK293T cells with a construct that creates functional MAC-tagged LSR

3) Simulation of canonical cleavage and oligomerization of CDTb via trypsinization

4) Confirmation of the interaction between CDTb and MAC-tagged LSR on the plasma membrane of transfected cells.

2 MATERIALS AND METHODS

2.1 In silico tools

In silico review of LSR and CDTb were performed prior to laboratory experiments of this thesis to help understand how best to optimize methods for proximity labeling-MS. The expression, structure and possible function of the proteins were elucidated utilizing the *in silico* tools listed in Table 1. The results of this *in silico* review were additionally employed to interpret the results of practical experiments as well as their possible meaning in a biological context.

Table 1. *In silico* tools utilized in this thesis.

In silico tool	Reference	Website
The Human Protein Atlas	Uhlén et al., 2015	https://www.proteinatlas.org/
Uniprot	The UniProt Consortium, 2025	https://www.uniprot.org/
Protparam	Gasteiger et al., 2005	https://web.expasy.org/protparam/
Scansite 4.0	Obenauer et al., 2003	https://scansite4.mit.edu/#home

In order to study intracellular protein-protein interactions caused by CDTb toxin docking to LSR utilizing proximity labeling-MS, it was essential to perform experiments on cells that have a low– or moderate endogenous LSR expression. Consequently, *in silico* tool, Human Protein Atlas (Uhlén et al., 2015; proteinatlas.org) was used to select the optimal cell line for experiments. To understand the biological context of CDTb and LSR interaction, the expression of LSR across different tissues and cells was also analyzed. Expression of LSR in tissues was regarded at the RNA and protein level, while expression in single cells was examined at the RNA level. From the results attained from Human Protein Atlas, the RNA expression (normalized transcript per million, nTPM) had been comprised of values from internally generated Human Protein Atlas (HPA) RNA-seq data and RNA-seq data from the Genotype-Tissue Expression (GTEx) project. Protein expression (score) of tissue groups were based on "true" protein expression estimates from a knowledge-based annotation, described in Assays & annotation. In the case of the gastrointestinal tract, a more specific immunohistochemistry (IHC) image of the colon was additionally examined.

To optimize the interaction between CDTb and LSR, the canonical cleavage and oligomerization of CDTb (Aktories 2022) was considered. The protein sequence of canonical

CDTb was examined using Uniprot (The UniProt Consortium, 2025) to assess possible cleavage sites leading to oligomerization during toxin docking. As recombinant CDTb containing residues 43-876 is used in all practical experiments of this thesis, the corresponding sequence from Uniprot was input into the ProtParam *in silico* tool (Gasteiger et al., 2005) to later fluorescently label the CDTb protein for the toxin-receptor interaction assay.

Uniprot was also utilized to estimate the molecular weights of proteins corresponding to the antibodies used in western blot analyses. For LSR, Uniprot was used to view protein structure and localization in the cell, and the attained protein sequence was entered into another *in silico* tool, Scansite 4.0 (Obenauer et al., 2003), for further study.

To elucidate possible protein-protein interactions between LSR and intracellular proteins after CDTb toxin docking, Scansite 4.0 was used. Plausible protein interactions could be interpreted as phosphorylation events (Nishi et al., 2011), and thus the potential sites for phosphorylation, e.g. domains and motifs, were elucidated from the LSR protein utilizing the "Motif Scan" program at Scansite. All mammalian motifs were searched at a high stringency and predicted domains were explored.

2.2 Transfection of cells with ultraID constructs

HEK293T cells were transfected with ultraID constructs, creating cells expressing MAC-tagged proteins that would enable proximity labeling. Transient transfection was first utilized to confirm the MAC-tagged proteins' expression and successful biotinylation of proteins. In addition, the biotin concentration for proximity labeling was also optimized. More extensive experiments were then performed on stable HEK293T cell lines (gift from Ida Shoultz, University of Turku). By using stably transfected cells, the aim was to reduce the number of steps required and the potential sources of error due to multiple steps, thereby optimizing proximity labeling-MS method.

2.2.1 Cell culturing

Based on *in silico* analysis, moderate endogenous LSR expression of HEK293T made them suitable for the experiments. The cells are easy to maintain, have reliable cell growth, and high transfectivity (Lin et al., 2014; https://www.cytion.com/). HEK293T cells are Human Embryonic Kidney cells that were developed in the early 1970s by transfection with sheared

adenovirus type 5 DNA fragments and genetically modified to express the large antigen T allele of simian virus 40 (SV40). This allele expression enables plasmids containing SV40 origin of replication to replicate in transfected cells, leading to amplified results of transient transfection (Tan et al. 2021; Lin et al. 2014).

HEK293T cell lines (ATCC code: CRL-3216) were cultured in T75 flasks at 37°C, 5% CO_2 and 90% humidity, from frozen stocks. Cell culture medium was Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% inactivated fetal bovine serum (iFBS), 1x Penicillin-Streptomycin, 2mM L-glutamine and 25mM HEPES. Cell lines were regularly split twice a week at ~80% confluency with the dilution ratio being 1:10. To ensure detachment of the adherent HEK293T cells during subculture, they were incubated with Trypsin-EDTA (0,05%) for 1 minute before resuspending them in 10ml of fresh media.

In some experiments, stably transfected HEK293T cell lines were utilized. These cells were grown as parental HEK293T cells, except that the culture medium was supplemented with hygromycin B (100µg/ml) to maintain the selection of transfected cells.

2.2.2 Transient transfection

The expression of MAC-tagged proteins and function of the ultraID tag was first confirmed with transiently transfected cells. HEK293T cells were transiently transfected with three different constructs, creating three cell lines: LSR, CAAX and pcDNAmock. The LSR cells were transfected with plasmids containing constructs to create LSR-MAC proteins, CAAX cells with plasmids containing constructs to create GFP– and MAC tagged CAAX proteins (CAAX-GFP-MAC), and pcDNAmock cells with empty plasmids as a mock control (Figure 6). MAC tags included epitope tags for affinity purification (StrepII), protein expression confirmation (anti-HA), proximity labeling (UltraID), membrane targeting (CAAX), and fluorescent detection (GFP). PcDNAmock cells functioned as controls throughout the experiments of this thesis, while LSR cells were utilized as the treatment group. For eventual proximity labeling-MS, CAAX cells will be utilized the control, thus the expression and function of CAAX-GFP-MAC in cells was also analyzed in this thesis. Plasmids had been constructed, propagated, extracted, and purified from *E. coli* cultures before the experiments of this thesis by Mika Savisalo (University of Turku) with Nucleobond® Midiprep kit, according to the manufacturer's protocol (Appendix 1).

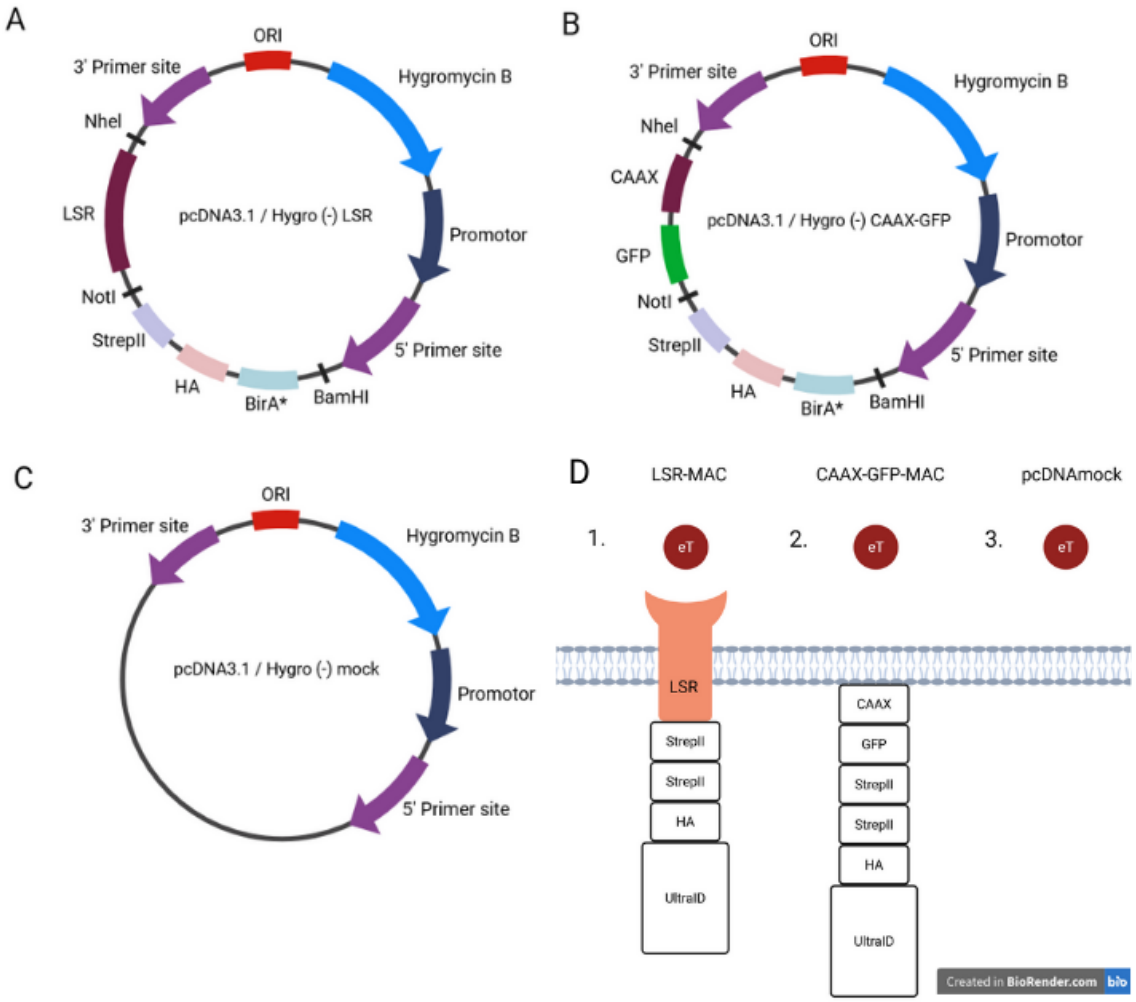


Figure 6. Plasmids and consequent MAC-tagged proteins. The plasmids contained sequences for creating cell lines expressing MAC-tagged proteins to be used in proximity labeling-MS. Sequences included the selection marker Hygromycin B, StrepII for affinity purification purposes, HA for confirming expression in cells with the primary antibody anti-HA, BirA* (UltraID) to enable biotinylation of proximal proteins, the LSR receptor, GFP for fluorescent marking, and CAAX for membrane targeting. The pcDNAmock group of cells were transfected with an empty plasmid to be used as a mock control. Mentioned NheI, NotI and BamHI are restriction sites. **(A)** Visualization of pcDNA3.1/Hygro (-) LSR plasmid for creating cell lines expressing the LSR-MAC protein. **(B)** Visualization of the pcDNA3.1/Hygro (-) CAAX-GFP plasmid for creating cell lines expressing the CAAX-GFP-MAC protein. **(C)** Visualization of pcDNA3.1/Hygro (-) mock plasmid for creating a mock control cell line. **(D)** MAC-tagged proteins created by transfecting HEK293T cells with visualized plasmids. (Created in Biorender, https://app.biorender.com, 16.07.2025.)

For the experiment, $1x10^6$ HEK293T cells were seeded in 10cm cell culture dishes and incubated for 24 h before transfection with either LSR, CAAX, or pcDNAmock plasmids. For each transfection mixture, 12,5ng/µl of DNA was resuspended in TE0.1X buffer (pH 7.05), followed by 10s of vortexing. After this, $CaCl_2$/HEPES (pH 7.05), was added in final concentrations of 0,25M and 12,5mM respectively, and the mixture was vortexed again for 30s. Finally, Hepes Buffered Saline (HBS) (pH 7.05) was added to the mixture dropwise, during vortexing, in final concentrations of 135mM NaCl, 1.15mM $Na_2HPO_4*2H_2O$, 25mM HEPES, 5,5mM Glucose, 0,5 mM KCl, and the mixture was vortexed for an additional 10s after HBS

addition. Each prepared mixture was then added dropwise to cell plates with fresh DMEM so that there was 4µg of DNA in each plate. After transfection, plates were incubated 24h before cell media was exchanged either to fresh media without biotin or to media containing 50µM or 200µM biotin.

2.2.3 Western blot for transiently transfected cells

To check MAC-tagged protein expression and ultraID function in transfected cells, western blot was utilized. 48 h after transfection and 24 h after biotin media exchange, cells were washed with Phosphate Buffered Saline (PBS), and then lysed with RIPA lysis buffer containing 400mM NaCl, 50mM Tris-HCl (pH 8.5), 0,5 % sodium-deoxycholate, 0,1 % Sodium dodecyl sulfate (SDS), 1 % Nonidet P-40 (NP-40), 1mM EDTA (pH 8.0) and 1x Pierce Protease and Phosphatase Inhibitor (cat#A32961). After a 20-minute incubation, cells were scraped and collected into 1,5ml microcentrifuge tubes and centrifuged at +4°C, 21000 rcf for 20 minutes. The supernatant was collected and stored at –20°C. Cell dishes and 1,5ml microcentrifuge tubes were handled on ice during the whole lysis process.

Protein concentrations were estimated with UV–1280 UV-Visible Spectrophotometer (Shimadzu). Samples were first centrifuged at +4°C, 21 000 rcf for 15 minutes, after which they were mixed with 1x Protein Assay Dye Reagent Concentrate (BioRad, cat#500-0006) in 10mm path length cuvettes and incubated for ~2 minutes before absorbance was measured at the wavelength of 595nm. Measurements were done in duplicates, and the average was calculated. A standard curve of bovine serum albumin (BSA) was used to calculate protein concentrations based on average absorbance measurements.

For SDS-PAGE, cell lysates containing approximately 7µg of total cellular protein per sample were diluted with 1xLAEMMLLI and 10% 2-Mercaptoethanol –loading dye and heated at 75°C for 10 minutes. The protein samples were run on 10% acrylamide gels (Appendix 2) at 70V for 30minutes, and 110V for the following 100 minutes. The proteins were transferred onto nitrocellulose membranes (Sartorious Stedim Biotech) via wet transfer in 20V for 19 hours at +4°C. Running buffer and Transfer buffer contents can be found in Appendix 3.

After transfer, to observe loading of proteins, membranes were stained with Ponceau S and imaged before the stain was washed off with two 10-minute washes with 1x Tris-buffered saline with Tween® 20 Detergent (1xTBST). Nitrocellulose membranes were then blocked with a buffer containing 5% skimmed milk powder (NFDM) in 1xTBST, for three hours at +4°C.

Following membrane blocking, immunostaining was done with antibodies according to Table 2 at +4°C for four hours in the case of streptavidin-HRP, and overnight (22 h) in the case of anti-hemagglutinin (anti-HA). After primary antibody incubation, the membranes were washed twice for 10 minutes with 1xTBST, and secondary antibody staining with anti-mouse was performed to anti-HA membrane according to Table 2 at +4°C for 2 hours. This was followed by three 7-minute washes with 1xTBST, after which membranes were imaged utilizing chemiluminescent substrates (WesternBright Quantum, Advansta, K-12042-D20; Luminol & Peroxidase) and LAS-4000 Luminescent Image Analyzer (FujiFilm).

Table 2. Antibodies used in western blot for transient transfection experiments.

Antibody	Host	Manufacturer, catalogue number	Dilution used
Streptavidin-HRP	*Streptomyces avidinii*	Thermo Scientific, RPN1231V-2mL	1 : 1000
Anti-HA	Mouse	Merck, H3663-200ul	1 : 1000
Anti-β-Actin	Mouse	Santa Cruz Biotechnology, sc-47778 HRP	1 : 2000
Goat Anti-Mouse Ig, Human ads-HRP	Goat	AH Diagnostics /SouthernBiotech, 1010-05	1 : 2500

After imaging, the anti-HA blot was stripped according to mild stripping protocol (Appendix 4), the membrane was blocked with 5% NFDM as before, and probed using anti-β-actin antibody (Table 2) for one hour at RT. The blot was washed and imaged as previously described.

2.2.4 Stably transfected cell lines

Stably transfected HEK293T cell lines created by Shoultz (University of Turku) were revived and cultured as described in chapter 2.2.1. Stable expression of MAC-tagged proteins was confirmed with anti-HA western blot as described in 2.2.3. One clone (K1/2/3) from all revived cell lines were chosen for further experiments; pcDNAmock K1 and CAAX K2 –cell lines were utilized for cell fractionation, flow cytometry and confocal microscopy experiments, while LSR K2 cells were utilized for cell fractionation and flow cytometry. Additionally, as two clones of the revived stably transfected LSR cells had low expression of LSR-MAC, new selections for creating stably transfected cell lines were done.

To create new stably transfected LSR cell lines, HEK293T cells were seeded onto plates and transfected with construct for the LSR-MAC protein as described in 2.2.2. After 24 h, the media was changed into fresh DMEM, and after 48 h, cells were moved to four 6-well plates, each well seeded with $2x10^5$ cells. Selection was begun simultaneously with varying hygromycin B concentrations to find optimal concentration for selecting stably transfected cells. Six different hygromycin B concentrations were used: 0, 5, 10, 50, 100, and 200µg/ml. Cells were cultured in 6-well plates with media exchanged every 2–3days for 20 days. Cell viability was observed with a microscope every few days, and images were taken on days six and nineteen. After 20 days, LSR cells from the wells containing 100µg/ml and 200µg/ml of hygromycin B were moved to T25 flasks for expansion, creating six new LSR clones (LSR K4–K9).

To confirm that newly created LSR clones (K4-K9) expressed LSR-MAC, western blot with anti-HA was once again performed as described in 2.2.3. The previously revived LSR K2 clone was also checked again for LSR-MAC expression to confirm stability of expression with increased passaging of the cell line. Previously revived pcDNAmock K1 clone, and normal HEK293T cells were utilized for the experiment to function as controls. Based on the results, LSR K5-K9 were used to create new cell stocks, while LSR K4 was cultured as described in chapter 2.2.1 and utilized in confocal microscopy experiments.

2.3 Localization of MAC-tagged proteins

After the expression of MAC-tagged proteins and successful biotinylation of proteins in transfected HEK293T cells had been confirmed, localization of the MAC-tagged proteins in cells were analyzed further. To enable the interaction between LSR-MAC and CDTb on the surface of cells, as well as confirm that CAAX-GFP-MAC will function as a reliable control during proximity labeling-MS, localization of MAC-tagged proteins to the plasma membrane needed to be confirmed. To achieve this, protein expression in the different cellular compartments was analyzed using western blot analysis and confocal microscopy.

