Effect of *mfsd8* deletion on the secretome and transcriptome of *Dictyostelium discoideum*

A Thesis Submitted to the Committee on Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Master of Science in the Faculty of Arts and Science

TRENT UNIVERSITY
Peterborough Ontario, Canada

Environmental and Life Sciences M.Sc. Graduate Program
May 2025

Abstract

Effect of *mfsd8* deletion on the secretome and transcriptome of *Dictyostelium discoideum*

Joshua Gray

Mutations in the *CLN7* (*MFSD8*) gene, causes CLN7 disease, a subtype of neuronal ceroid lipofuscinosis. MFSD8 is a lysosomal transmembrane protein that transports chloride across membranes. Experimentation regarding *Dictyostelium discoideum* revealed that mfsd8 deficiency altered lysosomal enzyme activity. During starvation, the aggregation of mfsd8$^-$ cells was delayed, and cells formed more mounds that were smaller in size, phenotypes that were attributed to reduced cell-substrate adhesion and altered lysosomal enzymatic activities. This study examines the possible transcriptomic and secretomic basis for these phenotypes. This work generated new datasets for examining the effect of mfsd8 loss on the transcriptome and secretome. The validity of these datasets was supported by use of western blotting and RT-PCR along with a set of assays probing relevant biological processes. Together these results elucidate the biological mechanisms behind the observed phenotypes and lay the foundation for future studies to further study the cellular role of MFSD8.

Keywords: NCLs, Battens disease, CLN7, MFSD8, Transcriptome, Secretome,

Dictyostelium discoideum

Acknowledgements

First, I would like to acknowledge and provide my continuing gratitude for my supervisor Robert Huber. Without his continuous support and guidance, I would not have been able to complete this thesis. His ability to allow for me to be me and provide constructive criticism of my work has allowed me to grow both personally and professionally as a researcher. His unwavering professional mentorship has been a metaphorical lighthouse when I was struggling under the tide of my work. He is a constant beacon for me to improve not only in my critical thinking but also my writing and public speaking in the opportunities he provided.

I would like to thank my community members, Barry Saville and Sanela Martic as without their time, effort, and comments during our meeting this work would not be what it is. I would further like to acknowledge that without the continual support of Doctors Huber, Martic and Saville along with their constant ability to keep me in good spirits during our meetings I would not have been able to complete this research.

I would like to thank both the past and present members of Huber lab I worked with, as without them my work on this topic would be less substantive. Their honest engagement created the ability to bounce ideas and discuss not only my thesis work but theirs which contributed substantially to this work. I would like to individually thank William Kim for his support and bioinformatics expertise, along with Ellicia Yap for teaching me how to work with *Dictyostelium discoideum*. Without them I would not have been able to begin this journey.

I would like to also thank my friends and family for being a constant port during any tumultuous times. Their constant support and encouragement have allowed me to breakthrough my bouts of writing block and complete this important work. Without them I would be a husk, without the ability to do anything and as such they have my continuing gratitude and admiration.
While this has been both a humbling and joyous experience in ways that I never expected at the start, I want to state that this has been a dream for me since I was young. The ability to do research, to find out more about the universe, to make a mark and contribute to the never-ending scientific march of discovery, is a dream come true. In conclusion I would like to thank anyone who reads this outside of my personal experience, as that is what will make this work worthwhile.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AMP(X)	Adhesion Modulation Protein X
ATP	Adenosine triphosphate
ATP13A2	ATPase 13A2
cAMP	Cyclic adenosine monophosphate
Cfa(X)	Counting Factor-Associated (X)
CSPα	Cysteine-string protein alpha
Ctn(X)	countin (X)
CTS(X)	Cathepsin (X)
Cpr(X)	Cysteine Proteinase (X)
DNAJC (5)	DnaJ homolog subfamily C member (5)
DEG	Differentially expressed genes
Exoc(2)	EXOcyst Component (2)
Gdc(A)	Gp64 and Disintegrin-like, Cysteine-rich protein (A)
GO	Gene ontology
irle	IRE-Like kinase
KCTD7	Potassium Channel Tetramerization Domain Containing 9
LAGO	Logically accelerated gene ontology
mbtps1	Membrane-Bound Transcription factor Peptidase, Site 1
MFS	Major Facilitator Superfamily
MFSD8	Major Facilitator Superfamily Domain 8
Mmm(1)	Maintenance of Mitochondrial Morphology (1)
mTOR	mammalian/mechanistic target of rapamycin
mTORC1	mTOR Complex
NCL	Neuronal Ceroid Lipofuscinosis
PdiA	PhosphoDiesterase Inhibitor A
PdsA	Phosphodiesterase A
Pig(O)	PhosphatidylInositol Glycan, class (O)
PGRN	Progranulin
PPT1	Palmitoyl-Protein Thioesterase 1
PsmB6	ProteaSoMe Beta 6
RT-PCR	Reverse transcript polymerase chain reaction
Sad(A)	Substrate Adhesion (A)
Sec(61)(b)	SECretory 61 complex Beta subunit
Sib(C)	Similar to Integrin Beta (C)
Sml(A)	SMaLl aggregates
Tgr(#)	tgr (tiger), **T**ransmembrane, **I**PT, I**G**, **E**-set, **R**epeat protein
Timm(17)	Translocase of Inner Mitochondrial Membrane (17)
TRPML(1)	Transient receptor potential-mucolipin (1)
Ucp(A)	Mitochondrial substrate carrier family protein Uncoupler protein (A)
Vat(X)	Vacuolar ATPase subunit C
Vsp(#)	Vacuolar Protein Sorting

Chapter 1 - General introduction

1.1 Neuronal ceroid lipofuscinoses

Neuronal ceroid lipofuscinoses (NCLs) are a family of inherited neurodegenerative lysosomal storage disorders (Mole & Coleman, 2015). NCLs occur in populations of all ethnicities and have a worldwide incidence rate ranging from 2 of 100,000 live births in the United States (Williams, 2011). In 2008 a study found that the rate of NCLs in Newfoundland Canada was 13 of 100,000 human live births (Moore et al., 2008), which could be attributed to a genetic founder effect where a small number of individuals passed on the gene to descendants and limited migration amplified the incidence (Mole and Cotman, 2015). The disorder is characterized by the accumulation of ceroid lipofuscin, which is an autofluorescent lipoprotein that normally accumulates as cells age. Ceroid lipofuscin is also generated from pathological conditions such as disease, malnutrition, and cell stress. The exact makeup of the storage material varies between the different subtypes of NCL. However, most of the accumulating ceroid lipofuscin storage material present within malfunctioning lysosomes is composed of sphingolipid activator proteins A and D and/or subunit C of mitochondrial ATP synthase (Haltia, 2006). This accumulation of materials in the lysosomes of cells leads to the deleterious effects on several cellular processes including cellular trafficking, which importantly is present as neuroinflammation, and eventual neuronal cell death. This loss of neuronal cells is one of the reasons behind the symptoms present in cases of Batten disease which include seizures, a progressive loss of both mental and motor capabilities and eventual death (Schulz et al., 2013).

In humans, Batten disease is due to mutations in one of the 13 distinct ceroid lipofuscinosis neuronal (*CLN*) genes with each gene generating its own NCL subclass (Butz et al., 2020) (Table 1.1). The protein products of *CLN* genes are found throughout the endomembrane system. CLN1/PPT1, CLN2/TPP1, CLN5, CLN10/CTSD and CLN13/CTSF are found in the lysosomal lumen, CLN3, CLN7/MFSD8 and CLN12/ATP13A2 are lysosomal transmembrane proteins, CLN6 and CLN8 exist within the endoplasmic reticulum, CLN4/DNAJC5/CSPα and CLN14/KCTD7 are cytosolic proteins associated with vesicular membranes, while CLN11/PGRN is a secreted glycoprotein (Cárcel-Trullols et al., 2015) (Figure 1.1). As the functions of these proteins are different and *CLN* genes localize throughout the endomembrane system, the categorization as CLN proteins is due to mutations in them triggering the hallmarks of Batten disease. Continued research into the effects of the Batten disease subtypes has elucidated interaction between CLN proteins, providing a framework for possible shared biological pathways that require their functionality, but it is currently unclear how the loss of these proteins triggers the disease progression.