2.3.1 Cell fractionation

To detect the localization of the MAC-tagged proteins in cells, a cell fractionation assay was performed on stably transfected LSR K2, CAAX K2, and pcDNAmock K1 HEK293T cells. Approximately $15x10^6$ plated cells from each cell line were lysed with hypotonic lysis buffer containing 20 mM HEPES (pH 8.0), 2.5 mM $MgCl_2$, 1 mM EDTA (pH 8.0), 1 mM DL-

Dithiothreitol (cat#43819-5G), 1x Pierce Protease and Phosphatase Inhibitor (cat#A32961), and 25 U / mL benzonase (cat#70664-10KUN), and scraped into 2ml microcentrifuge tubes. Samples were incubated in +4°C on spin for one hour, after which cells were further broken down by driving samples through 25G and 27G needles altogether 20 times. A portion from each sample was taken to function as a total lysate sample while the rest was centrifuged at +4°C, 800 rcf for 15 minutes. The supernatant was collected for further centrifugation, while the pellet was washed once with lysis buffer and stored as a nuclear sample. The collected supernatant was centrifuged at +4°C, 6000 rcf for 15 minutes, after which the supernatant was again collected, and the pellet washed before storing it as a mitochondrial sample. The collected supernatant was centrifuged at +4°C, 21000 rcf for 30 minutes, after which the pellet functioned as a plasma membrane sample, while the supernatant functioned as a cytosolic sample. All samples were stored at –20°C. Samples were handled on ice during the whole lysis process.

Western blot with the total lysate, nuclear, membrane, and cytosolic samples from all cell lines, as well as the mitochondrial sample from LSR K2 cells, was performed to verify the localization of the MAC-tagged proteins in cells. To prepare loading samples, lysis buffer was added to the collected pellets, and all samples were dyed with 1xSDS and 5% 2-Mercaptoethanol –loading dye. Gel run, wet transfer, Ponceau S staining, and membrane blocking were performed as previously described in 2.2.3.

Immunostaining was performed with overnight incubations of primary antibodies (Table 3) at +4°C, and 1 hour RT incubations of secondary antibodies (Table 4). Immunostaining with anti-Na,K-ATPase was done on a membrane that had been stripped according to mild stripping protocol (Appendix 4). Membrane washes and imaging were performed as described in 2.2.3.

Table 3. Information on primary antibodies used in western blot for cell fractionation experiment.

Primary antibody	Host	Manufacturer, catalogue number	Dilution and buffer	Purpose
MEK1/2	Mouse	Cell Signaling Technologies, CST-4694S	1 : 1000 in 5% NFDM	Cytosolic control
Anti-HA	Mouse	Merck, H3663	1 : 1000 in 5% NFDM	MAC-tagged protein identifier
Lamin B1	Rabbit	Cell Signaling Technologies, CST-17416S	1 : 1000 in 5% BSA	Nuclear control
Anti-Na,K-ATPase	Rabbit	Abcam, ab76020	1 : 1000 in 5% NFDM	Plasma membranous control
Anti-LSR	Rabbit	Atlas antibodies, HPA007270	1 : 1000 in 5% NFDM	Endogenous LSR & MAC-tagged protein identifier

Table 4. Information on secondary antibodies used in western blot for cell fractionation experiment.

Secondary antibody	Host	Manufacturer, catalogue number	Dilution	Corresponding primary antibody
Goat Anti-Mouse Ig, Human ads-HRP	Goat	AH Diagnostics / SouthernBiotech, 1010-05	1 : 2500	anti-HA & MEK1/2
Goat Anti-Rabbit Ig, Human ads-HRP	Goat	AH Diagnostics / SouthernBiotech, 4010-05	1 : 2500	Lamin B1, Anti-LSR & Anti-Na,K-ATPase

2.3.2 Confocal microscopy

To further study the subcellular localization of MAC-tagged proteins in cells, confocal microscopy was used. The experiment was done in two groups to separately analyze LSR-MAC and CAAX-GFP-MAC localization in cells. PcDNAmock K1 cells functioned as controls in both groups.

LSR K4, CAAX K2, and pcDNAmock K1 cells (50 000 cells/coverslip) were seeded onto glass coverslips, which had been coated with 50µg/ml collagen I (cat# A10483-01) during a one-hour incubation. Cells were grown for 24 hours before fixing them with 10% formaldehyde during a 30-minute incubation in RT. Fixed cells were stored in PBS in a 24-well plate covered with parafilm at +4°C until cells were stained.

Before antibody staining, the cells were washed thrice with PBS for 5 minutes, followed by perforation of cell membranes with 0,1% Triton X-100 with a 10-minute incubation. Coverslips were then washed again with PBS thrice for 5 minutes, after which they were incubated in 0,2% BSA in PBS for 5 minutes.

To analyze LSR-MAC localization in LSR K4 cells, anti-HA primary antibody and anti-mouse-Alexa Fluor647 secondary antibody were used. Moreover, cell membranes were stained with Alexa Fluor488-tagged WGA and nucleus with DAPI (Table 5).

To analyze CAAX-GFP-MAC localization in CAAX K2 cells, the colocalization of the GFP signal with stained cell membrane was analyzed. Primary antibody anti-Na,K-ATPase, and corresponding secondary antibody anti-rabbit-Alexa Fluor647 were used to stain cell membranes, while the nucleus of cells was again stained with DAPI (Table 5).

Staining of cells was done according to Table 5, with primary antibody incubation coinciding with WGA staining, and secondary antibody incubation with DAPI staining. Both incubations were done in RT protected from light and lasted 30 minutes, followed by three 5-minute PBS washes. Finally, coverslips containing samples were mounted onto microscope glass slides with Mowiol (cat#475904-100GM) and 2,5% DABCO (1,4-diazabicyclo[2.2.2]octane, cat#D27802-25G). Slides were stored in the dark at +4°C until imaging.

Table 5. Information on reagents used for cell staining in confocal microscopy experiment.

Reagent	Host, Dilution	Manufacturer, catalogue number	Purpose
Anti-HA	Mouse, 1:1000	Merck, H3663	MAC-tagged protein identifier
Anti-Na/K-ATPase	Rabbit, 1:1000	Abcam, ab76020	Plasma membrane staining
Wheat Germ Agglutinin-Alexa488 (WGA)	Wheat germ, 1µg/ml	Thermo Fisher Scientific, W11261	Membrane staining
Anti-mouse-Alexa647	Goat, 1:1000	Invitrogen Life Technologies, A21235	Alex Fluor647-tagged secondary antibody for anti-HA
Anti-rabbit-Alexa647	Goat, 1:1000	Invitrogen Life Technologies, A21245	Alex Fluor647-tagged secondary antibody for anti-Na,K-ATPase
4′,6-diamidino-2-phenylindole (DAPI)	–, 1µg/ml	Santa Cruz Biotechnology, sc-3598	Nucleus staining

Imaging of samples was done with the 3i Marianas CSU-W1 spinning disk (50µm pinholes) confocal microscope, using a 63x Zeiss Plan-Apochromat oil immersion objective, Hamamatsu sCMOS Orca Flash4.0– and Photometrics Evolve 10 MHz Back Illuminated EMCCD cameras, and the SlideBook 6 software. Imaging was done utilizing a 405nm excitation wavelength laser for DAPI, 488nm for WGA and GFP, and 640nm for secondary antibodies tagged with Alexa Fluor 647. The Z-step interval in all samples was 0,27µm.

Images from confocal microscopy were processed and analyzed with the ImageJ program (Schneider et al., 2012). To compare samples and remove background signals, the minimum pixel signal in relevant channels was modified. In the 640nm channel, the minimum signal was set to 150 in all samples utilized to analyze LSR-MAC localization. In the 488nm and 640nm channels, minimum signals were set to 100 and 150 respectively for all samples utilized to analyze CAAX-GFP-MAC localization.

2.4 Proteolytic cleavage and consequent oligomerization of CDTb

CDTb is known to canonically be clipped at the N-terminal followed by oligomerization during its binding to the LSR receptor on the cell surface, but the timing of the proteolytic cleavage and consequent oligomerization of the protein are still unknown (Aktories, 2022). CDTb binding to LSR-MAC on the surface of our transfected cells was first optimized by simulation of the canonical proteolytic cleavage and oligomerization of the protein by trypsinization of CDTb. Proteolytic cleavage was analyzed with western blot and Size Exclusion Chromatography (SEC). Oligomerization caused by simulated proteolytic cleavage was analyzed with Size Exclusion Chromatography and Multi-Angle Light Scattering (SEC-MALS). Recombinant CDTb protein containing residues 43-876 is used in all practical experiments of this thesis.

2.4.1 Western blot and Size Exclusion Chromatography (SEC) of CDTb

Recombinant CDTb protein samples containing residues 43-876 were obtained from Mika Savisalo (University of Turku). Prior to this thesis, the protein had been cloned into a pET15b plasmid, then transformed and expressed in BL21(DE3) *E. coli* cells for 24h in one liter of Auto Induction Medium (AIM) Terrific Broth at RT at 180rpm, and purified with Nickel-Nitriloacetic acid –affinity purification (Ni-NTA) and Size Exclusion Chromatography (SEC) with 20mM HEPES, 100mM NaCl, pH 8,0 –buffer.

To optimize the proteolytic cleavage of CDTb, the protein samples were incubated with trypsin in varying ratios and western blot was used to confirm that cleavage was achieved. Trypsin and CDTb were added in ratios of 0:1; 1:100; 1:20; 1:10; 1:5, and the volume of samples was kept constant by adding PBS. Samples were incubated for 45 minutes at 37°C, after which the reaction was stopped with Phenylmethanesulfonyl Fluoride (PMSF, cat#P7626-5G) in final concentration of 1mM.

Samples were then dyed with 1xSDS and 5% 2-Mercaptoethanol and divided into two groups; One group of samples was heated at 95°C for 10 minutes, and the other group of samples utilized as is. All samples were loaded onto a 1,5mm thick 7,5 % acrylamide gel (Appendix 2), and then ran and stained with Ponceau S as described in 2.2.3.

The membrane was blocked with 5% NFDM for 3,5h at +4°C, and incubated with Anti-CDTb antibody (1:2000 dilution) (Exalpha, ACdBTSB-HRP) for 2,5 hours at +4°C. Membrane washes and imaging were done as described in 2.2.3.

After optimization, a 1:5 ratio of trypsin to CDTb was used in size exclusion chromatography (SEC). The sample was prepared as before, and proteolytic cleavage was analyzed with SEC on an Äkta Pure 25 System (Cytiva) with a HiPrep 16/60 Sephacryl S-400 HR (Cytiva) –column. As an eluent, 20mM HEPES 100mM NaCl (pH 8.0) buffer was used, and signals were measured with a UV detector at 280nm. Collected fractions, as well as a control sample of trypsinated CDTb (1:5 ratio) without SEC, were prepared for gel run with loading dye and heating as before. Samples were run on 7,5 % acrylamide gels and western blot was performed as above except the anti-CDTb antibody was used at a dilution of 1:5000.

2.4.2 Size Exclusion Chromatography-Multi-Angle Light Scattering (SEC-MALS)

To confirm and study the degree of oligomerization of CDTb after proteolytic cleavage caused by trypsinization, Size Exclusion Chromatography with Multi-Angle Light scattering (SEC-MALS) was performed. Size exclusion chromatography (SEC) was used to separate different sized proteins in the sample, and the multi-angle light scattering (MALS) to determine molecular weights of analytes.

SEC-MALS was done with an untrypsinated and trypsinated (ratio of 1:5) CDTb sample. SEC separation was done in RT using a Superdex® 200 Increase 10/300 GL (Cytiva) –column, and 20mM HEPES 100mM NaCl (pH 8.0) –buffer. Samples were injected in 100µl volume with the flow rate of 0,75ml/min. BSA in the concentration of 1mg/ml was prepared in 20mM HEPES 100mM NaCl (pH 8.0) –buffer, and used as a standard before running CDTb samples. MALS analysis was performed with Wyatt Technologies, miniDAWN TREOS II and Optilab T-rEX dRI detector. Data collection and processing was done with Astra 7.3.0 (Wyatt technologies) software.

2.5 Interaction between CDTb and MAC-tagged LSR

Once the expression and localization of the MAC-tagged LSR were confirmed, and the oligomerization of CDTb protein was achieved, the interaction between CDTb and LSR-MAC proteins was analyzed by flow cytometry. As the proteolytic cleavage of CDTb is not

necessarily required for LSR docking (Chandrasekaran & Lacy, 2017), both oligomeric (trypsinated) and monomeric (untrypsinated) samples of CDTb were used in the experiment. The CDTb samples were first labelled with Alexa Fluor™ 647 NHS Ester (Thermo Fisher Scientific) according to Nanotemper Protein Labeling Kit RED-NHS 2nd Generation –protocol (Appendix 5). The extinction coefficient of CDTb was attained from ProtParam (Gasteiger et al., 2005), and excess dye was removed with PD MiniTrap G-10 columns (Cytiva).

After fluorescent labeling of CDTb, the flow cytometry experiment was performed. LSR K2 and pcDNAmock K1 cells ($2{,}5x10^5$ cells/cell line) were stimulated either with Alexa Fluor 647-labelled monomeric CDTb or with Alexa Fluor 647-labelled oligomeric CDTb in the concentrations of 100µg/ml, or left untreated. Samples were incubated for 30 minutes protected from light, centrifuged at 300rcf, at +4°C for 4 minutes, followed by three PBS washes before resuspension of the pellet in 500µl of PBS. Samples were kept on ice during the whole process to enable binding but prevent internalization of the CDTb protein. Results were analyzed with NovoCyte™ Flow Cytometer (ACEA Biosciences Inc.). The red laser (640nm) was used to excite Alexa Fluor 647 –labelled CDTb and the emitted fluorescence was detected using the 675/30nm bandpass filter. Analysis of data was done with the NovoExpress v1.6.2 software and FlowJo™ single cell analysis software v10 (BD Life Sciences). The experiment was later repeated with triplicate samples with 10µg/ml CDTb stimulations of cells.

3 RESULTS

3.1 In silico identifications

To understand the biological context of our results, i.e., the cells and tissues that the CDTb toxin can potentially target in the human body, as well as optimize the cell line selection for our experiments, the CDTb toxin and its cell surface receptor LSR, were studied with *in silico* tools: the Human Protein Atlas (Uhlén et al., 2015; proteinatlas.org), Uniprot (The UniProt Consortium, 2025), and Scansite 4.0 (Obenauer et al., 2003).

3.1.1 LSR is expressed in multiple tissues across the human body

Human Protein Atlas (HPA) was used to determine in which tissues and cell types the cell surface receptor LSR for CDTb toxin is expressed. HPA provided detailed information on the expression of the LSR protein in various human tissues both at mRNA and protein level (Figure 7A). The results showed that LSR is expressed in multiple tissues across the human body, and that LSR is most abundant at the protein level in the digestive tract and at the RNA level in the liver and gallbladder. In the gastrointestinal tract, LSR expression is highest in stomach and colon tissues (Figure 7B). To further assess LSR expression in the gastrointestinal tract, a colonic immunohistochemical (IHC) image was analyzed. As shown in Figure 7C, the epithelial cells, including goblet- and glandular cells of the human colon have high LSR expression, while endothelial cells have moderate LSR expression.

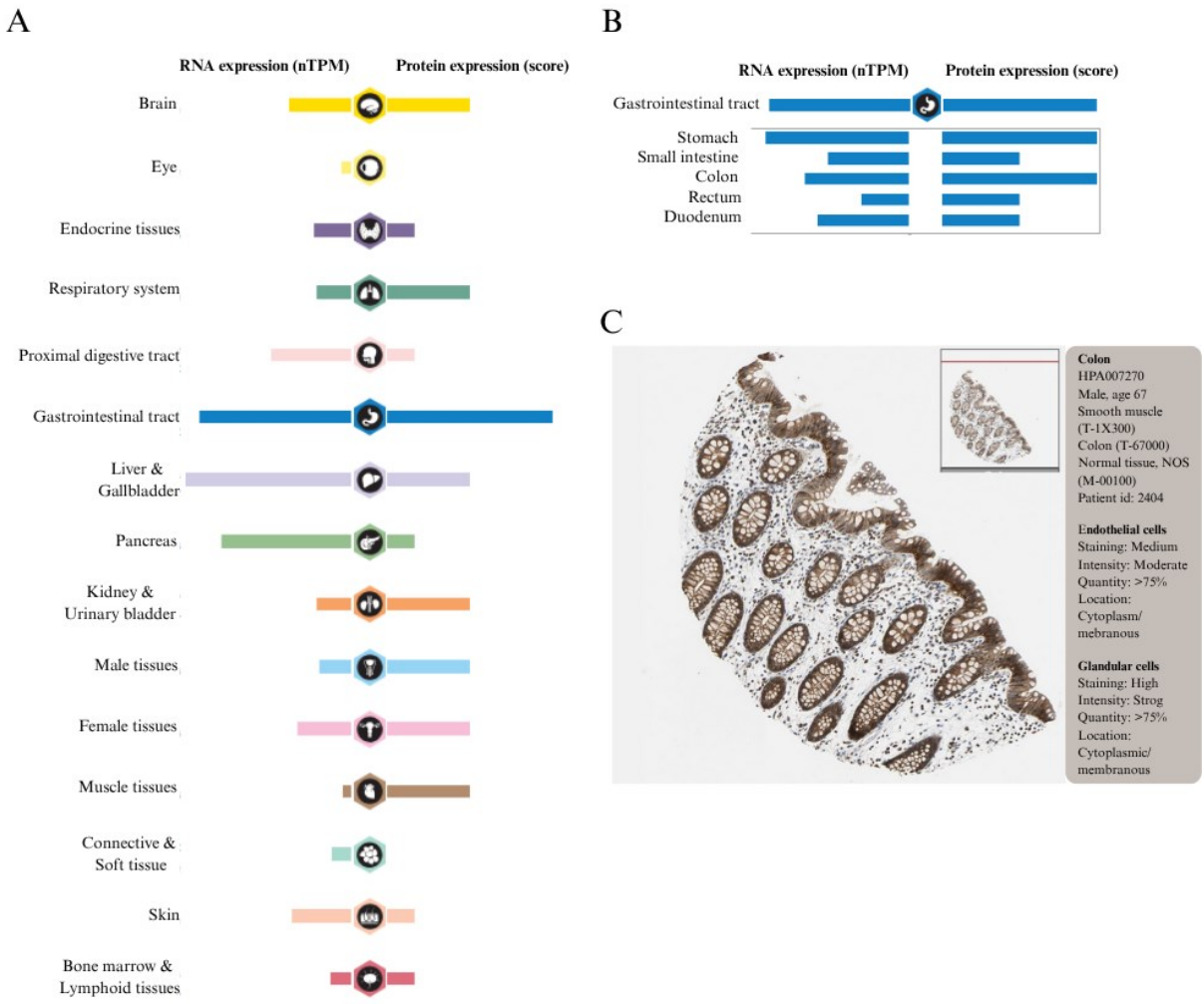


Figure 7. Overview of Lipolysis stimulated lipoprotein receptor (LSR) expression in humans. (A) Expression of LSR in different tissues is shown at the RNA and protein level. The analyzed tissues are grouped according to the functional features they have in common, e.g. Gastrointestinal tract group (blue) includes the stomach, small intestine, colon, rectum and duodeneum, and each group has it's own colour. The bars depict the highest expression score in each tissue group. **(B)** The RNA and protein expression of LSR in the human gastrointestinal tract. **(C)** Immunohistochemistry (IHC) image of the colon of a human male. Antibody staining in cell types (endothelial cells and glandular cells) is outlined as not detected, low, medium, or high, considering intensity of staining and fraction of stained cells (quantity) utilizing conventional immunohistochemistry profiling of the selected tissue. Location of the protein is mentioned in the case of both cells. Antibody (HPA007270) used in staining, as well as patient and tissue information, are listed on the right. (Human Protein Atlas, Uhlén et al. 2015; proteinatlas.org, modified 12.05.2025.)

Single cell expression of LSR in the human body was also analyzed utilizing HPA. In the colon, paneth cells, distal enterocytes, and intestinal goblet cells have the highest LSR expression at the RNA level (Figure 8). Other cells expressing LSR in the colon include enteroendocrine cells, undifferentiated cells, and T-cells.

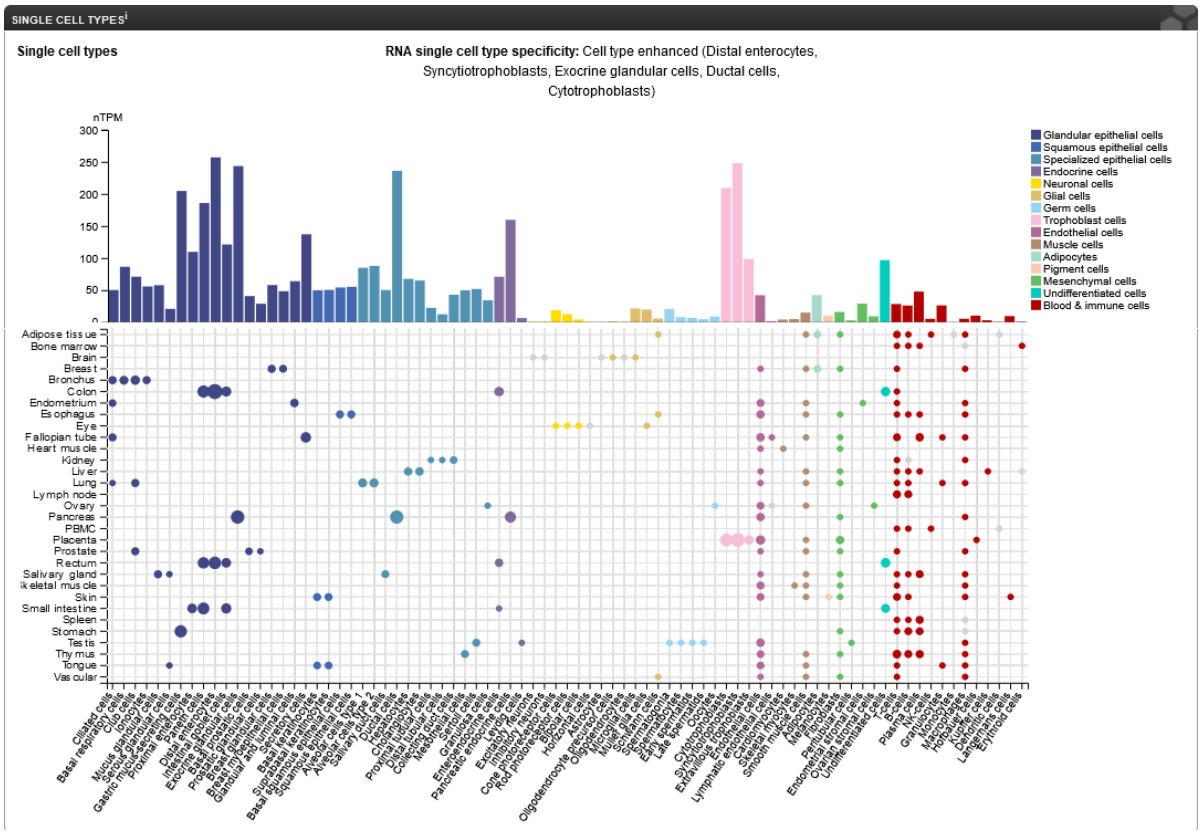


Figure 8. Single cell types expressing LSR in humans. This figure summarizes normalized single cell RNA (nTPM) from all single cell types. Colors represent into which cell type group the single cell is categorized into. Groups consist of cell types with common functional features. Bars represent the average of RNA expression (nTPM) of each cell type in all locations, while dots represent RNA expression (nTPM) of the cell type in specific tissues. Size of the bars and dots represent expression amount. (Human Protein Atlas, Uhlén et al. 2015; proteinatlas.org, 12.05.2025.)

To study the CDTb toxin-induced interactions between the LSR receptor and cellular proteins, the expression levels of LSR in different cell lines were analyzed with HPA to find cells with low or moderate endogenous LSR expression. Based on the analysis, the most suitable cell line for the actual experiments was HEK293T cells, as their endogenous LSR RNA expression was moderate and LSR protein expression was very low (Figure 9). Even though HeLa cells also have a low endogenous LSR expression, transfection of the desired constructs into HeLa cells prior to this thesis had been insufficient. Thus, HEK293T cells were considered the next best choice, and were thus utilized in the experiments of this thesis.

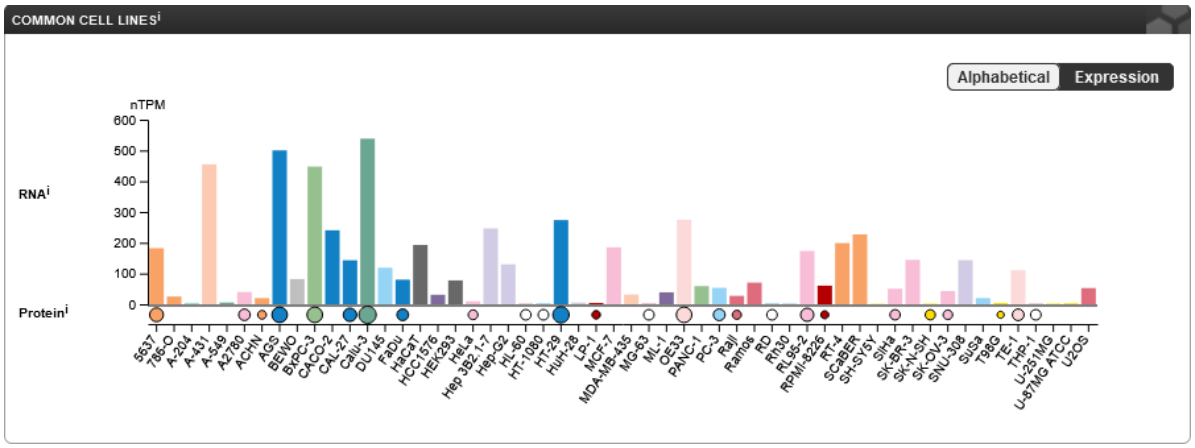


Figure 9. LSR expression in common cell lines. The figure relatively compares RNA level (nTPM), and protein level (nRPX) of each cell line. Levels are depicted as bars (RNA) and dots (protein), with size representing expression amount. Colors associated with cell lines represent the tissue types they are related to. RNA expression (normalized transcript per million, nTPM) is shown as the gene's individual nTPM in each listed cell line. Protein expression (Normalized Relative Protein Expression, nRPX) is based on Pan-Cancer Atlas project's MS proteomics dataset. A white dot means no detection of protein by MS, and the absence of a dot means that the MS data was not available. More information of RNA expression can be found at Cell line resource, and of protein expression at method. (Human Protein Atlas, Uhlén et al. 2015; proteinatlas.org, 12.05.2025.)

In conclusion, LSR is expressed in multiple tissues of the human body, at the protein level most prominently in the gastrointestinal tract, especially in the colon (Fig. 7A&B). A high expression of LSR at the protein level can be observed in the epithelial cells of the colon (Fig. 7C), while at the RNA level, distal enterocytes have the highest expression of LSR in the colon (Fig. 8). Endogenous LSR expression of cell lines was an essential factor for the experiments of this thesis; The moderate endogenous expression of LSR in HEK293T cells suited our experimental purposes (Fig. 9).

3.1.2 LSR has four predicted cytoplasmic motifs

To estimate possible interactions the LSR receptor may have with intracellular proteins after CDTb toxin docking, Scansite 4.0 (Obenauer et al., 2003) was utilized. These plausible protein-protein interactions can be studied as phosphorylation events (Nishi et al., 2011). Certain areas of proteins, e.g. domains or motifs, may indicate mediation of signaling via phosphorylation (Akiva et al., 2012). Thus, possible domains or motifs in LSR could potentially become phosphorylated after CDTb toxin docking, which could indicate their roles in transmitting essential downstream signals characteristic of *C. difficile* infection. To predict domains and motifs, the protein sequence of LSR attained from Uniprot (The UniProt Consortium, 2025) was entered into the "Motif Scan" program at Scansite. All mammalian motifs were searched at a high stringency and predicted domains were explored.

Nine predicted phosphorylation sites were found in the LSR protein (Figure 10). The best matching motifs corresponding to predicted sites were all listed and can be found in Appendix 6 with additional information. From the predicted sites with cytoplasm localized motifs, the leukocyte tyrosine kinase src homology 2 (ltk SH2) motif expressed the lowest percentile, meaning it scored in the best 0,019% of sites when compared to all motifs in the vertebrate subset of SWISS-PROT. The motif also has a moderately high surface accessibility (2,35) indicating a possible role in the downstream signaling of LSR.

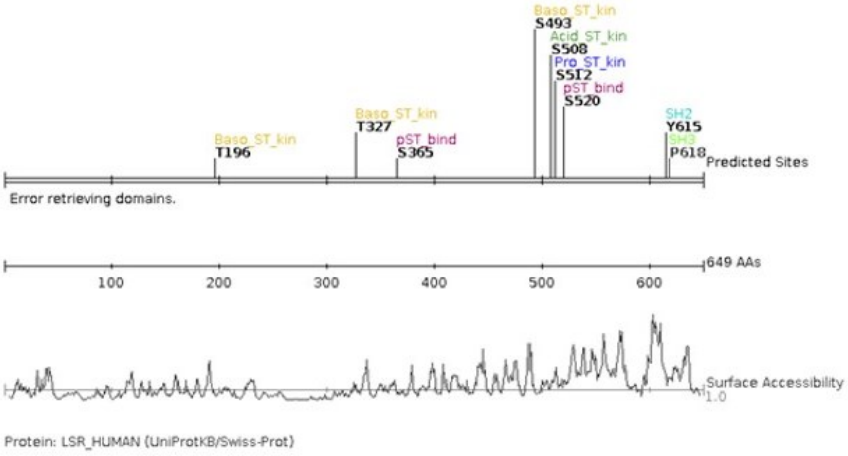


Figure 10. Graphical output of predicted protein phosphorylation sites in human LSR. On the upmost row, the LSR protein is represented as a line with rectangles constituting known domains. In this case, known domains could not be retrieved. Predicted phosphorylation sites and their corresponding motif family are labelled above the protein. On the second row, there is a ruler to follow the amino acid (AA) number progression along the protein for the bottom-most figure. On the last row, there is a plot depicting predicted surface accessibility at each residue. (Obenauer et al., 2003; http://scansite.mit.edu accessed 12.05.2025.)

Overall, the estimation of possible interactions between LSR and intracellular proteins as phosphorylation events utilizing Scansite 4.0, produced a list of plausible motifs, from which four were localized to the cytoplasm (Fig. 10 & Appendix 6). These four possible motifs in LSR could allude to their role in relaying downstream signals and indicate interactions between their canonical pathway target proteins after CDTb toxin docking to LSR.

3.2 Characterization of MAC-tagged protein expressing transfected HEK293T cells

HEK293T cells were transfected with ultraID constructs creating MAC-tagged proteins that enable proximity labeling-MS. LSR-MAC enables the study of interactions between LSR and intracellular proteins with CAAX-GFP-MAC functioning as a control in eventual proximity

labeling-MS. To optimize the protocol for proximity labeling-MS, transient transfection was done first to confirm MAC-tagged protein expression in cells. Additionally, optimal biotin concentration for ultraID (BirA*) function in proximity labeling was simultaneously considered. Stably transfected cell lines were then considered and new clones created to optimize sample preparation. Localization of MAC-tagged proteins to the plasma membrane of cells was additionally confirmed prior to proceeding to optimization of CDTb and LSR-MAC interaction.

3.2.1 Transient transfection

Transiently transfected and biotin stimulated HEK293T cells were checked for MAC-tagged protein expression and biotinylation of proteins. Western blot with anti-HA showed the antibody binding to proteins' (LSR-MAC 96,5kDa; CAAX-GFP-MAC 54,65kDa) HA tags, while cells without MAC-tagged proteins (pcDNAmock) showed no binding at these sites (Figure 11A). In cells grown in 50µM biotin media, autobiotinylation of MAC-tagged proteins was not unequivocal, while in the cells grown in 200µM biotin media, autobiotinylation was detected (Figure 11B). Additionally, biotinylated proteins could be observed in LSR– and CAAX cell samples, both in the 50µM– and 200µM biotin concentrations. Protein amounts between compared samples were reviewed with a loading control, anti-β-actin, as well as with Ponceau S staining of the membrane. Ponceau S was done only to observe and estimate protein loading between samples, thus further quantification or normalization of protein expression based on loading volumes was not done.

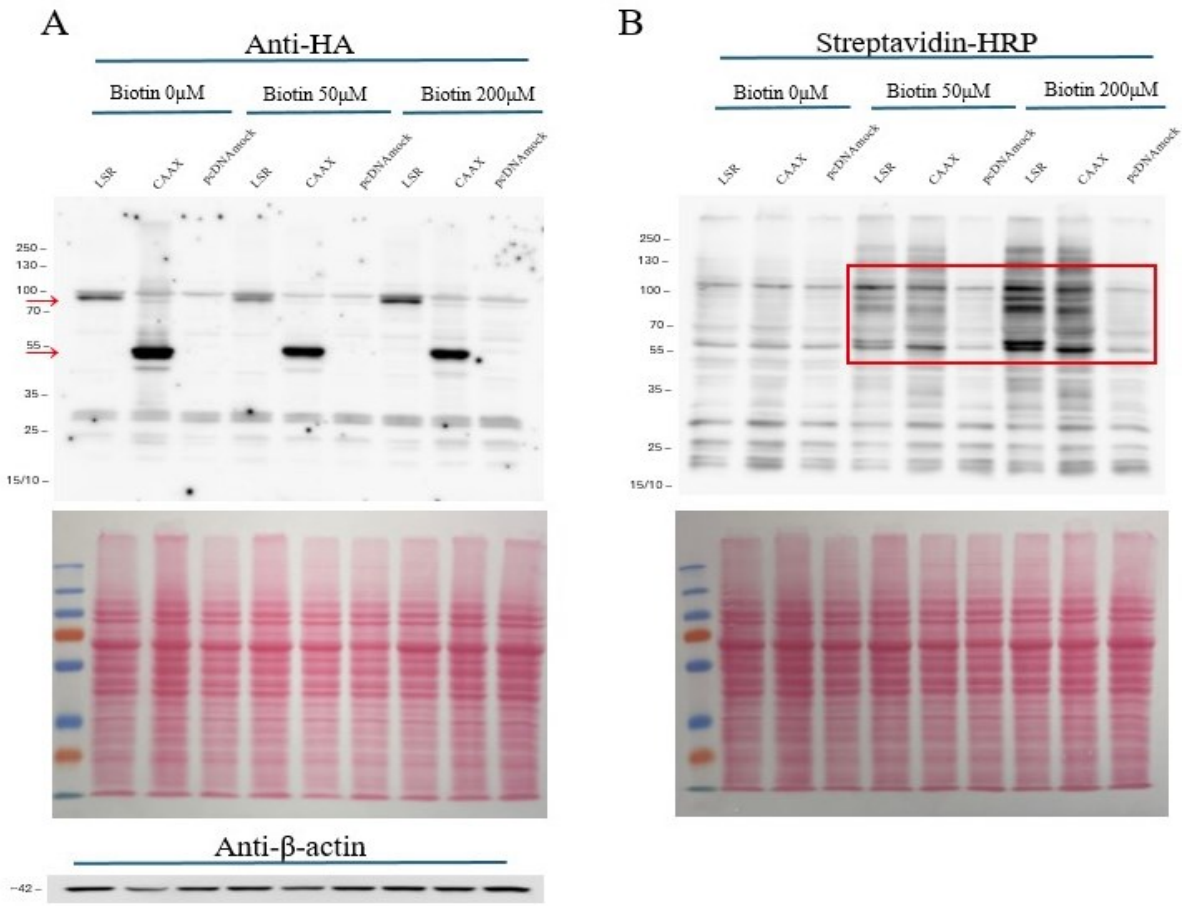


Figure 11. Confirming expression of MAC-tagged proteins and biotinylation of proteins. HEK293T cells were transiently transfected with plasmids containing constructs for LSR-MAC, CAAX-GFP-MAC, or pcDNAmock (contr.). 24 hours after transfection, media was changed into fresh DMEM with added supplements and biotin in the concentration of 0μM, 50μM, or 200μM. **(A)** Western blot with anti-HA (uppermost picture) was performed to check MAC-tagged protein expression. Expression of LSR-MAC (96,5kDa) and CAAX-GFP-MAC (54,65kDa) proteins in transfected cells was observed (red arrows). Ponceau S staining of the membrane (middle picture) was done to observe protein amounts between compared samples. Western blot with anti-β-actin as a loading control (lowermost picture) was additionally performed. **(B)** Western blot with streptavidin-HRP (upper picture) was done to confirm biotinylation of proteins due to MAC-tagged proteins' ultraID tag. Ponceau S staining of the membrane (lower picture) was done to observe protein amounts in compared samples. LSR and CAAX cells showed biotinylation of proteins (red) in cells grown in biotin 50μM– and 200μM media.

In conclusion, transfected HEK293T cells express their respective MAC-tagged proteins, LSR-MAC and CAAX-GFP-MAC (Fig. 11A). UltraID in the MAC-tagged proteins biotinylates proximal proteins in the presence of biotin, in both 50μM and 200μM concentrations (Fig. 11B).