Table 1.1. List of Genes associated with NCLs along with their functions and homologs in *D. discoideum*

NCL Gene	***Associated NCL disease***	***Function***	***Alternate name***	***Homolog in D. discoideum***
CLN1	Infantile	Palmitoyl-protein thioesterase	PPT1	Ppt1
CLN2	Late-Infantile	Serine protease	TPP1	Tpp1
CLN3	Juvenile	Transmembrane protein unknown function		Cln3
CLN4	Adult-onset, Kufs disease, Parry disease	Hsc70 co-chaperone	DNAAJC5/CSPα	Ddj1
CLN5	Late-infantile	Glycoside hydrolase		Cln5
CLN6	Variant late-infantile	Transmembrane protein unknown function		No homolog
CLN7	Late infantile	Transmembrane Cl^- transporter	MFSD8	Mfsd8/Cln7
CLN8	Northern epilepsy	Trafficking receptor		No homolog
CLN10	Congenital, neonatal and late infantile	Aspartic protease	CTSD	CtsD
CLN11	Adult-onset,	Secreted glycoprotein	PGRN	Grn
CLN12	Juvenile-onset	Transmembrane protein unknown function	ATP13A2	Kil2
CLN13	Adult onset	Cysteine protease	CTSF	CprA, CprB
CLN14	Infantile	Ubiquitin ligase linker	KCTD7/CUL3-E3	Kctd9

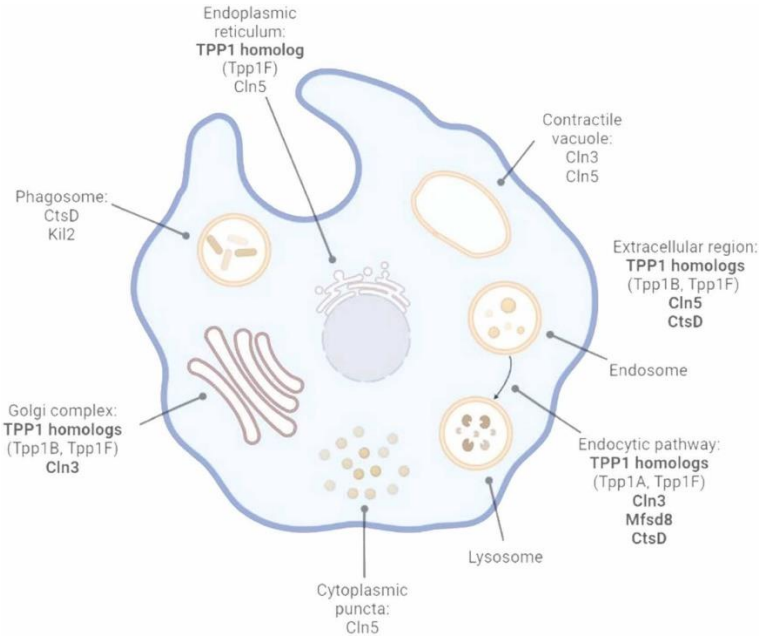

Figure 1.1. Localization of CLN proteins homologs within *D. discoideum* cells. CLN protein homologs with identical localization in mammalian cells have been bolded. Figure taken from Remtulla and Huber 2023.

While there are no cures for the different NCL subtypes, cerliponase alfa is a clinically approved enzyme replacement therapy for CLN2 disease, which is caused by mutations in *TPP1* (Schulz et al., 2018). Cerliponase alfa has been shown to drastically slow the progression of CLN2 disease. In CLN1 disease the use of recombinant PPT1 has been effective in mouse models of the disorder (Schulz et al., 2018). The use of Immunomodulatory Agents has shown beneficial therapeutic outcomes in *Ppt1* and *Cln3* mouse knockout models. Gene therapies have also been researched and in cases of CLN1, CLN2. CLN3 and CLN5 disease, the use of adeno-associated virus linked to the CLN proteins have shown beneficial therapeutic outcomes (Rosenberg et al., 2019).

The *CLN* gene of interest within this thesis is *CLN7,* also known as a major facilitator superfamily domain containing 8 (MFSD8). CLN7 disease is caused by mutations within the *CLN7* gene (Elleder et al., 2011). It is the late infantile onset version of Batten disease as the symptoms typically present between 1.5 and 6 years of age (Elleder et al., 2011). Symptoms typically progress from epileptic seizure to a loss of both motor and mental function, loss of the ability to speak, involuntary muscle contractions, changes in personality and eventual death (Elleder et al., 2011). Currently, in humans, a total of 88 distinct mutations have been cataloged (University College London, 2024).

1.2 MFSD8 and CLN7 disease

MFSD8 is a 518-amino acid, 58 kDa polytopic integral membrane protein that contains 12 transmembrane domains (Siintola et al., 2007). As denoted by its name MFSD8 is a member of the major facilitator superfamily of transporter proteins and as such contains an MFS domain (Figure 1.2) (Siintola et al., 2007). Members of MFS are known to transport small substrates such as ions, sugars, nucleosides, amino acids, and drugs across membranes using chemiosmotic ion gradients and MFSD8 was recently revealed to transport chloride ions (Wang et al., 2021). In humans MFSD8 is expressed in varying amounts throughout the body. The biochemical analysis of both human and rat lysosomes along with colocalization experiments in cultured mouse hippocampal neurons, COS-1 fibroblast-like cells, HeLa cervical cancer cells, *Drosophila melanogaster* and *Dictyostelium discoideum* have confirmed MFSD8 localizes to both late endosomes and lysosomes along with the localization being true for all observed species showing

that the localization is evolutionary conserved (Bagshaw et al., 2005; Schröder et al., 2007; Huber et al., 2020; Mohammed et al., 2017; Sharifi et al., 2010; Siintola et al., 2007). After translation and processing, human MFSD8 is targeted to the lysosome due to a N-terminal dileucine motif. At the lysosome, MFSD8 is proteolytically cleaved twice at two N-glycosylation sites, positions N371 and N376 (Steenhuis et al., 2012). The enzyme cysteine protease cathepsin L mediates one of these cleavages while the other enzyme is unknown (Steenhuis et al., 2012).

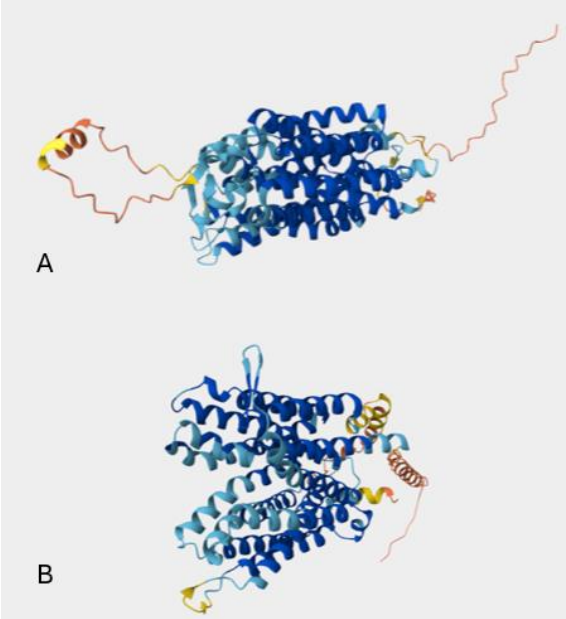

Figure 1.2. Protein structures of Human MFSD8 (A) and *Dictyostelium discoideum* Mfsd8 (B). Image generated using AlphaFold Protein Structure Database (Jumper et al 2021; Varadi et al 2023). The PDB id of MFSD8 and Mfsd8 are Q8NHS3 and Q8T2G9 respectively. of MFSD8 and Mfsd8 have a 29% and 47% match for identical and similar identity.

While MFSD8 functions as a lysosomal chloride channel, MFSD8 also has a host of alternate functions. It has roles in the regulation of lysosomal calcium content as it can reduce lysosomal membrane potential which activates the TRPML1 channel and triggers the release of lysosomal calcium ions. It also has a role in the regulation of endolysosomal compartment pH and research has shown that MFSD8 contributes to the acidification of the endosome as it transitions to a lysosome (Wang et al., 2021), there is a large amount of evidence indicating that the loss of MFSD8 negatively impacts lysosomal function. In mammalian studies the loss of MFSD8 was shown to alter the levels of CLN1, CLN5 and CLN10, which is indicative of lysosomal dysfunction as they are known soluble lysosomal enzymes (Huber et al., 2020). Studies in *Cln7*$^{-/-}$ mice also revealed impairments in lysosomal exocytosis (lysosomal β-hexosaminidase in mutant was ~40% of WT) and lysosomal motility (40% less lysosomes migrated to the periphery after reduction of cytoplasmic pH) along with reduced cell survival observed (von Kleist et al., 2019). Additionally in other mouse models of CLN7 disease, there were defects in autophagy and mTORC1 signaling (mutant had 30% of the tubular structures extending from autolysosomes observed in WT) (Danyukova et al., 2018).

Considerable effort has been made in understanding both MFSD8 and CLN7 disease (Wang et al., 2021; Brandenstein et al.,2015: Danyukova et al., 2018; von Kleist et al., 2019; Sharifi et al., 2010; Huber et al., 2020; Yap et al., 2022). This work has revealed many model organisms which can and have been used to not only study MFSD8 but also pathways the CLN genes share, as many of these organisms have homologs for other NCL proteins. These model organisms include mouse (*Mus musculus*), fruit fly

(*Drosophila melanogaster*), zebrafish (*Danio rerio*) and yeast (*Saccharomyces cerevisiae*) along with larger animals like Japanese macaques (*Macaca fuscata*), Chinese crested dog and chihuahuas. *Dictyostelium discoideum*, a simple eukaryote, is seen as a powerful model organism for studying a variety of biological processes and functions of proteins associated with human diseases including the NCLs (Huber, 2016; Huber, 2020; Mathavarajah et al., 2017).