3.2.2 Achieving stable expression of MAC-tagged proteins

Stably transfected pcDNAmock, CAAX, and LSR HEK293T cell lines previously created by Shoultz were considered to optimize eventual proximity labeling-MS protocol. Stable

expression of MAC-tagged proteins was checked with anti-HA western blot from three clones (K1-3) of each cell line (Figure 12A). Ponceau S staining of the membrane was performed only to observe and estimate protein loading between samples, thus further quantification or normalization of protein expression based on loading volumes was not done. All clones from the CAAX cell line stably expressed CAAX-GFP-MAC (54,65kDa). From the clones in LSR cell lines, only clone 2 (K2) stably expressed LSR-MAC (96,5kDa). Clones were at the following passage numbers: pcDNAmock K1 at passage 14, –K2 and –K3 at passages 18 and 8 respectively, CAAX K1 at passage 16, –K2 and –K3 at passages 10 and 21 respectively, and LSR K1 at passage 16, –K2 and –K3 at passages 5 and 10 respectively. One clone from all cell lines was selected to be used in further experiments: pcDNAmock K1 and CAAX K2 for cell fractionation, confocal microscopy and flow cytometry, and LSR K2 for cell fractionation and flow cytometry.

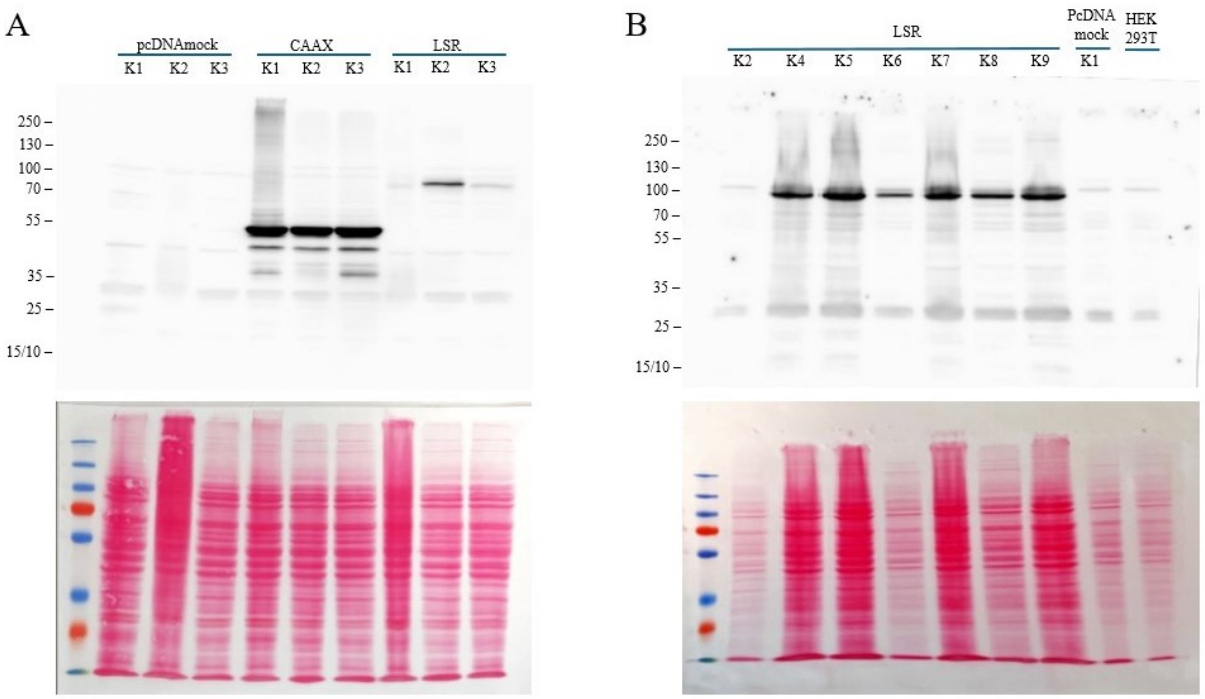


Figure 12. Confirming stable expression of MAC-tagged proteins. Western blot with anti-HA and membrane staining with Ponceau S were performed to analyze MAC-tagged protein expression in stably transfected cell lines. **(A)** Three clones (K1-K3) from stably transfected HEK293T pcDNAmock– (contr.), CAAX– and LSR cell lines were analyzed. All clones of CAAX cell lines stably expressed CAAX-GFP-MAC (54,65kDa). From the LSR cell lines, only clone two (K2) expressed LSR-MAC (96,5kDa) stably. **(B)** Western blot with anti-HA was done to check stable expression of LSR-MAC in newly created and selected clones (K4-K9), and in routinely passaged LSR K2, while pcDNAmock K1 and HEK293T cells functioned as controls. LSR K4-K9 cells expressed LSR-MAC (96,5kDa), while expression in LSR K2 cells was now absent.

As two clones of the previously created LSR cell lines by Shoultz, did not express LSR-MAC, the possibility of disappearing expression due to increased passaging was considered.

Additional LSR cell lines were created for further experiments when LSR K2 passage number increased. Selection of stably transfected cells was performed with hygromycin B in the concentration of 100µg/ml, and newly created cell lines were named LSR K4-K9. The expression of LSR-MAC in created cell lines was confirmed with western blot utilizing anti-HA (Figure 12B), and Ponceau S staining was performed to observe protein amounts between compared samples. Previously tested pcDNAmock K1 functioned as a control, while normal HEK293T cells were used as a negative control. LSR K2 cells were included to check expression of LSR-MAC after multiple passages of the cell line. At the time of the experiment, LSR K2 was at a passage number of 19, while LSR K4-K9 were at passage 5. All newly created and selected LSR clones K4-K9 expressed LSR-MAC (96,5kDa), while cells of the LSR K2 clone now showed no such expression.

In conclusion, utilizing cells stably expressing MAC-tagged proteins would be ideal for the proximity labeling-MS protocol. However, while the expression of CAAX-GFP-MAC seems to be stable in transfected HEK293T cells, the expression of LSR-MAC seems to decline and disappear with passaging (Fig. 12).

3.2.3 MAC-tagged proteins localize partly to the plasma membrane

Prior to further experiments to optimize methods for proximity labeling-MS, localization of MAC-tagged proteins in cells was elucidated. To determine LSR-MAC and CAAX-GFP-MAC subcellular localization, western blotting and confocal microscopy were used. Western blot analysis was done with LSR K2 cells at passage 12, and confocal microscopy analysis with LSR K4 cells at passage 7. The same precautions regarding passaging were not taken with pcDNAmock or CAAX cells, as there had not been any indication of decreased expression caused by passaging.

First, cellular components (cytosolic lysate (C), plasma membranous lysate (ME), mitochondrial lysate (M), nuclear lysate (N), and total lysate (T)) were separated from pcDNAmock K1, LSR K2 and CAAX K2 cells. Expression of MAC-tagged proteins and endogenous LSR were analyzed within C, ME, N, and T fractions using western blotting. From the LSR K2 cell line, mitochondrial lysate was also included in the analysis. Cell fractionation was confirmed by the ability to show the identity and purity of isolated protein fractions. The correct fractions were confirmed by applying antibodies targeting Na,K-ATPase (functioned as

a membrane loading control), lamin B1 (nuclear loading control), and MEK1/2 (a cytosolic loading control).

Endogenous LSR (~60kDa) was observed in the total lysate (T), nuclear (N), and the plasma membranous (ME) fractions of all cell lines, as well as in the mitochondrial (M) fraction of LSR cells (Figure 13A). The recombinant LSR-MAC protein (96,5kDa) could also be observed in LSR cells with anti-LSR staining. This result was further confirmed with anti-HA staining, where LSR-MAC could be observed in the mentioned fractions of LSR cells (Figure 13B). Anti-HA staining also showed recombinant CAAX-GFP-MAC protein (54,65kDa) in all fractions of CAAX cells. Successful fractionation was confirmed with a membrane– (anti-Na,K-ATPase), nuclear– (Lamin B1), and cytosolic (MEK1/2) loading control stainings. Isolation of membranes from cytosol was achieved in all cell lines (Figure 13C), but due to the absence of plasma membrane and mitochondrial loading controls, successful isolation of these membranous fractions could not be confirmed. As isolation of the nuclear fraction and the cytosolic fraction were equivocal (Figure 13D&E), results are not completely reliable. It's also important to note that protein amounts were not equalized between fractions and cell lines.

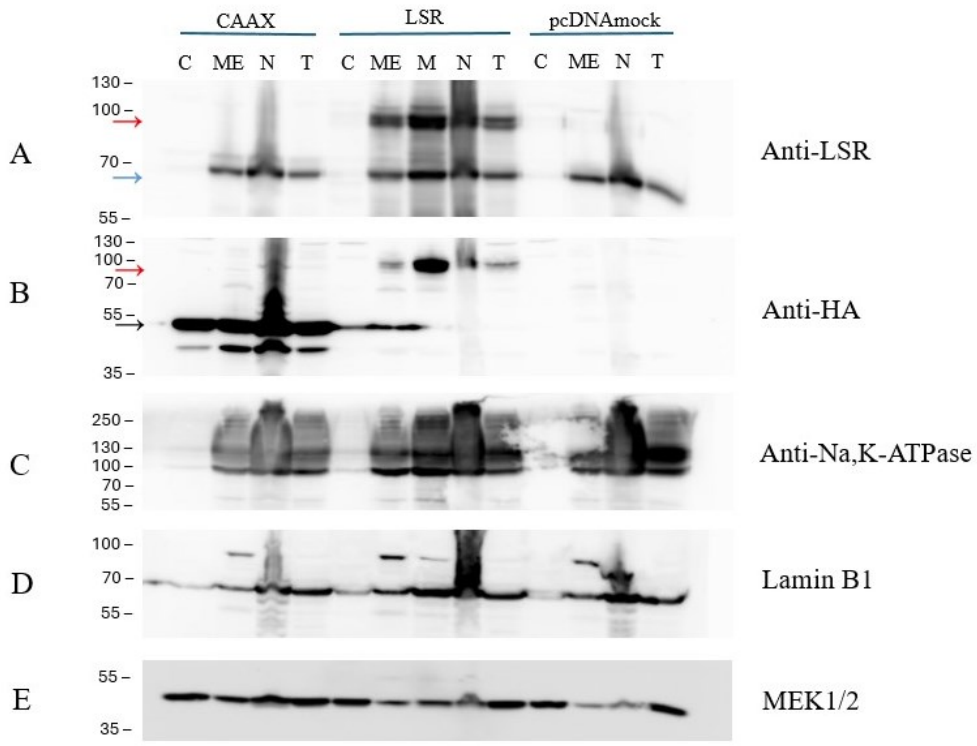


Figure 13. Discerning MAC-tagged protein localization in cells with fractionation. Stably transfected HEK293T CAAX K2–, LSR K2–, pcDNAmock K1 cells were fractionated and western blot was performed to discern LSR-MAC and CAAX-GFP-MAC localization. Anti-LSR staining was utilized to see endogenous LSR expression, while anti-HA staining was used to see MAC-tagged protein expression. Antibodies anti-Na,K-ATPase, Lamin B1, and MEK1/2 were used as loading controls to confirm successful fractionation. Protein amounts were not equalized between fractions and cell lines. **(A)** Endogenous LSR (~60kDa, blue arrow) and LSR-MAC (96,5kDa, red arrow) were observed in the plasma membranous and mitochondrial fractions (ME and M respectively). Results were attained with anti-LSR staining of the membrane. **(B)** CAAX-GFP-MAC (54,65kDa, black arrow) was observed in the cytosolic (C), plasma membranous (ME), and nuclear (N) fractions. LSR-MAC (96,5kDa, red arrow) was observed in the plasma membranous and mitochondrial fractions (ME and M respectively). Results were attained with anti-HA staining of the membrane. **(C)** Isolation of membranes from the cytosol was successful. Results were attained with anti-Na,K-ATPase staining. **(D)** Isolation of the nucleus was equivocal as there was leakage of saturation into other wells. Results were attained with Lamin B1 staining. **(E)** Isolation of the cytosol was successful with pcDNAmock– and LSR cells, while CAAX cells did not show clear results. Results were attained with MEK1/2 staining.

Confocal microscopy was used to further assess the subcellular localization of LSR-MAC and CAAX-GFP-MAC. For detecting MAC-tagged LSR protein, pcDNAmock K1 and LSR K4 cells were stained using a mouse anti-HA primary antibody and Anti-mouse-AlexaFluor647 (anti-mouse-AF-647) secondary antibody. DAPI served as a nuclear identifier, and Wheat Germ Agglutinin-Alexa488 (WGA) as plasma membrane labeling.

Nuclear staining with DAPI, membrane staining with WGA, and LSR-MAC staining with anti-HA and anti-mouse-AF647 were achieved in all samples (Figure 14A). LSR-MAC is observed localizing mostly to the intracellular region, but occasionally to the plasma membrane of HEK293T cells (Figure 14B). In the intracellular regions of cells, LSR-MAC could be observed colocalizing with membranous areas, especially in a region near the nucleus.

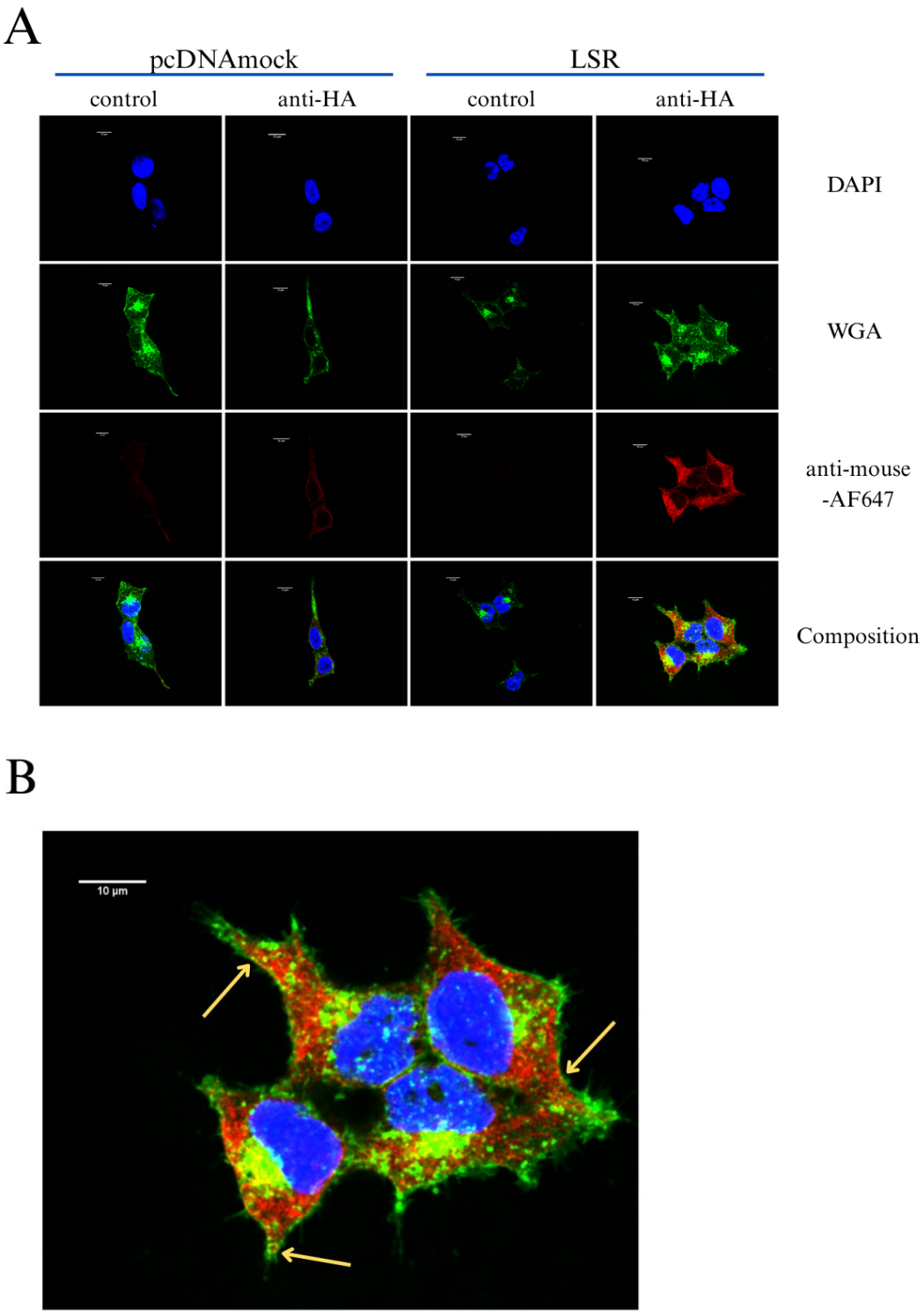


Figure 14. Elucidating LSR-MAC subcellular localization with confocal microscopy. LSR-MAC localized mainly to intracellular regions of HEK293T cells. PcDNAmock K1 and LSR K4 cells were immunofluorescence-stained and imaged with confocal microscope. **(A)** Images indicate location of nucleus (DAPI, blue), plasma membrane (WGA, green), and LSR-MAC (anti-HA + anti-mouse-AF647, red) in cells. There is a 10µm scale in each image. **(B)** The composition image of all stainings of LSR cells stained with anti-HA indicates LSR-MAC localization to the intracellular region (red and yellow), and occasionally to the plasma membrane (arrows) of cells.

For studying colocalization of CAAX-GFP-MAC and the plasma membrane, pcDNAmock K1 and CAAX K2 cells were stained using a rabbit anti-Na,K-ATPase primary antibody and Anti-rabbit-AlexaFluor647 (anti-rabbit-AF647) secondary antibody. The GFP tag in CAAX-GFP-MAC served as the protein identifier, and DAPI staining as a nuclear identifier.

Nuclear staining with DAPI, and membrane staining with anti-rabbit-AF647 was achieved in all samples (Figure 15A). Clear GFP signal could be observed in CAAX cells when compared to pcDNAmock cells. Additionally, the saturation difference of GFP signal between CAAX samples, as well as between individual cells in samples could be observed. CAAX-GFP-MAC is observed localizing mostly to the intracellular region and nucleus, but occasionally to the plasma membrane of cells (Figure 15B).

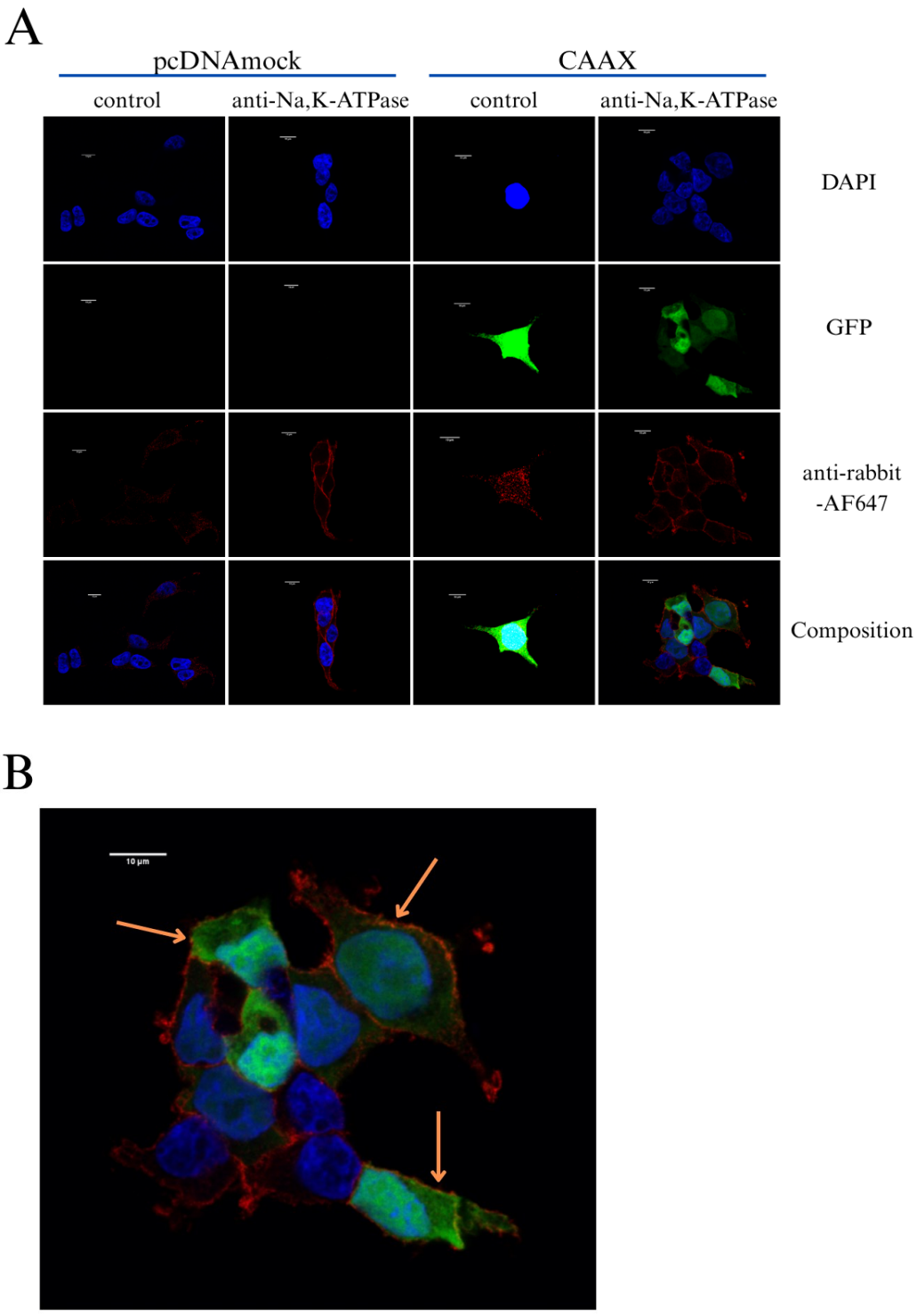


Figure 15. Elucidating CAAX-GFP-MAC subcellular localization with confocal microscopy. CAAX-GFP-MAC localizes mainly to intracellular regions and to the nucleus of cells. PcDNAmock K1 and CAAX K2 cells were immunofluorescence-stained and imaged with confocal microscope. **(A)** Images indicate location of nucleus (DAPI, blue), plasma membrane (anti-Na,K-ATPase + anti-rabbit-AF647, red), and CAAX-GFP-MAC (GFP, green) in cells. A 10µm scale is depicted in each image. **(B)** The composition image of all stainings in CAAX cells stained with Na,K-ATPase, indicates CAAX-GFP-MAC localization to the intracellular region and to the nucleus (green & teal), but occasionally to the plasma membrane (arrows) of cells.

In conclusion, based on cell fractionation and western blotting, LSR-MAC seems to localize to the presumed plasma membrane and mitochondria, and CAAX-GFP-MAC to the cytosol, plasma membrane and nucleus (Fig. 13A&B). However, the purity and isolation of all fractions could not be confirmed (Fig. 13C-E). Based on confocal microscopy, LSR-MAC and CAAX-GFP-MAC localized mainly to the intracellular regions, but occasionally also to the plasma membrane of HEK293T cells (Fig. 14B&15B). Inside the cells, LSR-MAC was mainly observed in a membranous area near the nucleus (Fig. 14B), while CAAX-GFP-MAC was observed all over the intracellular regions, including the nucleus (Fig. 15B).

3.3 Proteolytic cleavage and oligomerization of CDTb

During CDTb toxin docking to its cell surface receptor LSR, the protein is known to be proteolytically cleaved leading to its oligomerization into a heptameric prepore (Aktories 2022). To mimic this canonical cleavage and consequent oligomerization of CDTb during docking to LSR, experiments with trypsinization of CDTb were performed. Achieved proteolytic cleavage of CDTb was confirmed with western blot and SEC, and oligomerization with SEC-MALS.

3.3.1 CDTb is proteolytically cleaved by trypsin

Isolated and purified recombinant CDTb protein was incubated with trypsin in varying concentrations to optimize the ratio of trypsin and protein for proteolytic cleavage of CDTb, and results were analyzed with western blot and SEC. As some proteins are known to aggregate or degrade due to high temperatures during sample preparation (Burgess, 2009), both unheated and heated CDTb samples were analyzed.

Heated sample of trypsin and CDTb in the ratio of 1:5 resulted in the highest formation of cleaved protein (Figure 16A). Overall, the heated samples showed increased formation of cleaved protein corresponding with increasing concentration of trypsin. Unheated samples also showed a decrease of uncleaved CDTb corresponding with increasing trypsin concentration, but showed no clear increase in cleaved product of CDTb simultaneously. In all samples, the CDTb protein seemed to travel much slower in the gel than expected; Uncleaved recombinant CDTb should have showed as a band of 96,36kDa and cleaved as 73,6kDa, but corresponding sizes seemed to be at ~110kDa and ~85kDa.

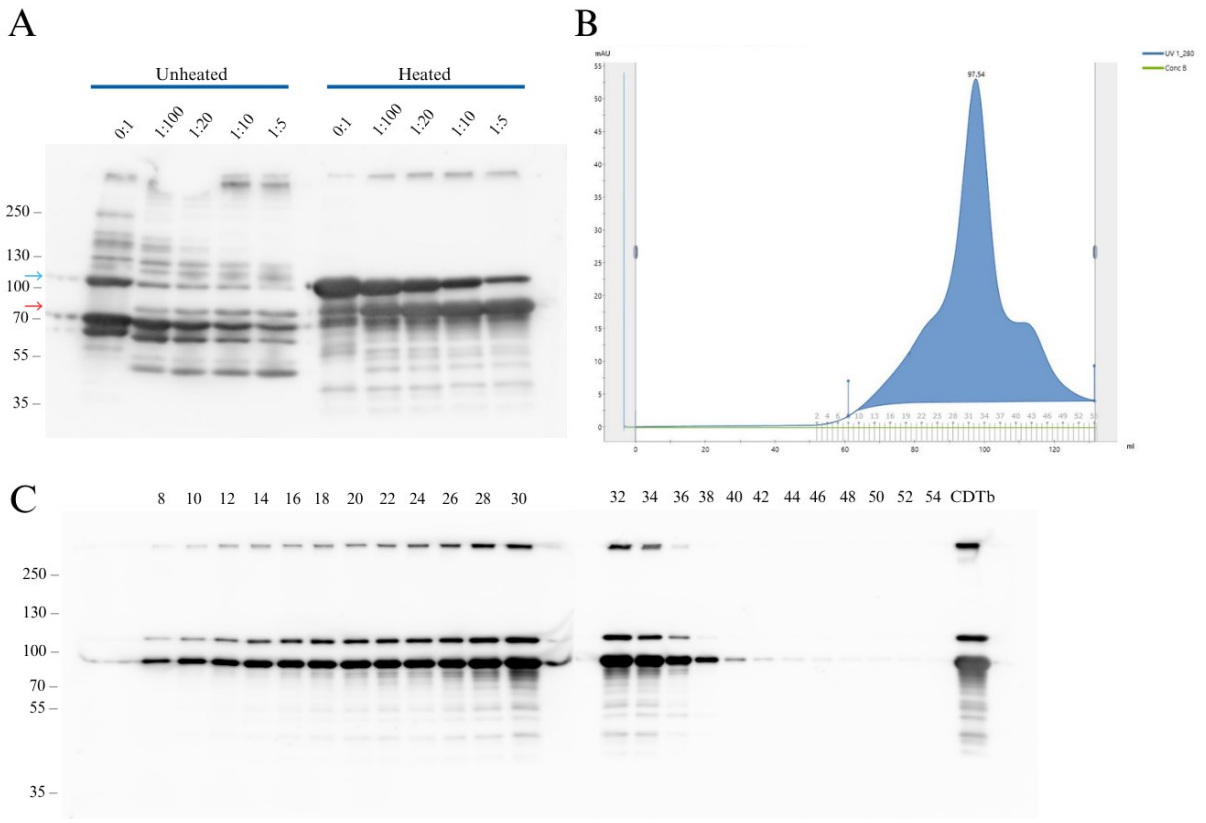


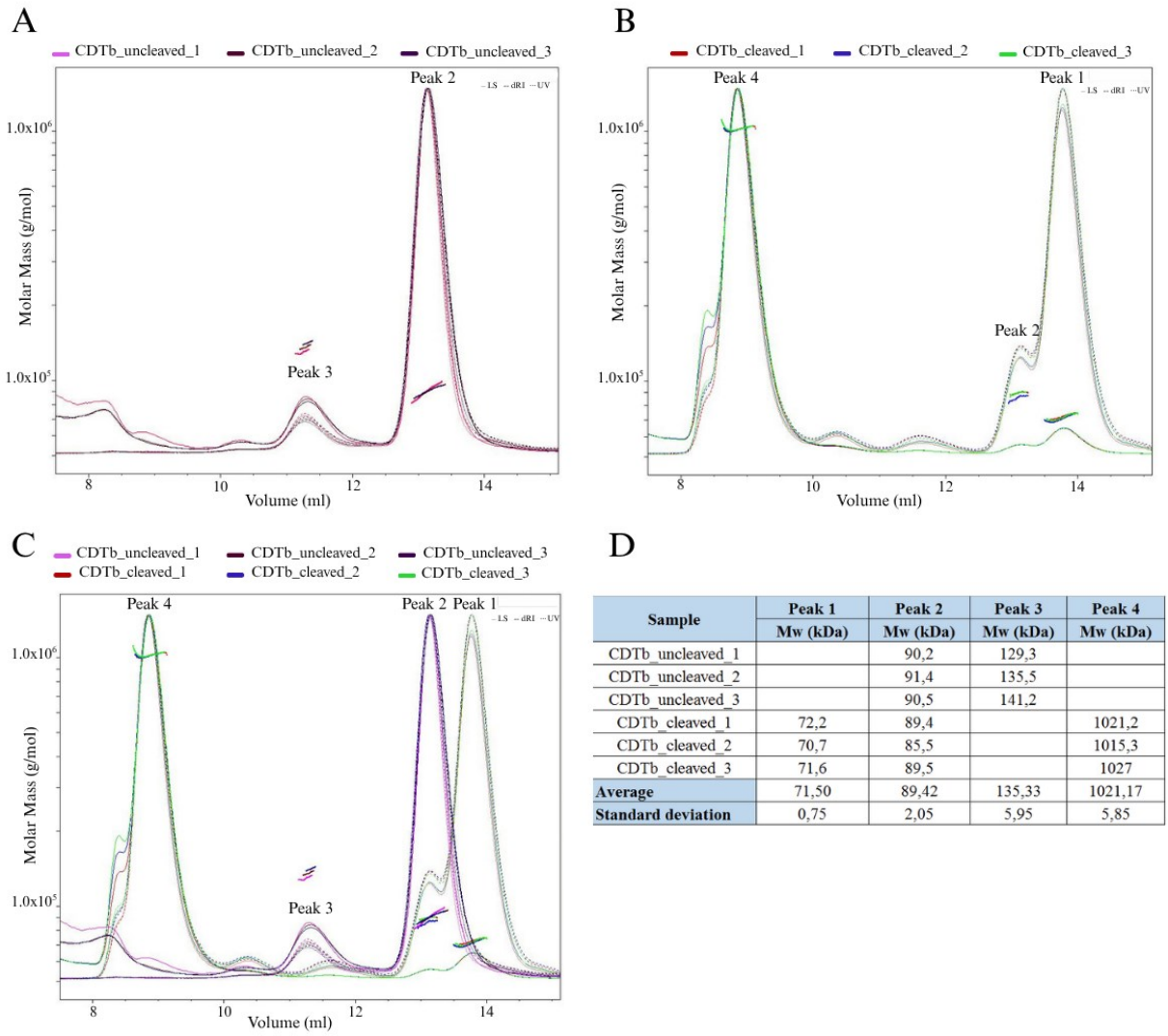
Figure 16. CDTb proteolytic cleavage via trypsin incubation. (A) Western blot of unheated– and heated samples (95°C) of CDTb with different ratios of trypsin (ratios marked as trypsin to CDTb). Isolated and purified CDTb protein (96,36kDa) was incubated with trypsin in varying concentrations for 45min in 37°C. Before gel run, each sample was divided in two, and one of them was heated in 95°C for 10minutes. Heated samples showed increased amounts of proteolytically cleaved CDTb (~85kDa, red arrow) corresponding to increasing trypsin concentrations. The CDTb protein (uncleaved 96,36kDa, proteolytically cleaved 73,6kDa) seemed to have travelled slower during gel run, as corresponding bands were observed during imaging at ~110kDa (blue arrow) and ~85kDa. **(B)** SEC profile (signal at 280nm) of a sample containing trypsin and CDTb in ratio of 1:5. Before SEC analysis, the sample was incubated for 45minutes at 37°C. Only one clear peak could be observed in the SEC profile corresponding to approximately the same sized proteins in the sample. Collected fractions (2-55) are shown under the peak. **(C)** Western blot of even numbered fractions (8-54) collected from SEC separation of trypsin and CDTb in ratio of 1:5 (B), as well as the same 1:5 sample, but without SEC separation. Samples were heated at 95°C for 10minutes prior to gel run based on results of (A). The imaged membrane showed both uncleaved CDTb (96,36kDa), as well as proteolytically cleaved CDTb (73,6kDa), as bands observed at ~110kDa and ~85kDa respectively.

To further analyze if proteolytic cleavage of CDTb was achieved with a trypsin incubation, SEC analysis of a sample containing the ratio of 1:5 of trypsin and CDTb was performed. SEC separates molecules based on size, allowing for the identification of cleavage products and their relative abundance. As results showed only one clear peak of approximately the same sized protein in the sample (Figure 16B), SEC fractions were then further analyzed with western blot to determine CDTb cleavage. Fractions 8–40 contained CDTb protein in uncleaved and cleaved forms (Figure 16C). Again, the CDTb protein seemed to have travelled slower during gel run than expected, but followed the same pattern as in the previous experiment (Figure 16A&C).

In conclusion, the optimal proteolytic cleavage of CDTb in these experiments was achieved with a 45-minute 37°C incubation in the ratio of 1:5 of trypsin and CDTb, followed by heating of the loading sample at 95°C for 10minutes (Fig.16A). While SEC analysis of the previously described sample containing 1:5 ratio of trypsin and protein, indicated only similar sized proteins in the sample (Fig. 16B), western blot analysis of SEC fractions revealed two different sized CDTb proteins, indicating proteolytic cleavage (Fig.16C).

3.3.2 CDTb oligomerizes after proteolytic cleavage

Next, the oligomerization of CDTb, which refers to the process where monomeric CDTb molecules assemble into larger complexes, was analyzed using Size-Exclusion Chromatography with Multi-Angle Light Scattering (SEC-MALS). Both uncleaved and trypsin (1:5) cleaved CDTb samples were used in the analysis. The measurements were repeated three times. In total, four different sized proteins were detected; Two from the uncleaved protein sample, and three in the cleaved protein sample, from which one was the same as in the uncleaved protein sample (Figure 17D). The uncleaved CDTb sample contained two peaks (peaks 2 and 3) corresponding to proteins of ~89kDa and ~135kDa, while the cleaved CDTb sample contained three peaks (peaks 1, 2, and 4) corresponding to proteins of ~71kDa, ~89kDa, and ~1021kDa (Figure 17).



Sample	Peak 1 Mw (kDa)	Peak 2 Mw (kDa)	Peak 3 Mw (kDa)	Peak 4 Mw (kDa)
CDTb_uncleaved_1		90,2	129,3	
CDTb_uncleaved_2		91,4	135,5	
CDTb_uncleaved_3		90,5	141,2	
CDTb_cleaved_1	72,2	89,4		1021,2
CDTb_cleaved_2	70,7	85,5		1015,3
CDTb_cleaved_3	71,6	89,5		1027
Average	71,50	89,42	135,33	1021,17
Standard deviation	0,75	2,05	5,95	5,85

Figure 17. SEC-MALS analysis of CDTb samples. SEC-MALS profiles (signal at 280nm) of **(A)** uncleaved CDTb samples, **(B)** cleaved CDTb samples utilizing trypsin to CDTb 1:5 ratio, and **(C)** an overlap image of previous two graphs. Overall, four peaks were detected in samples, corresponding to different sized proteins. In the uncleaved sample, two peaks could be detected (A), while the cleaved CDTb sample created three peaks (B), one of them overlapping with one peak from the uncleaved sample (A&C). **(D)** Table of SEC-MALS results showing protein sizes in each peak of (A, B, C). Peaks contained proteins with molecular weights of ca 71kDa, 89kDa, 135kDa, and 1021kDa. Standard deviation and averages of molecular weights within peaks were also listed.

In conclusion, SEC-MALS analysis of uncleaved and cleaved CDTb samples showed two different sized proteins in uncleaved CDTb sample, and three in cleaved CDTb sample indicating proteolytic cleavage and oligomerization of CDTb after trypsinization (Figure 17).