1.3 *Dictyostelium discoideum*

Dictyostelium discoideum is a soil dwelling social amoeba first described by Kenneth Raper in 1935 (Raper, 1935). *D. discoideum* has a 24-hour life cycle which has been revealed to depend on several extracellular signals (Figure 1.3). *D. discoideum* has 2 phases, a growth phase and a starvation phase. In the growth phase cells obtain nutrients through the phagocytotic ingestion of bacteria or through the utilization of micropinocytosis on liquid media. Cells then utilize the nutrients and proliferate by mitosis (Fey et al., 2007). In the second phase, a cascade event throughout several signaling pathways starts the development of multicellularity as a response to lack of nutrients, (Fey et al., 2007). Cells release and aggregate chemotactically to pulses of cyclic AMP (cAMP) (Mathavarajah et al., 2017). As the cells stream together over the course of 10 hours, they become multicellular aggregates, also known as mounds. The mound undergoes many morphological changes resulting in the formation of a finger as the cells differentiate into pre-stock and pre-spore cells. The finger then tips over becoming a motile slug which through various mechanisms transports itself to a location

where the process can end with the terminal differentiation of the pre-stalk and pre-spore which become a fruiting body. The fruiting body contains a stalk consisting of dead vacuolated cells and a mound of spores elevated by the stalk. If spores enter a nutrient rich environment, they germinate to release cells that can re-enter the *D. discoideum* life cycle as growth phase cells (Fey et al., 2007).

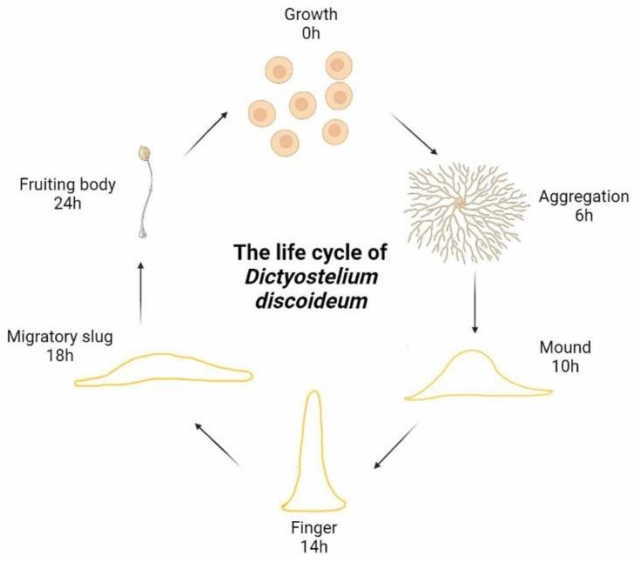

Figure 1.3. Image depicting the life cycle of *Dictyostelium discoideum*. In nutrient abundant environments divide mitotically and grow in a unicellular state however in response to starvation cells undergo a 24-hour developmental cycle which concludes in the cells differentiating into a multicellular fruiting body. If the spores withing the fruiting bodies receive nutrients they can germinate restarting the cycle. Figure taken from Remtulla and Huber 2023.

Dictyostelium discoideum is an excellent model organism for many conserved processes and both in modern times and historically it has been used to examine a host of human diseases including cancer, mitochondrial syndromes, and many neurological conditions (Huber et al., 2020a; Remtulla and Huber, 2023). One of the main reasons for

the common use of *D. discoideum* as a model organism, is the ease at which researchers are able to genetically modify the organism in combination with its relatively short life cycle (Remtulla and Huber, 2023). This combination allows researchers to examine a host of effects a mutation has on key biological processes within 24 hours. As a result of *D. discoideum* having conserved biological processes, findings in it are traceable to similar scenarios and findings within mammalian systems. This becomes readily apparent when examining CLN diseases as *D. discoideum* has homologs for 11 of the 13 total NCL proteins (CLN1-CLN5, CLN7, CLN10-CLN14) more than both fruit flies and yeast (Remtulla and Huber, 2023). As *D. discoideum* hosts 11 of 13 NCL proteins and shares biological processes with humans, *D. discoideum* is one of the efficacious model organisms for studying NCL protein related disease (Remtulla and Huber, 2023).

Regarding MFSD8 specifically *D. discoideum* has an ortholog of human MFSD8 called mfsd8 (DDB0307149) which carries a 29% and 47% match for identical and similar identity respectively (Huber et al., 2020). There is a large span of coverage between the two as amino acids 7–506 of human MFSD8 align with amino acids 30–483 of the Dictyostelium ortholog. This span includes 32 mutations documented in human CLN7 disease patients and 14 mutations are conserved in the Dictyostelium ortholog (Huber et al., 2020). Unlike human *MFSD8* which spans 13 exons, is 518 amino acids long, and has 33 computationally mapped possible isoforms, the *D. discoideum mfds8* is encoded as one exon and its protein product spans 498 amino acids (Huber et al., 2020). The peak of *mfsd8* expression is 8 hours post the induction of starvation and the start of multicellular

development, but it is transcribed throughout the *D. discoideum* life cycle (https://dictyexpress.research.bcm.edu/landing/) (Huber et al., 2020; Rot et al., 2009)

1.4 Secretion and Transcription

Secretion is incredibly important in our examination of NCL. It is known that in *D. discoideum*, 5 of the 11 homologs of human NCL-related proteins are secreted during development, those being palmitoyl protein thioesterase 1 (Human PPT1; *D. discoideum* Ppt1), tripeptidyl peptidase 1 (Human TPP1; *D. discoideum* Tpp1B, Tpp1C, Tpp1F), ceroid lipofuscinosis neuronal 5 (Human CLN5; *D. discoideum* Cln5), cathepsin D (Human CTSD; *D. discoideum* CtsD) and cathepsin F (Human CTSF; *D. discoideum* CprA , CprB, CprD, CprE, CprF, CprG, uncharacterized protein DDB0252831) (Bakthavatsalam and Gomer, 2010; Huber, 2017). Work examining other NCL disease models has shown that the deficiency in other NCL proteins can impact protein secretion like in the case of *cln3*-deficiency (Huber, 2017). Recent work has highlighted that *mfsd8* loss could have a similar effect as shown by Huber et al (2020) with modulations in the levels of secreted Cln5 and cathepsin D.

The transcriptome is also incredibly important in our examination of NCL conditions and the multicellular development of *D. discoideum*. The transcriptome is all the RNA transcripts that the cells are producing as such changes to it can have massive implications of the activities of the cells. Previous work showed that *cln3* and *cln5* deficiency altered the transcriptome and provided insight into the phenotypes caused by the loss of *cln3* and *cln5* (Kim and Huber, 2022; Huber and Mathavarajah, 2019). Thus,

examining the transcriptomes of *mfsd8*-deficient *D. discoideum* should elucidate new avenues of research (Kim and Huber, 2022; Huber and Mathavarajah, 2019).

1.5 Hypothesis and rationale

Recent research examining the effects of *mfsd8* loss on *Dictyostelium discoideum* has elucidated several new avenues of research. The loss of *mfsd8* causes several observable phenotypes that warrant further examination. During the unicellular growth stage, it was shown that loss of *mfsd8* repressed growth, impaired cytokinesis, increased cell proliferation, and increased cell size. While during multicellular development, loss of the gene delayed aggregation, reduced cell substrate adhesion and led to the aberrant secretion of several proteins including Cln5 and cathepsin D (Yap et al., 2022). All these phenotypical differences between the *mfsd8* knock out and wild type need to be examined, as understanding the biological mechanisms behind the development of these phenotypes will inform us about possible mechanisms that the NCL disorders and more specifically CLN7 disease, acts on. I hypothesize that alterations in the transcriptome and secretome lead to the development of these phenotypes and thus predict that there will be fluctuations in both the RNA transcribed for the transcriptome and protein released regarding secretome that are related to the phenotypes observed in previous research (Yap et al., 2022).

Based on historic and recent research, I hypothesized that both the transcriptome and the secretome will be perturbed in *Dictyostelium* after the loss of

mfsd8 and predict that the alterations observed in both systems will align with the previously observed phenotypes (Yap et al., 2022) (Figure 1.4.). As such, this study strives to uncover the transcriptomic and secretomic differences between a *mfsd8*$^{-}$ strain of *D. discoideum* and its wild type parental strain. Findings from this study could provide insight into the links between *mfsd8*-deficiency and CLN7 disease pathology in relation to the biological mechanisms and cellular pathways known to exist within *D. discoideum*. These links could then be translated to mammalian models of CLN7 disease and humans which could provide new insight for therapy development.