3.4 CDTb binds to MAC-tagged LSR receptor on the cell surface

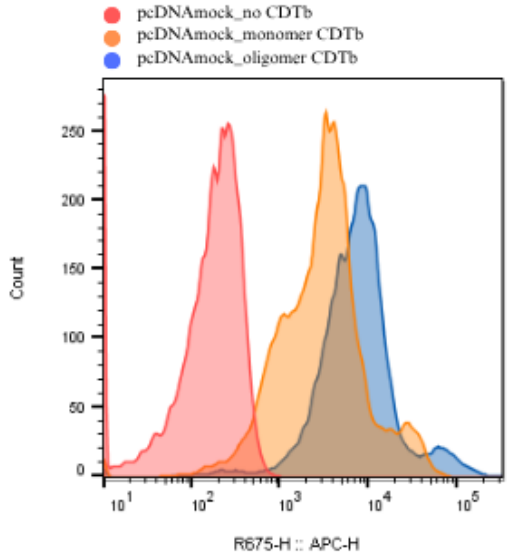
As MAC-tagged protein expression and biotinylation of proximal proteins was achieved, and the proteolytic cleavage and oligomerization of CDTb was confirmed, the interaction between

CDTb and LSR-MAC was studied next. Flow cytometry analysis of pcDNAmock K1 (contr.) and LSR K2 cells was first performed with LSR K2 cells at the point of passage 9 and repeated in triplicates at the point of passage 11.

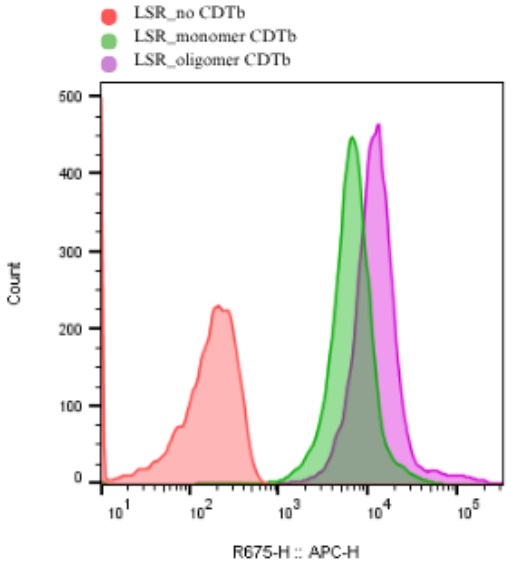
To assess the binding capability of CDTb to cell-surface receptors, pcDNAmock K1 and LSR K2 cells were incubated with Alexa Fluor 647 conjugated monomeric or oligomeric CDTb and analyzed by flow cytometry. Labeling of both monomeric CDTb and oligomeric CDTb with AlexaFluor647 was achieved with 3,25– and 3,52 degrees of labeling, respectively. Initially, the toxin concentration of 100µg/ml was used, but in the actual analyses, which were repeated 3 times, the toxin concentration was 10µg/ml.

The fluorescent signal of cells stimulated with Alexa Fluor 647 -tagged CDTb was observed to analyze binding affinity of the toxin to cells. As shown in Figure 18A&B, both in pcDNAmock K1 and LSR K2 cells, oligomeric CDTb had a higher binding affinity when compared to monomeric CDTb. With both monomeric or oligomeric CDTb stimulation of LSR cells, the average fluorescence intensity was increased compared to CDTb stimulated pcDNAmock cells (Figure 18C&D).

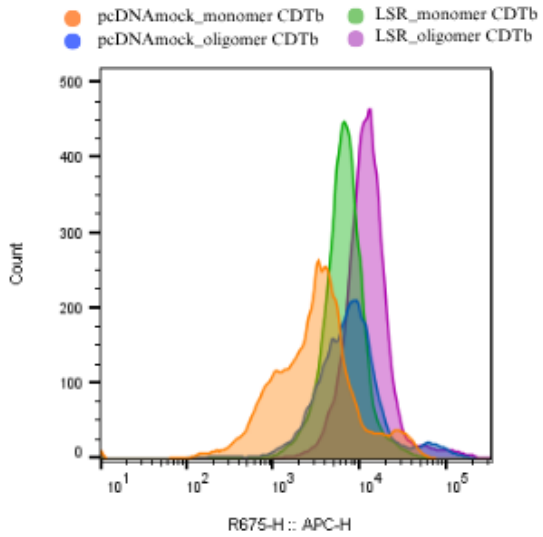
A



B



C



D

Cells	Average APC-H signal		
	No CDTb	Monomeric CDTb	Oligomeric CDTb
PcDNAmock K1	192	5448	12848
LSR K2	177	7632	15906

Figure 18. Flow Cytometry results of CDTb stimulation of transfected HEK293T cells. PcDNAmock K1 (contr.) and LSR K2 cells were stimulated with 10µg/ml Alexa Fluor 647 -tagged monomeric– or oligomeric CDTb. Observed APC-H signal, corresponding to fluorescently labelled CDTb binding to the cell surface of cells, was analyzed from all samples. Results were from a gated area created within the control, pcDNAmock_no CDTb, and copied to all samples. **(A)** Overlapping histograms of results obtained from CDTb stimulations of pcDNAmock K1 cells. Stimulation of pcDNAmock cells with CDTb monomer (orange) and –oligomer (blue) created a stronger APC-H signal when compared to stimulation without CDTb (red). **(B)** Overlapping histograms of results obtained from CDTb stimulations of LSR K2 cells. Stimulation of cells with monomeric– (green) and oligomeric (purple) CDTb created a stronger APC-H signal when compared to no CDTb (red) stimulation. **(C)** Overlapping histograms of stimulations performed with monomeric– and oligomeric CDTb to pcDNAmock K1– and LSR K2 cells. Stimulation of LSR cells created a stronger APC-H signal when compared to pcDNAmock cells in both monomeric (green vs orange) and oligomeric (purple vs blue) CDTb stimulations. The strongest signal could be observed with oligomeric CDTb stimulation of LSR cells. **(D)** Table of average APC-H signals detected in samples.

In conclusion, stimulation of LSR K2– and pcDNAmock K1 cells revealed higher binding affinity of oligomeric CDTb when compared to monomeric CDTb, indicated by observed APC-H signals. (Figure 18). Binding of both monomeric– and oligomeric CDTb was increased in LSR K2 cells when compared to pcDNAmock cells, indicated by observed APC-H signals.

4 DISCUSSION

As protein-protein interactions underlie cellular processes, understanding them can help decipher the complex pathways behind cellular function. Several methods, such as co-immunoprecipitation, crosslinking combined with mass spectrometry, and proximity labeling-MS, are currently used to study and elucidate the interactions of proteins. Proximity labeling enables the tagging of proteins transiently interacting with, or in the proximity of a protein of interest that has been fused with a biotin ligase. Thus, this method is especially useful in identifying weak or transient protein-protein interactions, which are often missed with canonical affinity purification MS approaches (Bosch et al. 2021). However, proximity labeling itself, as well as the following steps for the whole method of proximity labeling-MS, have limitations and require optimization. Previously developed biotin ligase enzymes such as TurboID and ultraID, address limitations in proximity labeling, and the created Multiple Approaches Combined –tag (MAC) by Kubitz et al. (2022) facilitates both proximity labeling and subsequent affinity purification of proteins. However, multiple factors in proximity labeling-MS still require consideration.

Clostridioides difficile infection (CDI) is currently one of the most significant hospital-acquired infections and is also classified as one of the top five urgent threats to human health by the CDC. One hallmark of hypervirulent strains of *C. difficile* is the production of *C. difficile* binary toxin (CDT) in addition to main virulence factors TcdA and TcdB (Martínez-Meléndez et al., 2022). To possibly prevent hypervirulent CDI, it is important to understand the cellular mechanisms behind it. While it has been shown that the lipolysis-stimulated lipoprotein receptor (LSR) functions as the host-receptor for CDT (Papatheodorou et al., 2024), the specific downstream signals of this interaction are yet unknown.

To study the possible protein interactions of LSR caused by toxin docking, LSR and CAAX-GFP (control) were fused with a MAC-tag and expressed in cells, after which cells were stimulated with CDTb and biotin. In cells, exogenously added biotin is converted by the ultraID ligase into a highly reactive compound, biotinyl-5'-AMP, which binds to the lysine residues of proteins in a nearby radius of 10nm. Subsequently, the biotinyl-5'-AMP –tagged proteins can be selectively affinity purified and then identified with mass spectrometry (Bosch et al. 2021). However, in order to achieve robust and reliable results, several parameters must be optimized. Characterization of LSR-MAC expressing cells and confirmation of CDTb and LSR-MAC interaction were thus essential goals of this thesis, and consequently, several key factors such

as selection of a suitable cell line, expression of plasma membrane localizing LSR-MAC in cells, the ideal exogenous biotin concentration for ultraID function, and simulation of canonical CDTb oligomerization were carefully considered.

4.1 Characterization of MAC-tagged protein expressing cells

One objective of this thesis was to transfect cells with constructs creating MAC-tagged proteins that would enable proximity labeling and affinity purification of proteins of interest. Moderate endogenous expression of LSR, high transfectability, and reliable maintainability of HEK293T cells (Lin et al., 2014; https://www.cytion.com/) led to the utilization of these cells in our experiments. Stable expression of MAC-tagged proteins and their localization to the plasma membrane of HEK293T cells was analyzed, and the optimal biotin concentration for ultraID function in proximity labeling was additionally considered.

4.1.1 Stable expression of MAC-tagged proteins

While expression of both LSR-MAC and CAAX-GFP-MAC (contr.) in transfected HEK293T cells was achieved, the stable expression of LSR-MAC seemed to decline and disappear with increased passaging of cells. At passage 5, LSR-MAC expression was observed, but at passage 10, the expression was observably decreased, and at passage 16 not observed at all.

Declining LSR-MAC expression could potentially be explained by receptor downregulation. While downregulation of receptors in cells is usually due to prolonged exposure to ligands, excess receptor expression alone might also negatively influence the function of cells (Su et al., 2024). In our experiments, the decreasing expression of LSR-MAC was observed prior to stimulation of cells with CDTb, which could indicate that the downregulation of LSR-MAC results from other negative impacts the protein has on the function of cells. As overexpression of LSR is linked to multiple cancer types (Papatheodorou et al., 2024; Kanda et al., 2023), it's possible that LSR-MAC expression is decreased and disappeared by cells to prevent dysregulation of signaling pathways leading to cancer. Additionally, the overexpression of LSR might metabolically burden cells (Oftadeh & Hatzimanikatis, 2024). Cells expressing low levels of LSR might exhibit increased overall health and cell proliferation when compared to cells with high LSR expression. This could lead to the selection of cells with a low expression of LSR in the population. However, similar problems with stable expression of transgenic canonical LSR isoform 1 were not reported by Reaves et al. (2014) with transfected Hs578t

claudin-low breast cell line. This indicates that the decreased expression of transgene LSR is cell line dependent or due to problems with the cloning vector used.

Since hygromycin B selection (100µg/ml) was maintained throughout cell culturing of all transfected cells, the hygromycin B gene was successfully integrated into genome of HEK293T cells. Cells that survived hygromycin B selection have thus likely integrated the LSR-MAC gene as well. Considering the previously discussed potential metabolic burden and linkage to cancer of LSR overexpressing cells, as well as successful hygromycin B gene integration, the survival of hygromycin B and LSR gene expressing cells during cell line selection is probable, with decreasing LSR expression explained by previously discussed selection of cells with low LSR expression during cell culturing. However, as LSR-MAC expression decreased over time, the possibility of partial integration cannot be ruled out. In partial integration, only some parts of the transfected DNA are stably incorporated into the HEK293T cell's genome, resulting in a mixed population of cells, some with the full LSR construct integrated, some with only parts of it, and some with no integration at all. Though Guo et al. (2021) have previously found that a hygromycin marker gene did not seem to significantly impede transgene expression of recombinant protein in HEK293 cells.

Based on our western blot results, the control protein, CAAX-GFP-MAC expression was not reduced with passaging similar to LSR-MAC. However, when analyzing protein localization from confocal images, a difference in GFP signal intensity between individual cells was observed. This heterogeneity in the amount of CAAX-GFP-MAC expression could indicate that the protein expression was not as stable as previously thought. At the time of the confocal imaging, CAAX K2 cells were at passage number 27, and during western blot to confirm stable expression they were at passage 21, which might suggest changes in expression between these passage numbers. However, as cells came from a mixed pool selection, it’s more likely that heterogeneity in CAAX-GFP-MAC expression is explained by differences in transgene expression levels between polyclonal cells.

Mori et al. (2020) have shown that adding a localization signal to GFP can alter its expression pattern in HEK293 cells, potentially by affecting the stability of the protein, translation, or interactions with cellular components. As CAAX is considered a localization motif, the combination of CAAX-GFP could lower the expression limit of our MAC-tagged protein. Thus, expression limits of CAAX-GFP-MAC might also differ between individual cells based on transgene expression. Overall, expression limit, differences in transfection efficiency between

cells, and previously discussed cell selection due to metabolic burden might all possibly explain the observed heterogeneity in CAAX-GFP-MAC expression.

4.1.2 Subcellular localization of MAC-tagged proteins

To enable the study of LSR interactome during toxin docking, it was essential to confirm MAC-tagged proteins' localization to the plasma membrane of HEK293T cells, where CDTb could bind to LSR-MAC. Furthermore, the localization of CAAX-GFP-MAC was analyzed to enable its utilization as a reliable control for LSR-MAC.

Analysis of LSR-MAC localization revealed expression mostly in intracellular regions of cells. Interestingly, LSR-MAC seemed to co-localize mainly with intracellular membranous structures. High saturation of LSR-MAC in the membranous structures of the cytosol could be explained with previously discussed receptor down-regulation. Internalization of excessive LSR-MAC to vesicles or even lysosomes might prevent dysregulation of cell function due to LSR overexpression. The highest saturation of LSR-MAC could be observed at a membranous structure near the nucleus, presumably the Golgi apparatus, which is responsible for processing proteins before they are transported to their final destinations in the cell. As the Golgi apparatus functions as a quality control organelle for proteins (Schwabl & Teis, 2022), LSR-MAC retention into this organelle could indicate problematic folding of the protein. Overall, protein retention and localization to the Golgi apparatus, vesicles, and lysosomes might indicate the degradation process of LSR-MAC.

Another explanation for the intracellular localization of LSR-MAC is the use of an LSR protein sequence, which is not canonically targeted to the plasma membrane. LSR has six isoforms with slightly different sequences (The UniProt Consortium, 2025), and it is possible that the subcellular localization might differ between them. Previously, Hemmasi et al. (2015) reported expression of LSR splice variants and the absence of the canonical LSR (isoform 1) in the colon carcinoma cell line HCT116. In the same study, plasma membrane targeting of LSR variants lacking a cysteine-rich region or a cytoplasmic tail were observed in HCT116 LSR KO cells. Additionally, Papatheodorou et al. (2011) reported CDT entry in HeLa cells via an LSR variant that has a shortened N-terminus (GenBank protein accession number: AAH04381.2) identical to the human LSR isoform 6 (Uniprot, Q86X29-6). Thus, a sequence for LSR with a clipped N-terminus might be essential for improving plasma membrane localization of the protein. Overall, since it was evident that LSR also localizes to the plasma membrane of HEK293T

cells, it was decided to proceed to the next stage, i.e., to the investigation of the interaction between CDTb and LSR-MAC using flow cytometry.

Localization of CAAX-GFP-MAC was also analyzed to enable its utilization as a control in proximity labeling-MS. Results indicated expression of CAAX-GFP-MAC mainly at intracellular regions but occasionally at the plasma membrane of cells. Previously, Choy et al. (1999) have postulated that a post-translational modification, i.e. prenylation, of CAAX leads to its localization to the endomembranous system, where further methylation occurs. In the same study, plasma membrane targeting is posited to happen only after methylation and be dependent on palmitoylation. The observed intracellular localization of CAAX-GFP-MAC in this master's thesis might thus indicate that the CAAX motif has been prenylated and has localized to the presumed ER and Golgi, yet methylation and subsequent plasma membrane localization have not, for some reason, occurred. However, as CAAX-GFP-MAC was occasionally observed on the plasma membrane of cells in our confocal imaging results, it was still plausible to consider it for use as a control for LSR-MAC in proximity labeling-MS.

4.1.3 Optimal biotin concentration for ultraID in proximity labeling

To enable proximity labeling, the function of ultraID to biotinylate proximal proteins of MAC-tagged LSR and CAAX-GFP was confirmed. The optimal biotin concentration for elucidating LSR interactome was also considered with 50µM and 200µM biotin concentrations based on previous research (Liu et al., 2020; Cho et al., 2020; Mair et al., 2019). Biotinylation of proteins was observed with both biotin concentrations, but clear autobiotinylation of the MAC-tagged proteins was only observed with the 200µM biotin treatment. However, it's important to note that high biotin concentration might also lead to labeling of non-target proteins (Cho et al., 2020). Additionally, higher than 50µM concentrations of biotin have been proposed to cause issues in affinity purification steps, possibly due to excess free biotin competing with biotinylated proteins for binding spots on streptavidin beads (Kim et al., 2023; Mair et al., 2019). Park & Kim (2025) also report that desalting to remove excess biotin is essential prior to affinity purification in cases of high biotin concentration during proximity labeling in plants. Thus, to optimize the proximity labeling-MS protocol for LSR interactome elucidation, biotin concentration of 50µM or lower could be used.

4.2 Optimizing CDTb and LSR-MAC interaction

An essential goal of this thesis was to optimize the interaction of the toxin of interest, CDTb, and the LSR-MAC protein expressed in transfected HEK293T cells. To achieve this, the quaternary structure of CDTb needed to be considered first. Canonically, CDTb oligomerizes during docking to LSR (Aktories, 2022), which might indicate the need to use oligomerized CDTb for successful interaction between the toxin and its receptor. It has also been hypothesized that receptor binding itself does not induce proteolytic cleavage and consequent oligomerization of CDTb (Anderson et al., 2020; Mullard et al., 2025). As recombinant CDTb was used in the practical experiments of this thesis, it was essential to review its quaternary structure and interaction with LSR to confirm its reliable use in proximity labeling-MS, in addition to reviewing the need for toxin oligomerization prior to receptor docking. Thus, simulation of CDTb cleavage and oligomerization was performed first, after which monomeric (untrypsinated) and oligomeric (trypsinated) CDTb were both used to stimulate LSR-MAC expressing cells to confirm the interaction and examine the necessity of oligomerization prior to LSR docking.

4.2.1 Simulation of canonical CDTb oligomerization

To confirm and optimize CDTb and LSR-MAC interaction, simulation of canonical cleavage and oligomerization of recombinant CDTb were performed first. Based on our results, optimal proteolytic cleavage of CDTb was achieved with trypsinization in the ratio of 1:5 of trypsin to toxin, and consequent oligomerization of CDTb was confirmed. Previously, Anderson et al. (2020) also accomplished proteolytic activation and oligomerization of recombinant CDTb (residues 43-876) by incubation with bovine trypsin with a 1:5 w/w ratio of trypsin to toxin.

When analyzing results of CDTb oligomerization, the canonical size of different CDTb quaternary structures needed consideration. The canonical monomeric CDTb subunit (~98kDa) (The Uniprot consortium, 2025) is activated via a ~20kDa proteolytic cleavage at the N-terminal, forming cleaved monomeric CDTb (~75kDa) (Xu et al., 2020). As cleaved monomeric CDTb is known to canonically oligomerize into a heptamere (Papatheoderou et al., 2024), the corresponding size of the heptameric CDTb protein could be calculated as ~525kDa. Additionally, a di-heptameric CDTb (~1000kDa) has been previously identified by Xu et al. (2020) in chymotrypsin activated CDTb samples (residues 212-876).