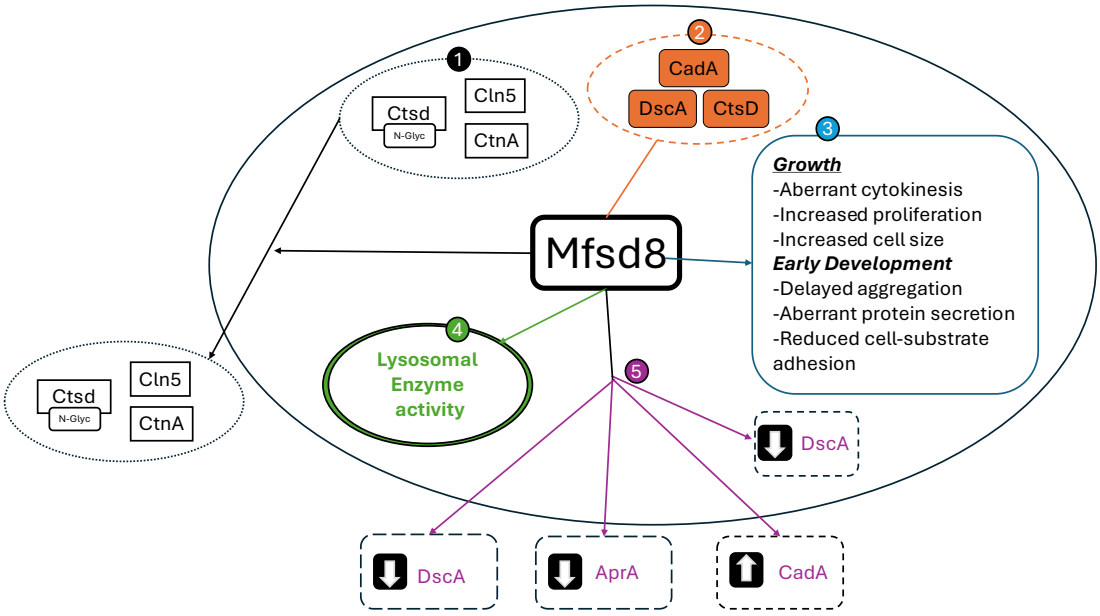

Figure 1.4. Model of our current understanding of the functions of Mfsd8 in *D. discoideum*. (1 Black) Mfsd8 regulates the secretion of Cln5, CtsD and CtnA during early stages of development. (2 Orange) Mfsd8 interacts with CadA,CtsD and DscA during growth and the early stages of development. (3 Blue) The loss of *mfsd8* causes several phenotypical including increased proliferation during growth and delayed aggregation during the early stages of development. (4 Green) The loss of *mfsd8* affects the activity

of lysosomal enzyme activity during growth and the early stages of development. (5 Purple) During growth and 4 hours into multicellular developments the loss of *mfsd8* lowers extracellular and intracellular amounts of DscA. The loss of *mfsd8* also lowers the extracellular amount of AprA during growth and increases the amount of extracellular CadA 4 hours into development. (Yap et al., 2022)

Chapter 2- Differential gene expression caused by *mfsd8* loss in *D. discoideum*

2.1 Introduction

Mutations to one of the 13 ceroid lipofuscinosis neuronal (CLN) genes (CLN1-8, CLN10-14) leads to neuronal ceroid lipofuscinosis (NCL), also known as Batten disease (Mole & Cotman, 2015). The CLN proteins fulfill a vast number of functions within the cell; PPT1/CLN1, TPP1/CLN2, CLN5, CTSD/CLN10 and CTSF/CLN13 are all soluble lysosomal proteins; CLN3, MFSD8/CLN7, and ATP13A2/CLN12 are endolysosomal membrane proteins; CLN6 and CLN8 are components of the endoplasmic reticulum; and DNAJC5/CLN4 is a component of the synaptic vesicle (Huber 2020). The NCL family of neurodegenerative diseases are characterized by the accumulation of autofluorescence lipid-protein aggregates called lipofuscin, in the lysosomal compartment of cells. lipofuscin aggregates can present in various tissues including skeletal muscle, kidney and retinal cells and for cells, this accumulation of lipofuscin leads to dysfunction. In animals the accumulations of these ceroid lipofuscin aggregates lead to a host of symptoms as neuronal cells are lost. The symptoms include progressive loss of motor and cognitive function, epilepsy, vision loss and a reduced lifespan (Schulz et al., 2013). The severity and onset of these symptoms are directly related to both the specific gene affected and the specific mutation in the gene.

Mutations in *CLN7,* also known as a major facilitator superfamily domain containing 8 protein (*MFSD8*), causes a late infantile version of NCL (Siintola, 2007). Symptoms generally present between ages 2-7. The MFSD8 protein is an integral membrane protein. MFSD8 functions as an endolysosomal chloride channel (Wang et

al., 2021). While the exact function of MFSD8 in relation to CLN7 disease is not known, in mammals, it has been associated with autophagy, neuronal cell survival, and lysosomal size, trafficking, and pH regulation, all important cellular process in the progression of NCL pathology (Wang et al.,2021).

Dictyostelium discoideum (*D. discoideum*) is a social amoeba that has been used as a model organism for a variety of disorders and conserved cellular and developmental processes since its identification in 1935 (Raper, 1935). *D. discoideum* has homologs of 11 of the 13 *CLN* genes and has been previously used to study the pathology underlying several forms of NCL (Remtulla & Huber, 2023). The *D. discoideum* life cycle begins in a haploid single cell state in which it can phagocytise bacterium, intake nutrients and multiply by mitosis. In nutrient poor condition over the course of 24 hours, *D. discoideum* undergoes a series of morphological changes which ends with the production of a fruiting body of viable spores and a stalk that supports the body (Mathavarajah et al., 2017). Previous work regarding the role of *mfsd8* in *D. discoideum* has illustrated several phenotypic differences playing an essential role in several conserved cellular processes such as proliferation, pinocytosis, cytokinesis, adhesion, delayed aggregation, and protein secretion (Yap et al., 2022).

In this study, comparative transcriptomics was used to examine the genes affected by the loss of *mfsd8* during the growth stage and 8 hours post starvation, which is when *mfsd8* is most highly expressed during the *D. discoideum* life cycle (Figure 2.1). There were no genes implicated that directly related to previously reported *mfsd8*$^{-}$ phenotypes during growth (Yap et al., 2022). During starvation, genes affecting protein

secretion and aggregation were observed in aberrant amounts in the *mfsd8* knockout strain. The effect on gene expression was also observed in RT-PCR and subsequent assays, which together further support our RNA-seq analysis. This study elucidates the effect that *mfsd8* deletion has on *D. discoideum* and furthermore indicates possible avenues in regard to the study of CLN7 disease on mammals.

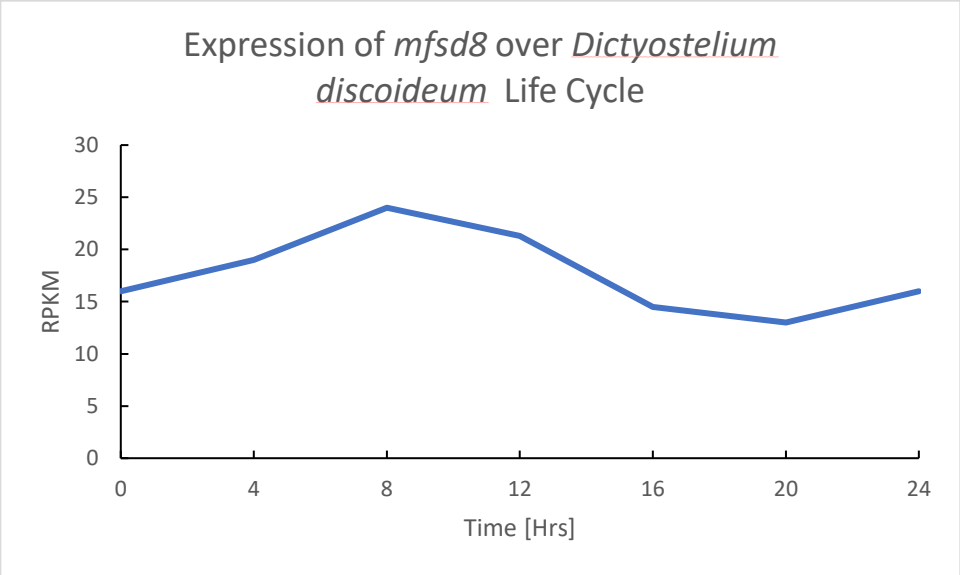

Figure 2.1. Gene expression profile of *mfsd8* throughout the life cycle of *D. discoideum. Data* was obtained from Dictyexpress (https://app.dictyexpress.org/) (Rot et al., 2009). RPKM, Reads per kilo base per million mapped reads.