We observed two different sized proteins in untrypsinated CDTb samples: ~89kDa, 135kDa, and three in trypsinated CDTb samples: ~71kDa, ~89kDa, and ~1021kDa. Even though observed molecular weights in our results differ from previously mentioned CDTb protein sizes, the differences can be explained by the utilization of a recombinant CDTb protein that contains residues 43-876. Based on this, the observed ~89kDa protein would represent uncleaved monomeric CDTb, ~71kDa cleaved monomeric CDTb, and ~1021kDa di-heptameric CDTb. While the observed 135kDa protein matches the calculated size of a possible CDTb dimer, the protein was only observed in the untrypsinated CDTb sample, meaning it's unlikely to be an intermediate quaternary structure of CDTb. A possibility remains that this 135kDa protein is simply an impurity which isn't present in the trypsinated CDTb sample because it has been cleaved by trypsin.

Successful proteolytic cleavage and oligomerization of recombinant CDTb, as well as prior usage of recombinant CDTb containing similar residues (Anderson et al., 2020; Xu et al., 2020; Mullard et al., 2025) indicates reliable usage of the protein as a stimulator for LSR in future experiments studying CDTb and LSR interaction.

4.2.2 Confirmation of LSR-MAC and CDTb interaction

After confirming recombinant CDTb protein oligomerization via optimal proteolytic cleavage, both monomeric and oligomeric CDTb were considered for the interaction with LSR-MAC. If oligomerization of the toxin is not a requirement for receptor binding, the protocol for sample preparation for proximity labeling-MS could once again be optimized by removing excess steps and consequent error points involved in CDTb oligomerization.

Additionally, when considering previously discussed results of LSR-MAC localization, comparison of CDTb binding between LSR– and control cells was crucial. Consequently, our flow cytometry results indicated sufficient LSR-MAC expression on the plasma membrane of cells. When comparing the binding affinity of monomeric and oligomeric CDTb, our results indicated binding of the toxin in both of its quaternary structures. Even though oligomerized CDTb had a higher binding affinity when compared to monomeric CDTb, the binding affinity of monomeric CDTb seemed to be sufficient to indicate toxin docking even without prior proteolytic cleavage and oligomerization.

As the receptor binding domain (RBD) of CDTb alone has been shown to be sufficient for interaction with the Ig-like domain of LSR (Hemmasi et al., 2015) and even cause LSR

clustering into lipid rafts (Papatheodorou et al., 2013), monomeric CDTb seems to have potential as a stimulator of LSR. Additionally, multiple intermediate quaternary structures of trypsinated CDTb (residues 45-876) have been shown to have equal intoxication potency on Caco-2 cells when mixed with CDTa (residues 50-463) in a 1:1 ratio (Mullard et al., 2025) indicating that oligomerization of CDTb prior to receptor interaction could be redundant.

Now that characterization of MAC-tagged protein expressing cells and confirmation of CDTb and LSR-MAC interaction have been done, and included parameter have been optimized for proximity labeling-MS, following experiments on elucidating the interactome of LSR upon CDT docking can follow. This research into protein interactions downstream of LSR during CDT toxin docking, could then lead to a more comprehensive understanding of the intoxication pathway of CDT and subsequently give an outlook into the prevention of hypervirulent strains of *C. difficile*.

4.3 Challenges and limitations

Although the optimization of parameters for proximity labeling-MS was successful, there are still steps that can be improved.

Firstly, the choice of cell line could be considered more carefully, as the moderate endogenous LSR expression of HEK293 cells might still influence the analysis of CDTb and LSR-MAC interaction. To closely study canonical signaling after CDTb binding to LSR, epithelial cells of the colon might be the best choice. On the other hand, before starting the experiments, it should be ensured that LSR is not expressed endogenously in these cells or that its expression is minimal. There are a few studies showing that LSR can be removed from Caco-2 and HCT116 cell lines originating from the colon (Czulkies et al., 2017; Hemmasi et al., 2015). Thus, to avoid endogenous LSR, these LSR knockout (LSR KO) cell lines for example, could be transfected with MAC-tagged LSR and used in interaction assays.

Secondly, as LSR expression and localization can differ depending on splice variants, cellular context and experimental conditions (Reaves et al., 2014; Hemmasi et al., 2015; Papatheodorou et al., 2011), it must be ensured that the transfected LSR-MAC is localized to the cell membrane before performing proximity labeling-MS analysis.

Thirdly, to investigate the cellular responses mediated by CDTb binding to LSR, the quaternary structure and potency of the CDTb oligomer should be further tested. While Mullard et al.

(2025) recently demonstrated that intermediate quaternary structures of recombinant CDTb show equal potency in cell intoxication of Caco-2 cells, it is worth investigating whether these results apply without CDTa present.

Finally, the experiment on the interaction between CDTb and LSR-MAC could be improved by using a similarly fluorescently labelled control protein to estimate unspecific binding of our toxin. When considering the lack of this control in our experiment, the results might not relay the same information on monomeric CDTb binding. More evidence of monomeric CDTb binding to LSR should thus be done prior to proximity labeling-MS. Additionally, optimizing the CDTb concentration for stimulation of LSR-MAC expressing cells should be considered prior to proximity labeling-MS.

4.4 Conclusions and future perspectives

The optimization of parameters for proximity labeling-MS was essential for the elucidation of LSR interactome upon CDT docking. Optimization of cell line, MAC-tagged protein expression in cells, biotin concentration, CDTb oligomerization and toxin-receptor interaction were all integral steps in this process.

Based on our *in silico* review and the results of practical experiments, stably transfected HEK293T cells at the lowest possible passage number, passage 10 at most could be used in proximity labeling-MS. However, unstable expression, inconsistent results on localization of MAC-tagged proteins, as well as insufficient evidence of monomeric CDTb interaction with LSR-MAC give credence to the possibility that the sequences for MAC-tagged proteins need modifications. Additionally, cell line selection could be modified to LSR KO cells to optimize analysis of CDTb and LSR-MAC interaction.

While optimization of cell line selection and characterization of MAC-tagged protein expressing cells needs further consideration, optimization of biotin concentration for proximity labeling-MS was achieved. Combining our results of optimal biotin concentration for ultraID function and previous knowledge of affinity purification (Kim et al., 2023; Mair et al., 2019), biotin concentration of 50µM seems to be optimal for proximity labeling-MS. However, biotin incubation duration could still be considered and optimized for proximity labeling.

While CDTb oligomerization was achieved, it did not seem essential for receptor docking. However, further research needs to be done to confirm results of monomeric CDTb and LSR-

MAC interaction to truly optimize proximity labeling. Additionally, the concentration of CDTb to stimulate LSR-MAC expressing cells should still be optimized.

Overall, we achieved our objectives to elucidate expression, structure, and possible function of LSR and CDTb with *in silico* tools. We also successfully transfected HEK293T cells with a construct creating MAC-tagged LSR and simulated canonical cleavage and oligomerization of CDTb. However, we were not able to conclusively confirm the interaction of CDTb and LSR-MAC on the plasma membrane of transfected cells.

When taking into account the discussed results, the optimization of methods for proximity labeling-MS to elucidate protein-protein interaction caused by CDT docking to LSR will need to continue. The results of this thesis have given insight into the weak points of the optimization process, and thus the work continues. To address the most prevalent challenges concerning construct expression and toxin-receptor interaction, a new construct for LSR-MAC will be developed and transfected into LSR KO HCT116 cell. Functional validity of transfected LSR KO HCT116 cell lines, and the toxicity of different CDTb quaternary structures will be confirmed by combined CDTa and CDTb stimulation of cells. Optimal toxin and biotin stimulation concentrations and durations for proximity labeling-MS will additionally be considered, and any other future bridges will be crossed when they are reached.

ACKNOWLEDGEMENTS

I would like to thank Arto Pulliainen for giving me the opportunity to carry out my thesis in his research group, Turku Cellular Microbiology Laboratory (TCML), and for his guidance both in the writing process and in the laboratory throughout the whole thesis. I would also like to thank Tiina Henttinen for her help in the writing process. A special thank you goes to Madhukar Vedantham and Rajendra Bhatane for guiding my progress in the laboratory. Additionally, a thank you to all other TCML group participants, including Mika Savisalo and Ida Schoulz for their contributions in the materials utilized in this thesis work. Finally, a thank you to the Turku University Biology Department's professor Katja Anttila for reviewing this thesis.

REFERENCES

Abcam. (2025, May 27). *Anti-Sodium Potassium ATPase antibody [EP1845Y] – Plasma Membrane Loading Control (ab76020)*. https://www.abcam.com/en-us/products/primary-antibodies/sodium-potassium-atpase-antibody-ep1845y-plasma-membrane-loading-control-ab76020

Abt, M. C., McKenney, P. T., & Pamer, E. G. (2016). *Clostridium difficile* colitis: Pathogenesis and host defence. *Nature Reviews Microbiology*, *14*(10), 609–620. https://doi.org/10.1038/nrmicro.2016.108

Akiva, E., Friedlander, G., Itzhaki, Z., & Margalit, H. (2012). A dynamic view of domain-motif interactions. *PLoS Computational Biology*, *8*(1), e1002341. https://doi.org/10.1371/journal.pcbi.1002341

Aktories, K. (2023). From signal transduction to protein toxins—a narrative review about milestones on the research route of *C. difficile* toxins. *Naunyn–Schmiedeberg's Archives of Pharmacology*, *396*(2), 173–190. https://doi.org/10.1007/s00210-022-02300-9

Anderson, D. M., Sheedlo, M. J., Jensen, J. L., & Lacy, D. B. (2020). Structural insights into the transition of *Clostridioides difficile* binary toxin from prepore to pore. *Nature Microbiology*, *5*(1), 102–107. https://doi.org/10.1038/s41564-019-0601-8

Bosch, J. A., Chen, C. L., & Perrimon, N. (2021). Proximity-dependent labeling methods for proteomic profiling in living cells: An update. *Wiley Interdisciplinary Reviews: Developmental Biology*, *10*(1), e392. https://doi.org/10.1002/wdev.392

Burgess, R. (2009). Important but little known (or forgotten) artifacts in protein biochemistry. In R. R. Burgess & M. P. Deutscher (Eds.), *Methods in Enzymology* (Vol. 463, pp. 813–820). Academic Press. https://doi.org/10.1016/S0076-6879(09)63044-5

Carter, G. P., Chakravorty, A., Pham Nguyen, T. A., Mileto, S., Schreiber, F., Li, L., Howarth, P., Clare, S., Cunningham, B., Sambol, S. P., Cheknis, A., Figueroa, I., Johnson, S., Gerding, D., Rood, J. I., Dougan, G., Lawley, T. D., & Lyras, D. (2015). Defining the roles of TcdA and TcdB in localized gastrointestinal disease, systemic organ damage, and the host response during *Clostridium difficile* infections. *mBio*, *6*(3), e00551-15. https://doi.org/10.1128/mBio.00551-15

Chandrasekaran, R., & Lacy, D. B. (2017). The role of toxins in *Clostridium difficile* infection. *FEMS Microbiology Reviews*, *41*(6), 723–750. https://doi.org/10.1093/femsre/fux048

Cho, K. F., Branon, T. C., Udeshi, N. D., et al. (2020). Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nature Protocols*, *15*, 3971–3999. https://doi.org/10.1038/s41596-020-0399-0

Choy, E., Chiu, V. K., Silletti, J., Feoktistov, M., Morimoto, T., Michaelson, D., Ivanov, I. E., & Philips, M. R. (1999). Endomembrane trafficking of ras: The CAAX motif targets proteins to the ER and Golgi. *Cell*, *98*(1), 69–80. https://doi.org/10.1016/S0092-8674(00)80607-8

Cytion. (2025, July 21). *HEK293T cells*. https://www.cytion.com/HEK293T-Cells/300189

Czepiel, J., Dróżdż, M., Pituch, H., et al. (2019). *Clostridium difficile* infection: Review. *European Journal of Clinical Microbiology & Infectious Diseases*, *38*, 1211–1221. https://doi.org/10.1007/s10096-019-03539-6

Czulkies, B. A., Mastroianni, J., Lutz, L., Lang, S., Schwan, C., Schmidt, G., Lassmann, S., Zeiser, R., Aktories, K., & Papatheodorou, P. (2017). Loss of LSR affects epithelial barrier integrity and tumor xenograft growth of CaCo-2 cells. *Oncotarget*, *8*, 37009–37022. https://www.oncotarget.com/article/10425/text

Dieckmann, M., Dietrich, M. F., & Herz, J. (2010). Lipoprotein receptors—An evolutionarily ancient multifunctional receptor family. *Biological Chemistry*, *391*(11), 1341–1363. https://doi.org/10.1515/BC.2010.129

Ernst, K., Landenberger, M., Nieland, J., Nørgaard, K., Frick, M., Fois, G., Benz, R., & Barth, H. (2021). Characterization and pharmacological inhibition of the pore-forming *Clostridioides difficile* CDTb toxin. *Toxins*, *13*(6), 390. https://doi.org/10.3390/toxins13060390

Fabien Sohet, C., Lin, C., Munji, R. N., Lee, S. Y., Ruderisch, N., Soung, A., Arnold, T. D., Derugin, N., Vexler, Z. S., Yen, F. T., & Daneman, R. (2015). LSR/angulin-1 is a tricellular tight junction protein involved in blood–brain barrier formation. *Journal of Cell Biology*, *208*(6), 703–711. https://doi.org/10.1083/jcb.201410131

FlowJo™ Software (Version 10) [Computer software]. (2023). Becton, Dickinson and Company.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In J. M. Walker (Ed.), *The proteomics protocols handbook* (pp. 571–607). Humana Press. https://doi.org/10.1007/978-1-59745-228-1_36

Gerding, D. N., Johnson, S., Rupnik, M., & Aktories, K. (2014). *Clostridium difficile* binary toxin CDT: Mechanism, epidemiology, and potential clinical importance. *Gut Microbes, 5*(1), 15–27. https://doi.org/10.4161/gmic.26854

Guo, C., Fordjour, F. K., Tsai, S. J., Morrell, J. C., & Gould, S. J. (2021). Choice of selectable marker affects recombinant protein expression in cells and exosomes. *Journal of Biological Chemistry, 297*(1), 100838. https://doi.org/10.1016/j.jbc.2021.100838

Hemmasi, S., Czulkies, B. A., Schorch, B., Veit, A., Aktories, K., & Papatheodorou, P. (2015). Interaction of the *Clostridium difficile* binary toxin CDT and its host cell receptor, lipolysis-stimulated lipoprotein receptor (LSR). *Journal of Biological Chemistry, 290*(22), 14031–14044. https://doi.org/10.1074/jbc.M115.650523

Higashi, T., & Furuse, M. (2022). Tricellular tight junctions. In L. González-Mariscal (Ed.), *Tight junctions* (pp. 11–26). Springer. https://doi.org/10.1007/978-3-030-97204-2_2

Human Protein Atlas. (2025, February 26). RNA and protein expression summary (v24.0). https://www.proteinatlas.org/ENSG00000105699-LSR/tissue

Human Protein Atlas. (2025, April 16). Single cell types (v24.0). https://www.proteinatlas.org/ENSG00000105699-LSR/single+cell

Human Protein Atlas. (2025, April 16). Common cell lines (v24.0). https://www.proteinatlas.org/ENSG00000105699-LSR/cell+line

Human Protein Atlas. (2025, April 16). Colon antibody staining (v24.0). https://www.proteinatlas.org/ENSG00000105699-LSR/tissue/colon#imid_2077894

Kanda, M., Serada, S., Hiramatsu, K., Funauchi, M., Obata, K., Nakagawa, S., Ohkawara, T., Murata, O., Fujimoto, M., Chiwaki, F., Sasaki, H., Ueda, Y., Kimura, T., & Naka, T. (2023). Lipolysis-stimulated lipoprotein receptor-targeted antibody-drug conjugate demonstrates potent antitumor activity against epithelial ovarian cancer. *Neoplasia, 35,* 100853. https://doi.org/10.1016/j.neo.2022.100853

Kawamoto, A., Yamada, T., Yoshida, T., Sato, Y., Kato, K., Tsuge, H. (2022). Cryo-EM structures of the translocational binary toxin complex CDTa-bound CDTb-pore from *Clostridioides difficile. Nature Communications, 13*, 6119. https://doi.org/10.1038/s41467-022-33888-4

Kim, T. W., Park, C. H., Hsu, C. C., Kim, Y. W., Ko, Y. W., Zhang, Z., Zhu, J. Y., Hsiao, Y. C., Branon, T., Kaasik, K., Saldivar, E., Li, K., Pasha, A., Provart, N. J., Burlingame, A. L., Xu, S. L., Ting, A. Y., & Wang, Z. Y. (2023). Mapping the signaling network of BIN2 kinase using TurboID-mediated biotin labeling and phosphoproteomics. *The Plant Cell, 35*(3), 975–993. https://doi.org/10.1093/plcell/koad013

Kordus, S. L., Thomas, A. K., & Lacy, D. B. (2022). *Clostridioides difficile* toxins: Mechanisms of action and antitoxin therapeutics. *Nature Reviews Microbiology, 20*, 285–298. https://doi.org/10.1038/s41579-021-00660-2

Kowarz, E., Löscher, D., & Marschalek, R. (2015). Optimized Sleeping Beauty transposons rapidly generate stable transgenic cell lines. *Biotechnology Journal, 10*, 647–653. https://doi.org/10.1002/biot.201400821

Kubitz, L., Bitsch, S., Zhao, X., et al. (2022). Engineering of ultraID, a compact and hyperactive enzyme for proximity-dependent biotinylation in living cells. *Communications Biology, 5*, 657. https://doi.org/10.1038/s42003-022-03604-5

Kuehne, S. A., Collery, M. M., Kelly, M. L., Cartman, S. T., Cockayne, A., & Minton, N. P. (2014). Importance of toxin A, toxin B, and CDT in virulence of an epidemic *Clostridium difficile* strain. *Journal of Infectious Diseases, 209*(1), 83–86. https://doi.org/10.1093/infdis/jit426

Landenberger, M., Nieland, J., Roeder, M., Nørgaard, K., Papatheodorou, P., Ernst, K., & Barth, H. (2021). The cytotoxic effect of *Clostridioides difficile* pore-forming toxin CDTb. *Biochimica et Biophysica Acta (BBA) - Biomembranes, 1863*(6), 183603. https://doi.org/10.1016/j.bbamem.2021.183603

Lin, Y.-C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., Speleman, F., Marynen, P., Aerts, S., Deprest, J., Van Nieuwerburgh, F., Deforce, D., Mestdagh, P., Carmeliet, P., Vandesompele, J., ... Speleman, F. (2014). Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nature Communications, 5*, 4767. https://doi.org/10.1038/ncomms5767

Liu, X., Salokas, K., Tamene, F., Jiu, Y., Weldatsadik, R. G., Öhman, T., Varjosalo, M. (2018). An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. *Nature Communications, 9*, 1188. https://doi.org/10.1038/s41467-018-03523-2

Liu, X., Salokas, K., Weldatsadik, R. G., Öhman, T., & Varjosalo, M. (2020). Combined proximity labeling and affinity purification−mass spectrometry workflow for mapping and visualizing protein interaction networks. *Nature Protocols, 15*(10), 3182–3211. https://doi.org/10.1038/s41596-020-0365-x

Lourenço, L. C., Prosty, C., Huttner, A., Lee, T. C., & McDonald, E. G. (2025). Approach to *Clostridioides difficile* diarrheal infection. *CMI Communications, 2*(2), 105079. https://doi.org/10.1016/j.cmicom.2025.105079

Mair, A., Xu, S. L., Branon, T. C., Ting, A. Y., & Bergmann, D. C. (2019). Proximity labeling of protein complexes and cell-type-specific organellar proteomes in Arabidopsis enabled by TurboID. *eLife, 8*, e47864. https://doi.org/10.7554/eLife.47864

Martínez-Meléndez, A., Cruz-López, F., Morfin-Otero, R., Maldonado-Garza, H. J., & Garza-González, E. (2022). An update on *Clostridioides difficile* binary toxin. *Toxins, 14*(5), 305. https://doi.org/10.3390/toxins14050305

Mori, Y., Yoshida, Y., Satoh, A., et al. (2020). Development of an experimental method of systematically estimating protein expression limits in HEK293 cells. *Scientific Reports, 10*, 4798. https://doi.org/10.1038/s41598-020-61646-3

Morris, J., Knudsen, G., Verschueren, E., Johnson, J. R., Cimermancic, P., & Gross, M. (2014). Affinity purification–mass spectrometry and network analysis to understand protein-protein interactions. *Nature Protocols, 9*(12), 2539–2554. https://doi.org/10.1038/nprot.2014.164

Mostowy, S., & Cossart, P. (2012). Septins: The fourth component of the cytoskeleton. *Nature Reviews Molecular Cell Biology, 13*, 183–194. https://doi.org/10.1038/nrm3284

Mullard, R. M., & Sheedlo, M. J. (2025). Oligomerization of the *Clostridioides difficile* transferase B component proceeds through a stepwise mechanism. *PLoS Pathogens, 21*(7), e1013186. https://doi.org/10.1371/journal.ppat.1013186

Nishi, H., Hashimoto, K., & Panchenko, A. R. (2011). Phosphorylation in protein-protein binding: Effect on stability and function. *Structure, 19*(12), 1807–1815. https://doi.org/10.1016/j.str.2011.09.021

Obenauer, J. C., Cantley, L. C., & Yaffe, M. B. (2003). Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Research, 31*(13), 3635–3641. https://doi.org/10.1093/nar/gkg584

Oftadeh, O., & Hatzimanikatis, V. (2024). Genome-scale models of metabolism and expression predict the metabolic burden of recombinant protein expression. *Metabolic Engineering, 84*, 109–116. https://doi.org/10.1016/j.ymben.2024.06.005

Park, T.-K., & Kim, T.-W. (2025). A brief guide to analyzing TurboID-based proximity labeling-mass spectrometry in plants. *Molecules and Cells, 48*(8), 100236. https://doi.org/10.1016/j.mocell.2025.100236

Pfeiffer, C. T., Paulo, J. A., Gygi, S. P., & Rockman, H. A. (2022). Proximity labeling for investigating protein-protein interactions. *Methods in Cell Biology, 169*, 237–266. https://doi.org/10.1016/bs.mcb.2021.12.006

Papatheodorou, P., Carette, J. E., Bell, G. W., Schwan, C., Guttenberg, G., Brummelkamp, T. R., & Aktories, K. (2011). Lipolysis-stimulated lipoprotein receptor (LSR) is the host receptor for the binary toxin *Clostridium difficile* transferase (CDT). *Proceedings of the National Academy of Sciences of the United States of America, 108*(39), 16422–16427. https://doi.org/10.1073/pnas.1109772108

Papatheodorou, P., Hornuss, D., Nölke, T., Hemmasi, S., Castonguay, J., Picchianti, M., & Aktories, K. (2013). *Clostridium difficile* binary toxin CDT induces clustering of the lipolysis-stimulated lipoprotein receptor into lipid rafts. *mBio, 4*, e00244-13. https://doi.org/10.1128/mbio.00244-13

Papatheodorou, P., Minton, N. P., Aktories, K., & Barth, H. (2024). An updated view on the cellular uptake and mode-of-action of *Clostridioides difficile* toxins. In P. Mastrantonio & M. Rupnik (Eds.), *Updates on Clostridioides difficile in Europe* (Vol. 1435, pp. [pages]). Springer. https://doi.org/10.1007/978-3-031-42108-2_11

Pourliotopoulou, E., Karampatakis, T., & Kachrimanidou, M. (2024). Exploring the toxin-mediated mechanisms in *Clostridioides difficile* infection. *Microorganisms, 12*(5), 1004. https://doi.org/10.3390/microorganisms12051004

Reaves, D. K., Fagan-Solis, K. D., Dunphy, K., Oliver, S. D., Scott, D. W., & Fleming, J. M. (2014). The role of lipolysis stimulated lipoprotein receptor in breast cancer and directing breast cancer cell behavior. *PLoS One, 9*(3), e91747. https://doi.org/10.1371/journal.pone.0091747

Samavarchi-Tehrani, P., Samson, R., & Gingras, A. C. (2020). Proximity dependent biotinylation: Key enzymes and adaptation to proteomics approaches. *Molecular & Cellular Proteomics, 19*(5), 757–773. https://doi.org/10.1074/mcp.R120.001941

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods, 9*(7), 671–675. https://doi.org/10.1038/nmeth.2089

Schwabl, S., & Teis, D. (2022). Protein quality control at the Golgi. *Current Opinion in Cell Biology, 75*, 102074. https://doi.org/10.1016/j.ceb.2022.02.008

Sicari, B., Londono, R., Dziki, J., & Badylak, S. (2023). Extracellular matrix as a bioscaffold for tissue engineering. In J. De Boer, C. A. Van Blitterswijk, J. A. Uquillas, & N.

Malik (Eds.), *Tissue engineering* (3rd ed., pp. 137–172). Academic Press. https://doi.org/10.1016/B978-0-12-824459-3.00005-6

Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H., & Kuijper, E. J. (2016). *Clostridium difficile* infection. *Nature Reviews Disease Primers, 2*, 16020. https://doi.org/10.1038/nrdp.2016.20

Spigaglia, P., Mastrantonio, P., & Barbanti, F. (2018). Antibiotic resistances of *Clostridium difficile*. In P. Mastrantonio & M. Rupnik (Eds.), *Updates on Clostridium difficile in Europe* (Vol. 1050, pp. [pages]). Springer. https://doi.org/10.1007/978-3-319-72799-8_9

Su, J., Song, Y., Zhu, Z., et al. (2024). Cell–cell communication: New insights and clinical implications. *Signal Transduction and Targeted Therapy, 9*, 196. https://doi.org/10.1038/s41392-024-01888-z

Tan, E., Chin, C. S. H., Lim, Z. F. S., & Ng, S. K. (2021). HEK293 cell line as a platform to produce recombinant proteins and viral vectors. *Frontiers in Bioengineering and Biotechnology, 9*, 796991. https://doi.org/10.3389/fbioe.2021.796991

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., ... Pontén, F. (2015). Tissue-based map of the human proteome. *Science, 347*(6220), 1260419. https://doi.org/10.1126/science.1260419

The UniProt Consortium. (2025). UniProt: The universal protein knowledgebase in 2025. *Nucleic Acids Research, 53*(D1), D609–D617. https://doi.org/10.1093/nar/gkae1010

Yen, F. T., Masson, M., Clossais-Besnard, N., André, P., Grosset, J. M., Bougueleret, L., Dumas, J. B., Guerassimenko, O., & Bihain, B. E. (1999). Molecular cloning of a lipolysis-stimulated remnant receptor expressed in the liver. *Journal of Biological Chemistry, 274*(19), 13390–13398. https://doi.org/10.1074/jbc.274.19.13390

Wahlström, A., Brumbaugh, A., Sjöland, W., et al. (2024). Production of deoxycholic acid by low-abundant microbial species is associated with impaired glucose metabolism. *Nature Communications, 15*, 4276. https://doi.org/10.1038/s41467-024-48543-3

Wigelsworth, D. J., Ruthel, G., Schnell, L., Herrlich, P., Blonder, J., et al. (2012). CD44 promotes intoxication by the *Clostridial* Iota-family toxins. *PLoS ONE, 7*(12), e51356. https://doi.org/10.1371/journal.pone.0051356

Xu, X., Godoy-Ruiz, R., Adipietro, K. A., Peralta, C., Ben-Hail, D., Varney, K. M., Cook, M. E., Roth, B. M., Wilder, P. T., Cleveland, T., Grishaev, A., Neu, H. M., Michel, S. L.

J., Yu, W., Beckett, D., Rustandi, R. R., Lancaster, C., Loughney, J. W., Kristopeit, A., Christanti, S., ... Weber, D. J. (2020). Structure of the cell-binding component of the Clostridium difficile binary toxin reveals a di-heptamer macromolecular assembly. *Proceedings of the National Academy of Sciences of the United States of America, 117*(2), 1049–1058. https://doi.org/10.1073/pnas.1919490117

APPENDICES

Appendix 1

Macherey-Nagel, Endotoxin-free plasmid DNA purification.

NucleoBond® Xtra Midi EF
NucleoBond® Xtra Maxi EF
NucleoBond® Xtra Midi Plus EF
NucleoBond® Xtra Maxi Plus EF

January 2022/Rev. 13

https://seafile.utu.fi/f/2a84cdf9c8ee4adb9c49/?dl=1

Appendix 2

PAGE for proteins

Resolving (lower) gel

Reagent	5%	6%	7.5%	10%	12%	15%
4x lower gel buffer	1250	1250	1250	1250	1250	1250
AA-MBAA (30 : 0.8)	835	1000	1250	1665	2000	2500
H_2O	2915	2750	2500	2085	1750	1250
APS (10%)	25	25	25	25	25	25
TEMED	5	5	5	5	5	5

All volumes in µl for 1 gel with a 1mm spacer (typical spacers 0.75, 1 and 1.5mm).

For example, to get 4 gels multiply volumes by 5 (1mm spacer) and by 6 (1.5mm spacer).

Use n-Butanol (saturated 1:1 with water) to cover the gel during polymerization.

Polymerize approx. 2h !!! Preferably overnight.

Stacking (upper) gel)

Reagent	Volume (µl)
4x upper gel buffer	750
AA-MBAA (30 : 0.8)	300
Sucrose (30%)	1950
APS (10%)	15
TEMED	3

All volumes in µl for 1 gel with a 1mm spacer.

Polymerize approx. 1h

4x upper gel buffer: 0.5M Tris / 0.4% SDS / adjust the pH to 6.8 with HCl in ddH_2O

4x lower buffer gel: 1.5M Tris /0.4% SDS / adjust the pH to 8.8 with HCl in ddH_2O

AA-MBAA (30 : 0.8): 300g acrylamide / 8g N,N'-methylenebisacrylamide in 1l of ddH_2O

APS: 100mg of APS in 1ml of ddH_2O (prepare new one weekly, or use frozen stocks)

MW (g/mol)

Tris: 121.14g/mol

SDS: 288,38g/mol

Glycin: 75.07g/mol

Separation (acrylamide concentration/kDa)

15% / 10-43kDa, 12% / 12-60kDa, 10% / 20-80kDa, 7.5% / 36-94kDa, 5% / 57-212kDa

Appendix 3

10x Running/Transfer buffer

Tris-base: 30g

Glycine: 144g

ddH2O: 1 L

- Prepare 800 mL of distilled water in a suitable container.
- Add 30.3 g of Tris base to the solution.
- Add 144.4 g of Glycine to the solution.
- Add distilled water until the volume is 1 L.

1x Running buffer

10x Running buffer: 100ml

10% SDS: 10ml

ddH2O: 1 L

- Prepare 800 mL of distilled water in a suitable container.
- Add 100ml of 10x Running buffer to the solution.
- Add 10ml of 10% SDS to the solution.
- Add distilled water until the volume is 1 L.

1x Transfer buffer

10x Running buffer: 100ml

Methanol: 200ml

ddH2O: 1 L

- Prepare 800 mL of distilled water in a suitable container.
- Add 100ml of 10x Running buffer to the solution.
- Add 200ml of Methanol to the solution.
- Add distilled water until the volume is 1 L.

Appendix 4

Stripping for reprobing (Abcam®)

Stripping is the term used to describe the removal of primary and secondary antibodies from a Western blot membrane. Stripping is useful when one wants to investigate more than one protein on the same blot, for instance a protein of interest and a loading control. When probing for multiple targets, stripping and re-probing a single membrane instead of running and blotting multiple gels have the advantage of saving samples, materials, and time.

It is not advisable to make quantitative comparisons of targets probed before and after stripping since the procedure removes some sample protein from the membrane. For the same reason, a stripped membrane should not be probed to demonstrate the absence of a protein.

A PVDF membrane is highly recommended to minimize loss of sample protein. Note also that colorimetric/chromogenic detection reagents will leave a permanent visible stain on the membrane that can interfere with subsequent detection of targets of similar molecular weights. Chemiluminescent reagents such as ECL are recommended as they will not leave a stain and are more sensitive than colorimetric reagents.

The following two protocols differ in harshness of treatment. As a rule of thumb, try the gentler one first and then proceed to the harsher one if there is still a signal from the antibody that one is trying to strip. These steps can be repeated for probing with several antibodies, though the potential signal may be weaker and the background higher after each round of stripping. Some researchers report successfully staining a membrane after stripping ten or more times.

Efficiency of stripping can be checked by incubating the membrane with chemiluminescent detection reagent. If stripping is judged to be satisfactory, rinse the membrane several times with buffer, then block before proceeding to the antibody incubation.

Mild stripping

Buffer, 1 liter

15 g glycine

1 g SDS

10 ml Tween20

Adjust pH to 2.2

Bring volume up to 1 L with ultrapure water.

Membrane incubation

Use a volume that will cover the membrane. Incubate at room temperature for 5-10 minutes.
Discard buffer.
5-10 minutes fresh stripping buffer.
Discard buffer.
10 minutes PBS
10 minutes PBS
5 minutes TBST
5 minutes TBST
Ready for blocking stage.

Harsh stripping
Prepare buffer and strip membranes under a fumehood.
Buffer, 0.1 liters
20 ml SDS 10%
12.5 ml Tris HCl pH 6.8 0.5M
67.5 ml ultra pure water
Add 0.8 ml ß-mercaptoethanol under the fumehood.
Discover more at abcam.com/technical

Procedure
1. Warm the buffer to 50°C.
2. Add the buffer to a small plastic box which has a tight lid. Use a volume that will cover the membrane.
3. Add the membrane. Incubate at 50°C for up to 45 minutes with some agitation.
4. Dispose of the solution as required for ß-mercaptoethanol based buffers.
5. Rinse the membrane under running water tap for 1-2 hours.
6. Traces of ß-mercaptoethanol will damage the antibodies. Wash extensively for 5 minutes in TBST.
Ready for blocking stage

Appendix 5

Nanotemper Protein Labeling Kit RED-NHS 2nd Generation For Dianthus, Monolith and NT. Automated Instruments with a RED Detector (Cat# MO-L011)

https://seafile.utu.fi/f/c0b641dad7ff48329341/?dl=1

Appendix 6

Table of results showing additional information on high-scoring motif sites found in human LSR. Each motif family with a predicted site is listed in this table under "motif group" and is accompanied by information of the best matching domain motif under the title "motif". Under the title "site" and "sequence", the position of the motif in the protein is indicated. The score and percentile indicated with each motif are ranks achieved with comparing the motif with every possible motif site included in the vertebrate subset of SWISS-PROT. Calculated surface accessibility and localization of each motif site is additionally listed. (Obenauer et al., 2003 ; http://scansite.mit.edu accessed 12.05.2025.)

Motif Group	Motif	Site	Sequence	Score	Percentile	Surface Accessibility	Localization
Phosphoserine/ threonine binding group (pST_bind)	14-3-3 Mode 1 (1433_m 1)	S365	GDVDRSS sAGGQGS Y	0.294	0.107%	0.5457	cytoplasm
Src homology 2 group (SH2)	Itk SH2 (Itk_SH2)	Y615	KEEEEEA yYPPAPPP	0.319	0.019%	2.3536	cytoplasm
Proline-dependent serine/threonin e kinase group (Pro_ST_kin)	CDK1 motif 2 - [ST]Pxx K (CDK1_ 2)	S512	TAESGSRs PTSNGGR	0.332	0.191%	2.4107	cytoplasm
Acidophilic serine/threonin e kinase group (Acid_ST_kin)	GSK3-improved (GSK3b)	S508	TPPSTAEs GSRSPTS	0.332	0.063%	1.3957	nucleus
Basophilic serine/threonin e kinase group (Baso_ST_kin)	Aurora B (AuroB)	T196	YYQGRRIt ITGNADL	0.348	0.040%	0.4546	nucleus
Basophilic serine/threonin e kinase group (Baso_ST_kin)	Akt Kinase (Akt_Kin)	S493	ARRPRAR sVDALDD L	0.402	0.063%	0.7609	cytoplasm
Acidophilic serine/threonin e kinase group (Acid_ST_kin)	GSK3 Kinase (GSK3_ Kin)	S508	TPPSTAEs GSRSPTS	0.435	0.183%	1.3957	nucleus
Src homology 3 group (SH3)	Nck 2nd SH3 (Nck_2n d_SH3)	P618	EEEAYYP pAPPPYSE	0.439	0.186%	2.0738	nucleus
Basophilic serine/threonin e kinase group (Baso_ST_kin)	Clk2 Kinase (Clk2_Ki n)	S493	ARRPRAR sVDALDD L	0.497	0.144%	0.7609	nucleus

Motif Group	Motif	Site	Sequence	Score	Percentile	Surface Accessibility	Localization
Basophilic serine/threonine kinase group (Baso_ST_kin)	AMP Kinase (AMPK)	T327	PSIYAPSt YAHLSPA	0.515	0.133%	1.4762	nucleus
Phosphoserine/ threonine binding group (pST_bind)	Nek10 S (Nek10_ S)	S520	PTSNGGRs RAYMPPR	0.607	0.196%	1.8455	unknown / NA