2.2 Materials and methods

2.2.1 Cell culture, chemicals, media and antibodies

AX4 (DBS0237637, hereafter referred to as wild type, WT (GWDI_448_B_4) and *mfsd8*$^{-}$ Cells were maintained on SM/2 agar containing *Kebsiella Aerogenes* (Fey et al., 2007). Experimental cells were cultured Axenically in flasks containing nutrient rich HL5

medium at 22C and 150rpm (For medium, Hunstanton, Norfolk, United Kingdom). The HL5 medium was supplemented by ampicillin (100ul/ml) and streptomycin sulfate(300ug/ml) as a preventative for bacterial growth. Blasticidin S hydrochloride (10 µg/ml) was used to select for *mfsd8*$^{-}$ cells. For experimentation Mid-log phase cells (1-5x10^{6}cell/ml) were harvested and added to flasks containing fresh HL5 media containing ampicillin (100 µg/ml) and streptomycin sulphate (300 µg/ml) and left for two doubling times. The next day mid-log phase cells were taken from the flasks and added to falcon six well plates. The cells were allowed to adhere for an hour at which time the cells were washed with 2ml of KK2 buffer (0.7 g/L K2HPO4, 2.2 g/L KH2PO4, pH 6.5) two times. After the repeated washing, 1ml of KK2 was added to the well and it was left for 8 hours. Cells were imaged at each time point using a Nikon Ts2R-FL inverted microscope equipped with a Nikon 10 Digital Sight Qi2 monochrome camera (Nikon Canada Incorporated Instruments Division, Mississauga, Ontario, Canada). After 8 hours the wells were imaged again, then washed again with 2 ml two times and cells were harvested from the well with 1ml of ice cold kk2 and added to 1.5ml microfuge tubes. The cells were then pulsed in a Sorvall legend microfuge 21 (Thermo Fisher Scientific, Whitby, Ontario, Canada) and media was removed from the tubes. Cells were then frozen using liquid nitrogen and stored at -80˚C for future processing. The same process was used for growth phase samples. Growth-phase and starved cells were recovered from the -80˚C. and allowed to thaw at which point RNA was extracted using the Monarch Total RNA miniprep kit according to the manufacturer's instructions (New

England Biolabs, Whitby, Ontario, Canada). Three biological reps were used for growth and starvation and submitted for RNA sequencing analysis.

2.2.2 RNA sequencing and bioinformatics

RNA sequencing of 3 biological replicates was performed by The Centre for Applied Genomics at the Hospital for Sick Children (Toronto, Ontario, Canada). RNA samples were first prepared using a stranded poly(A) mRNA library preparation kit (New England Biolabs, Whitby, Ontario, Canada), which removed rRNA. RNA quality was determined using Bioanalyzer (Agilent Technologies, Santa Cruz, California, United States of America). The RNA-seq analysis for all mRNA libraries were done through an Illumina NovaSeq SP flowcell PE100 sequencer (Illumina Inc, San Diego, Calafornia, United States of America) and adaptor sequences were removed using the Trimmomatic application (Bolger et al., 2014). The quality of the paired reads was assessed using fastQc (Babraham bioinformatics, Babraham, Cambridge, United Kingdom). The reference genome was acquired from ensemble database and indexed using HISAT2 (hierarchical indexing for spliced alignment of transcripts 2) (Kim et al., 2019). HISAT2 was then used to map paired reads onto the indexed reference genome. Features count (Liao et al., 2014) was used to obtain the read counts. Differential gene expression analysis was preformed using the SARSTools (Varet et al., 2016) package with an alpha value of 0.05. The Benjamin-Hochberg (BH) multiple-testing correction were used in the differential gene expression analysis. The resulting fold change value was converted to $\log_2$ fold change to allow for a symmetrical value around 0 which making negative values

represent lowered amount of transcript and positive values represent increased amounts of transcript. A fold change threshold of 1.25 (Log2FC=0.0.3219) was then applied on the list of differentially expressed genes (DEGs). Logically accelerated gene ontology term finder (LAGO) was used for gene ontology (GO) term enrichment analyses of the DEGs identified by the LAGO database (Boyle et al., 2004). A p-value threshold of 0.05 and the BH correction was applied.

2.2.3 Reverse transcription polymerase chain reaction

Total RNA from WT and *mfsd8*$^{-}$ cells was isolated from frozen samples using a commercially available Monarch® Total RNA Miniprep Kit according to the manufacturer's instructions (New England Biolabs, Whitby, Ontario, Canada). Total RNA of the samples was then quantified using an Agilent BioTek Take3 microvolume plate and applicable software (Agilent Technologies, Santa Clara, Calafornia, United States of America). Oligonucleotides specific for each gene were designed and were aligned against the *D. discoideum* coding sequence database by BLAST to ensure that they were specific for the gene tested. Table 2.1 includes a list of all genes examined. RT-PCR was then preformed using a Luna® universal one-step RT-qPCR kit (New England Biolabs Whitby, ON, Canada). 20 µl reactions contained luna universal one step reaction mix, luna warm start 0.04 µM forward primer, 0.04uM reverse primer, and template RNA.

Table 2.1. List of primers used in reverse transcription polymerase chain reaction

Gene ID	Gene name	Primer
DDB_G0294034	*rnlA*	GAGCGCTAAGGTCCATAGGTCTAAAGGGAAAC
		CTGTACATTGTTGGCTAGAGAACGCCTTAAATTGG
DDB_G0272192	*mfsd8*	GGATTTGCATTCCCAGCAATTATGGGACC
		GACAAGTAACCTGGTGCGGTGTATTTATTGATT3
DDB_G0282397	*mbtps1*	GTTGAAAAGGTGTGGGAAGTGTCTGCAAATAAG
		GACCATTCAATGGCGAGCCGTAAACAC
DDB_G0288195	*sibC*	CAGTCACCACAGCACTTCGTGGTACTTTAT
		GAATCGAGACCACCATTATCTGTCATGAAAAATGG
RT-PCR reactions were performed using the following parameters: 55˚C/10 minutes, 95˚C/1minute, 45 cycles of 95˚C/10 seconds, 60˚C/30 seconds. Data were collected from three biological replicates, with two technical replicates for each condition.		

2.2.4 Autophagy assay

The autophagic capacity of *mfsd8*$^{-}$ cells and WT cells were examined using a Sigma Aldrich autophagy kit (MAK-138) which uses a proprietary florescent autophagosome marker as proxy for autophagic activity (Sigma Aldrich Canada, Oakville, Ontario, Canada) following the manufacturers guidelines. Seven biological replicates of $2.5x10^7$ Mid-log phase cells were harvested from flasks and added to 15ml tubes which were then centrifuged. The pellet was resuspended in 2ml of kk2. This wash was repeated twice and then the pellet was resuspended to a concentration of $4x10^6$ cells per ml. Four millilitres of this was then added to 50ml flasks which were allowed to starve on a 150rpm stationary rotor. After 8 hours, $2.0x10^5$ cells were removed from the flask and added to 1.5 ml microfuge tube this was then centrifuged, and the supernatant was removed. 80 µl of MAK138 working solution was added to the tube and it was allowed to incubate for 45 minutes. After centrifugation of the samples, the pellet was washed with 100 µl of MAK138 wash buffer three times. The pellet was then

resuspended in 100 µl of wash buffer and 50 µl was added to separate wells of a black translucent bottom 96 well plate and florescence (em 360 ex 528 120 gain) was examined using a synergy HTX multimode plate reader. Autophagy values are standardized against activities from WT cells.

2.2.5 Enzyme assays

Protein concentrations were quantified using a Qubit 2.0 (Fisher Scientific Company, Ottawa, Ontario, Canada). Assays were performed in triplicate, with 4 biological replicates and enzyme activities were subtracted by values obtained from blank solutions. Activates values are standardized against activities of Lysates from WT cells. Data were statistically analysed using one-sample t-tests and p-values < 0.05 were considered significant.

2.2.5.1 CTSD assay

CTSD activity was examined using Cathepsin D Activity Assay Kit (Fluorometric) (ab65302) (Abcam, Cambridge, United Kingdom). Both WT and *mfsd8*$^{-}$ cells starved in KK2 for 8 hours were lysed using the provided lysis buffer. All assays used the same 27.5 ug of cell lysate and added to 100 µl of reaction buffer. 2 µl of substrate (GKPILFFRLK(Dnp)-D-R-NH2) labeled with MCA) was then added to this reaction mixture. Samples were then incubated at 37˚C for 2 hours and deposited onto 96-well opaque black bottom plates. The fluorescence of samples was then measured (360±40nm excitation 460±40 nm emission)

2.2.5.2 α-Glucosidase activity assay

α-Glucosidase activity was examined using methods adapted from Wimer et al. (1997) WC lysates in 0.1% NP40 in 0.05 M 2-(N-morpholino) ethanesulfonic acid (MES) (pH 6.5) were added to 0.1 M sodium succinate (pH 6.0) containing 2 mM p-nitrophenyl-α-D-glucopyranoside substrate (487506, Sigma Aldrich Canada, Oakville, Ontario, Canada). Samples were then incubated at 65°C for 2 h and then quenched with two equal volumes of 1 M Na_2CO_3. Activity was measured at 395 nm absorbance. Equal protein masses of 150ug were used.

2.2.5.3 β-Galactosidase activity assay

β-galactosidase activity assay was examined using methods adapted from Maruhn (1976), the reaction contained a reaction solution of 5 mM ortho-nitrophenyl-β-D-galactopyranoside substrate (48712-M, Sigma Aldrich Canada, Oakville, Ontario, Canada) in 100 mM citrate buffer (pH 4.0). The reaction solutions were incubated at 37°C for 1 hour and quenched using an equal volume of 2-amino-2-methyl-propanol/HCl buffer.

2.3 Results

2.3.1 Effect of *mfsd8*-deficiency on the transcriptome

Previous work examining *D. discoideum* linked the function of Mfsd8 to proliferation, pinocytosis, cytokinesis, adhesion during growth and aggregation and

protein secretion during the early stages of multicellular development (Huber et al., 2020; Yap et al., 2022). To gain insight into the molecular mechanisms underlying these *mfsd8*-deficiency phenotypes in *D. discoideum* and the biological pathways impacted by the loss of *mfsd8*, RNA-seq analysis on growth-phase cells and cells starved for 8-hour in KK2 buffer was performed. The 8-hour time point was chosen as it is the point at which *mfsd8* is the highest (Stajdohar et al., 2017). A volcano plot was generated to visualize the profile of DEGs (Figure 2.2). It showed the genes up and down regulated genes with red plotted dots representing the statistically significant DEGs and the black dots representing the genes determined to not be significant ($p<0.05$). Triangles represent genes above the graphs p value limit. The list of DEGs was filtered to include those genes with a fold change of positive and negative 1.25 along with a p-value < 0.05. RNA-sequencing showed that during growth 445 unique genes were differentially expressed (Table S1), while during early multicellular development 4954 unique genes were differentially expressed (Table S2). For growth and early multicellular development this represents 4% and 40% of *D. discoideum* genome which contains an estimated 12000-13000 protein-coding genes (Fey et al., 2019). 280 genes seen in the growth dataset were also observed in the early multicellular development dataset. During growth, several ubiquitin-related genes were upregulated while several cup genes, a family of highly conserved proteins up regulated by the presence of Ca^{2+} were downregulated (Table S1) (Coukell et al., 2004). In early multicellular development the only shared ubiquitin was *ubqD* which was upregulated in both the shared cup proteins were seen to be downregulated. During starvation, the gene families most perturbed were the short-

chain dehydrogenase/reductase family proteins as they were included in the most up and down regulated genes along with. Several mitochondrial related genes (*ucpA, mmm1, mhsp70, timm17*) had some of the most reduced expression. *pig* family genes, which play roles in GPI biosynthesis pathway (AgBase, 2011), and *tgr* family genes (*trgb1, TrgC1)*, which play roles in aggregation during development, along with several of its suppressor genes (*trgC1)* were amongst the highest upregulated expression (Kibler et al., 2003).

During starvation, several of the homologs for other *CLN genes* (*cln2,3 and 5*) were impacted (Table S2). It is known that *mfsd8* is associated with autophagy which is corroborated by our data set. *atg1* and *atg9* were upregulated in growth conditions, while in early multicellular development *atg5, atg7, atg8b, atg9, atg13 and atg17* (Kim et al., 2022) were all upregulated while *atg4-1, atg4-2* and *atg8a* were all downregulated (Kim et al., 2022). During growth, there were no genes that could be directly implicated in phenotypes observed. However, in starvation condition many genes with links to *mfsd8*-deficiency phenotypes were observed (Table 2.2). Regarding delays in aggregation members of the countin factor complex (Tang et al., 2001) were impacted along with genes regulating proper cAMP mediated chemotaxis. For reduced cell substrate adhesion, increases in the expression of *paxB* and decreases in the expression of *vasP* were observed. *paxB* encodes paxillin which is found in the focal adhesions at the cell-substrate interface and allows for the cell to properly adhere to substrates (Duran et al., 2009). *vasP* encodes vasodilator-stimulated phosphoprotein (VasP) which localizes to the leading edge of migrating cells (Bukharova et al., 2005).

VasP is associated with actin filament formation and adhesion. PaxB and VasP have been experimentally observed to affect cell substrate adhesion as over expression lines of *paxB* and null mutants have been observed to decrease cell substrate adhesion (Bukharova et al., 2005, Duran et al., 2009). Regarding altered protein secretion, *vsp*11, *vps*13, *vps*33, *vps*37, *exoc*2, and *sec61b* proteins involved with the transport and secretion of material were also impacted. Interestingly several proteins linked to Parkinson's were affected such ass *deej* and *vps35* (König and McBride, 2024)

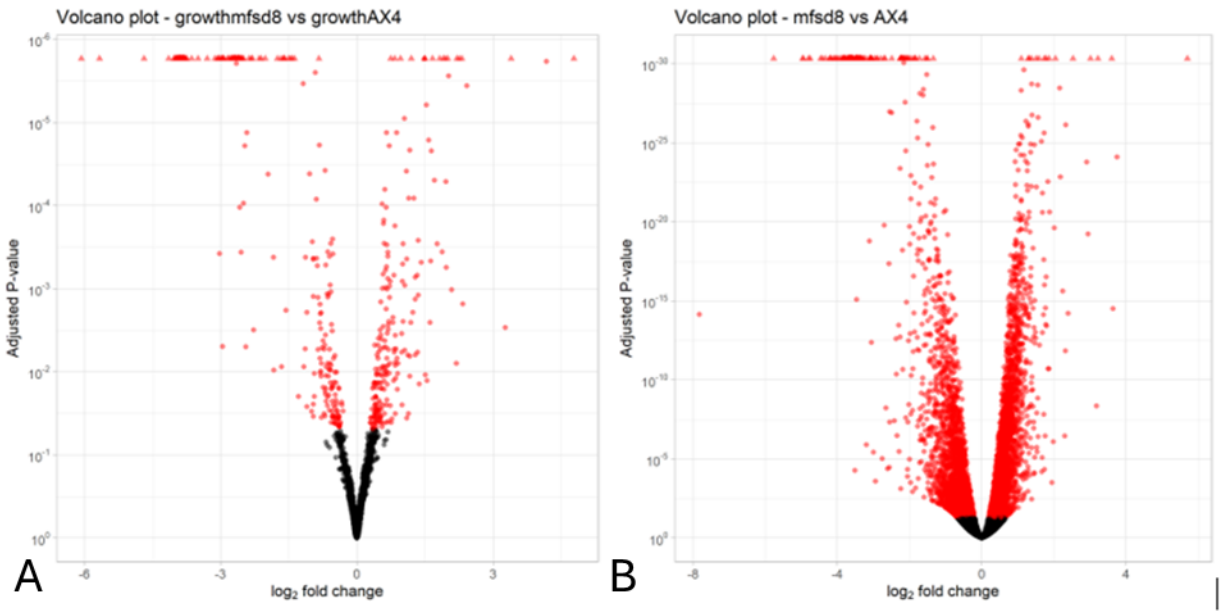

Figure 2.2. Volcano plot depicting differential gene expression in mfsd8$^-$ cells during growth (A) and after 8 h of starvation (B). The plotted red dots represent differentially expressed genes observed with p values above 0.05 and black plotted dots represent genes observed with p values below 0.05. Triangles represent genes with p values above the p value limit in SARTools. Generated by SARTools following differential expression analysis.

Table 2.2. List of genes relating to phenotypical differences observed in the 8-hour starvation transcriptomic dataset

Gene ID	UniProt ID	Gene name	Protein name	log2 Fold Change
DDB_G0289533	*Q54HE9*	*act27 DDB_G0289533*	*Putative actin-27*	*-1.486*
DDB_G0277809	*Q54Z64*	*cfaA DDB_G0277809*	*Counting factor-associated protein A*	*-1.254*
DDB_G0287981	*Q54JL1*	*DDB0187718*	*Uncharacterized protein*	*-1.245*
DDB_G0286893	*Q54L54*	*act29 DDB_G0286893*	*Putative actin-29*	*-0.935*
DDB_G0289157	*Q9U641*	*cmfB cmfr1 DDB_G0289157*	*Conditioned medium factor receptor 1*	*-0.864*
DDB_G0267956	*Q55FT9*	*abiA DDB_G0267956*	*Abl interactor homolog*	*-0.814*
DDB_G0277831	*Q9TX43*	*carD car4 DDB_G0277831*	*Cyclic AMP receptor 4, cAMP receptor 4*	*-0.78*
DDB_G0269234	*P07830*	*act1 act1a DDB_G0289553; act2 act2-1 DDB_G0274133; act4 DDB_G0289005; act5 DDB_G0289663; act6 DDB_G0274135; act7 DDB_G0280545; act8 actA8 DDB_G0269234; act9 DDB_G0274601; act11 DDB_G0288879; act12 DDB_G0274129; act13 DDB_G0274599;*	*Major actin (Actin A1) (Actin A12) (Actin A8) (Actin III) (Actin M6) (Actin-1) (Actin-11) (Actin-12) (Actin-13) (Actin-14) (Actin-15) (Actin-16) (Actin-19) (Actin-2) (Actin-2-sub 1) (Actin-20) (Actin-21) (Actin-3a) (Actin-4) (Actin-5) (Actin-6) (Actin-7) (Actin-8) (Actin-9) (Actin-IEL1)*	*-0.776*

		act14 actB1 DDB_G0274137; act15 actA1 DDB_G0272520; act16 actM6 DDB_G0272248; act19 DDB_G0274727; act20 DDB_G0274285; act21 DDB_G0274561		
DDB_G0290087	*Q54GK9*	*chmp2a1 vps2A DDB_G0290087*	*Charged multivesicular body protein 2a homolog 1 (Vacuolar protein-sorting-associated protein 2A)*	*-0.759*
DDB_G0286545	*Q54LN2*	*vps13D DDB_G0286545*	*Intermembrane lipid transfer protein vps13D (Putative vacuolar protein sorting-associated protein 13D)*	*-0.752*
DDB_G0269902	*Q55CU2*	*act26 DDB_G0269902*	*Putative actin-26*	*-0.684*
DDB_G0287993	*Q54JK4*	*chmp5 vps60 DDB_G0287993*	*Charged multivesicular body protein 5 (Vacuolar protein-sorting-associated protein 60)*	*-0.655*
DDB_G0289541	*Q5TJ65*	*vasp DDB_G0289541*	*Protein VASP homolog (DdVASP)*	*-0.653*
DDB_G0287021	*Q54KZ4*	*chmp6 vps20 DDB_G0287021*	*Charged multivesicular body protein 6 (Vacuolar protein-sorting-associated protein 20)*	*-0.62*
DDB_G0285253	*Q54NF8*	*scrA DDB_G0285253*	*Protein SCAR (Suppressor of cAMP receptor)*	*-0.604*

DDB_G0267708	*Q55GD9*	*vps25 DDB_G0267708*	*Vacuolar protein-sorting-associated protein 25 (ESCRT-II complex subunit VPS25)*	*-0.511*
DDB_G0291728	*Q54E71*	*arpE actr5 DDB_G0291728*	*Actin-related protein 5 (Actin-related protein E)*	*-0.451*
DDB_G0286131	*Q54M91*	*DDB0186819, sec63*	*J domain-containing protein*	*-0.436*
DDB_G0267394	*O96552*	*chmp1 DG1118 vps46 DDB_G0267394*	*Charged multivesicular body protein 1 (Developmental gene 1118 protein) (Vacuolar protein-sorting-associated protein 46)*	*-0.433*
DDB_G0269510	*Q55DV8*	*vps37*	*vps37||Vacuolar Protein Sorting*	*0.028151*
DDB_G0269942	*Q55CR0*	*DDB_G0269942*	*Vesicle-trafficking protein SEC22b (SEC22 vesicle-trafficking protein homolog B)*	*0.342*
DDB_G0272106	*O96621*	*arpB arp2 DDB_G0272106*	*Actin-related protein 2 (Actin-like protein 2) (Actin-related protein B)*	*0.394*
DDB_G0274109	*Q8MML5*	*paxB DDB_G0274109*	*Paxillin-B*	*0.425*
DDB_G0283755	*P42528*	*arpC aclA act DDB_G0283755*	*Actin-related protein 3 (Actin-like protein 3) (Actin-related protein C)*	*0.468*
DDB_G0291097	*Q54F53*	*DDB0189239*	*Uncharacterized protein*	*0.522*
DDB_G0287881	*B0G163*	*exoc5 sec10 DDB_G0287881*	*Exocyst complex component 5 (Exocyst complex component Sec10)*	*0.588*

DDB_G0277863	*P22549*	*pdiA pdi DDB_G0277863*	*Cyclic nucleotide phosphodiesterase inhibitor, PDI*	*0.682*
DDB_G0275007	*P34090*	*cmfA DDB_G0275007*	*Conditioned medium factor, CMF (Density-sensing factor)*	*0.683*
DDB_G0272991	*Q558Z9*	*exoc7 exo70 DDB_G0272991*	*Exocyst complex component 7 (Exocyst complex component Exo70)*	*0.821*
DDB_G0279991	*Q54VZ8*	*exoc8 exo84 DDB_G0279991*	*Exocyst complex component 8 (Exocyst complex component Exo84)*	*0.832*
DDB_G0280081	*Q54VX5*	*exoc2 sec5 DDB_G0280081*	*Exocyst complex component 2 (Exocyst complex component Sec5)*	*0.895*
DDB_G0293520	*Q54BP6*	*exoc3 sec6 DDB_G0293520*	*Exocyst complex component 3 (Exocyst complex component Sec6)*	*1.052*

2.3.2 Gene ontology (GO) term enrichment analyses

GO term enrichment analyses were carried out utilising LAGO (Boyle et al., 2004). During these analyses, genes that were uncharacterized (genes with annotations but unknown functions) were excluded. For growth (Table 2.3) out of the 445 genes, a list of 201 were for biological processes, 197 for cellular components and 208 for molecular function. For starvation out of the 4954 genes, a list of 3192 were for biological processes, 3405 for cellular components and 3312 for molecular function. During growth, differentially expressed genes associated with the loss of *mfsd8* were linked to several biological processes including, responses to stimulus (51), response to

stress (30), response to cation/salt stress (4), socially cooperative development (28) DNA integration (12). Differentially expressed genes primarily localized to the membrane (100), and cell periphery (28). Differentially expressed genes primarily were involved in binding (112), (carbohydrate (11) ubiquitin protein (5)) and protein C-terminal carboxyl O-methyltransferase activity (2) while others had ABC-type transporter activity (6)). There were no GO terms directly affecting the observed growth phenotypes observed, however there could be indirect interactions contributing to the observed phenotypes which should be examined in future experimentation.

Table 2.3. GO term enrichment analysis of differentially expressed genes during growth

Biological processes			
GOID	***Term***	***# of genes***	***% of total***
Martic GO:0050896	*Response to stimulus*	*51*	*11*
GO:0051716	*Cellular response to stimulus*	*40*	*8.9*
GO:0007154	*Cell communication*	*31*	*7.0*
GO:0006950	*Response to stress*	*30*	*6.7*
GO:0048856	*Anatomical structure development*	*29*	*6.5*
GO:0032502	*Developmental process*	*29*	*6.5*
GO:0030587	*Sorocarp development*	*28*	*6.3*
GO:0051703	*Biological process involved in intraspecies interaction between organisms*	*28*	*6.3*
GO:0099120	*Socially cooperative development*	*28*	*6.3*
GO:0006259	*DNA metabolic process*	*23*	*5.2*
GO:0015074	*DNA integration*	*17*	*3.8*
GO:0043157	*Response to cation stress*	*4*	*0.9*
GO:0009651	*Response to salt stress*	*4*	*0.9*
Molecular function			
GOID	***Term***	***# of genes***	***% of total***
GO:0005488	*Binding*	*113*	*25.3*
GO:0030246	*Carbohydrate binding*	*11*	*2.4*
GO:0140359	*ABC-type transporter activity*	*6*	*1.3*
GO:0031386	*Protein tag*	*5*	*1.1*
GO:0031625	*Ubiquitin protein ligase binding*	*5*	*1.1*
GO:0044389	*Ubiquitin-like protein ligase binding*	*5*	*1.1*
GO:0004671	*Protein C-terminal S-isoprenylcysteine carboxyl O-methyltransferase activity*	*2*	*0.5*
GO:0003880	*Protein C-terminal carboxyl O-methyltransferase activity*	*2*	*0.5*
Cellular component			
GOID	***Term***	***# of genes***	***% of total***
GO:0110165	*Cellular anatomical entity*	*154*	*34.6*
GO:0016020	*Membrane*	*100*	*22.4*
GO:0031224	*Intrinsic component of membrane*	*73*	*16.4*
GO:0016021	*Integral component of membrane*	*70*	*15.7*

GO:0071944	*Cell periphery*	*28*	*6.3*
GO:0031225	*Anchored component of membrane*	*5*	*1.1*

During starvation (Table 2.4, Table S3) differentially expressed genes were involved in many biological processes including establishment of localization (605), gene expression (291), nitrogen compound metabolic process (303), catabolic processes (351), vesicle mediated transport (267), developmental processes (310), locomotion (157) actin filament-based processes (146) and response to endoplasmic reticulum stress (42). Differentially expressed genes were localized cytoplasm (1636), nucleus (762), mitochondrion (219), endoplasmic reticulum (211). Like in growth many genes associated with the loss of *mfsd8* related to binding (organic cyclic binding (1117), protein binding (532), small molecule (624)) catalytic activity (hydrolase activity (584), transferase activity (558), peptidase activity (128), kinase activity (184)). During starvation several GO terms directly relating to the observed phenotypical differences between WT and *mfsd8*$^{-}$ cells were observed. Aggregation, locomotion, chemotaxis and aggregation involved in sorocarp development were just some of the enriched GO terms. The GO terms directly relating to altered secretion include Export from cell, secretion from cell and exocytosis. Together these results show that during growth and starvation the loss of *mfsd8* effect affects several response processes.

Table 2.4. Selection of GO term enrichment analysis of differentially expressed genes after 8 hours of starvation

Biological process			
GOID	***Term***	***# of genes***	***% of total***
GO:0065007	*Biological regulation*	*881*	*17.8*
GO:0006810	*Transport*	*595*	*12.0*
GO:0006950	*Response to stress*	*359*	*7.2*
GO:0009056	*Catabolic process*	*351*	*7.1*
GO:0032502	*Developmental process*	*310*	*6.4*
GO:0016192	*Vesicle-mediated transport*	*267*	*5.4*
GO:0006508	*Proteolysis*	*235*	*4.7*
GO:0040011	*Locomotion*	*157*	*3.2*
GO:0030029	*Actin filament-based process*	*146*	*2.9*
GO:0006935	*Chemotaxis*	*104*	*2.1*
GO:0007005	*Mitochondrion organization*	*61*	*1.25*
GO:0046903	*Secretion*	*48*	*1.05*
GO:0006914	*Autophagy*	*40*	*0.85*
Molecular function			
GOID	***Term***	***# of genes***	***% of total***
GO:0016787	*Hydrolase activity*	*584*	*11.8*
GO:0016740	*Transferase activity*	*558*	*11.3*
GO:0140096	*Catalytic activity, acting on a protein*	*435*	*0.8*
GO:0016462	*Pyrophosphatase activity*	*178*	*0.4*
GO:0030234	*Enzyme regulator activity*	*177*	*0.4*
GO:0022857	*Transmembrane transporter activity*	*161*	*0.3*
GO:0004672	*Protein kinase activity*	*132*	*0.3*
GO:0008233	*Peptidase activity*	*128*	*0.3*
Cellular component			
GOID	***Term***	***# of genes***	***% of***
GO:0012505	*Endomembrane system*	*439*	*8.9*
GO:0005739	*Mitochondrion*	*219*	*4.4*
GO:0005783	*Endoplasmic reticulum*	*211*	*4.3*
GO:0030139	*Endocytic vesicle*	*192*	*3.9*
GO:0005777	*Peroxisome*	*39*	*0.1*

GO:0000502	*Proteasome complex*	*24*	*0.1*

2.3.3 Loss of *mfsd8* affects the activity of lysosomal enzymes

The degradation of intracellular material during growth and starvation is carried out by lysosomal enzymes. Previous work examining the early stages of multicellular development showed that several lysosomal enzymes had aberrant activity at the 4-hour time point (Yap et al., 2022). Our comparative transcriptomics showed that the transcription of several known lysosomal enzymes was modified after the 8-hour timepoint in our knockout mutant (Table 2.5). These included a decrease of tripeptidyl-peptidase *1* and increases in β-*galactosidase (1 and 2)* along with *ctsD* and α -glucosidase (1 and 2). As such, we assessed the effect of *mfsd8* deletion on the activity of some of the lysosomal enzymes shown to have abnormal expression at 8 hours (Figure 2.3). We observed a significant increase in activity of α-glucosidase, β-glucosidase and CtsD which mirrored the transcriptomic upregulation observed. Combined, these results show that the loss of *mfsd8,* alters the activity of several known lysosomal enzymes during aggregation, and this supports the results found in the comparative transcriptomic assessment.

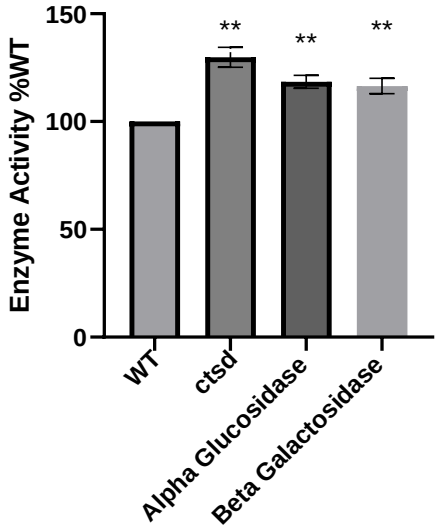

Figure 2.3 Loss of mfsd8 increases the activity of lysosomal enzymes during starvation. WT and mfsd8-deficient cells during 8-hour KK2 starvation were lysed and the activities of various carbohydrate enzymes were assessed including Ctsd, α-Glucosidase and **β** -Galactosidase using relevant assays as described in the Materials and Methods. Statistical significance was determined using the one-sample t-test. Data obtained from plate reader represented as mean enzyme activity +/- SEM (n=4 biological replicates) **p<0.01.

Table 2.5. List of genes encoding lysosomal enzymes observed in 8-hour starvation transcriptomic dataset.

Gene ID	***UniProt ID***	***Gene name***	***Gene***	***Protein name***	***log2 Fold Change***
DDB_G0269914	*Q55CT0*	*tripeptidyl-peptidase 1*	*Tpp1* *cln2 = similar to Ceroid-Lipofuscinosis, Neuronal 2*	*Tripeptidyl-peptidase 1, TPP-1, EC 3.4.14.9 (Tripeptidyl aminopeptidase) (Tripeptidyl-peptidase I, TPP-I)*	*-0.944*
DDB_G0285637	*Q54MV6*	*beta-galactosidase 2*	*glb2*	*Probable beta-galactosidase 2, Lactase 2, EC 3.2.1.23*	*0.368*
DDB_G0285533	*Q54N35*	*palmitoyl-protein thioesterase 1*	*ppt1*	*Palmitoyl-protein thioesterase 1, EC 3.1.2.22 (Palmitoyl-protein hydrolase 1)*	*0.419*
DDB_G0278653	*Q54YC4*	*alpha-mannosidase*	*manD*	*Alpha-mannosidase D, EC 3.2.1.24*	*0.492*
DDB_G0287659	*Q54K56*	*beta-N-acetylhexosaminidase* *glycoside hydrolase family 20 protein* *beta-hexosaminidase*	*nagD*	*Beta-hexosaminidase subunit B2, EC 3.2.1.52 (Beta-N-acetylhexosaminidase subunit B2) (N-acetyl-beta-glucosaminidase subunit B2)*	*0.53*

		N-acetylglucosaminidase			
DDB_G0290957	*P04988*	*cysteine proteinase 1*	*cprA*	*Cysteine proteinase 1, EC 3.4.22.-*	*0.539*
DDB_G0287231	*Q54KN4*	*alpha-mannosidase*	*manF*	*Alpha-mannosidase F, EC 3.2.1.24*	*0.547*
DDB_G0292206	*P34098*	*alpha-mannosidase A||lysosomal alpha-mannosidase*	*manA*	*Lysosomal alpha-mannosidase, Laman, EC 3.2.1.24 (Alpha-D-mannoside mannohydrolase) (Alpha-mannosidase A) [Cleaved into: Alpha-mannosidase 60 kDa subunit; Alpha-mannosidase 58 kDa subunit]*	*0.604*
DDB_G0290217	*Q54GE1*	*beta-galactosidase 1*	*glb1*	*Beta-galactosidase 1, EC 3.2.1.23 (Acid beta-galactosidase 1, Lactase 1)*	*0.652*
DDB_G0269154	*Q94502*	*alpha-glucosidase II*	*modA*	*Neutral alpha-glucosidase AB, EC 3.2.1.207 (Alpha-glucosidase 2) (Glucosidase II subunit alpha) (Protein post-translational modification mutant A)*	*0.675*

DDB_G0281823	*Q54TD0*	*tripeptidyl-peptidase 1F*	*tpp1F* *v4-7*	*Peptidase S53 domain-containing protein*	*0.662*
DDB_G0279411	*O76856*	*cathepsin D*	*ctsD* *similar to Ceroid-Lipofuscinosis, Neuronal 5*	*Cathepsin D, EC 3.4.23.5 (Ddp44)*	*0.699*
DDB_G0269790	*Q55D50*	*alpha-glucosidase* *alpha-glucoside hydrolase*	*gaa*	*Maltase*	*0.737*
DDB_G0279799	*P04989*	*cysteine proteinase 2*	*cprB*	*Cysteine proteinase 2, EC 3.4.22.- (Prestalk cathepsin)*	*0.752*
DDB_G0282539	*Q54SC9*	*beta-N-acetylhexosaminidase* *glycoside hydrolase family 20 protein*	*nagB*	*Beta-hexosaminidase subunit A2, EC 3.2.1.52 (Beta-N-acetylhexosaminidase subunit A2) (N-acetyl-beta-glucosaminidase subunit A2)*	*0.761*

		beta-hexosaminidase			
		N-acetylglucosaminidase			
DDB_G0285647	*Q54MU9*	*glycoside hydrolase family 20 protein*	*nagE*	*Beta-N-acetylhexosaminidase, EC 3.2.1.52*	*0.804*
		N-acetylglucosaminidase			
DDB_G0283033	*Q54RM4*	*tripeptidyl-peptidase 1E*	*tpp1E*	*Peptidase S53 domain-containing protein*	*0.914*
DDB_G0278259	*Q54YF7*	*alpha-mannosidase*	*manB*	*Alpha-mannosidase B, EC 3.2.1.24*	*1.026*
DDB_G0287577	*Q54K67*	*alpha-mannosidase*	*manG*	*Alpha-mannosidase G, EC 3.2.1.24*	*1.059*

2.3.4 Autophagy assay

During multicellular development amoebae rely on autophagy to provide cells with energy for aggregation and fruiting body generation (Mesquita et al., 2016). Pervious experimentation has shown that there are linkages between NCL diseases and disruptions in autophagy (Brandenstein et al, 2015). By examining the comparative transcriptomics dataset, several autophagy related genes were observed to be

differentially expressed (Table 2.6) including decreases in both *atg4* and *atg8* which is indicative of dysfunctions in autophagy. Thus, I measured autophagic activity in *mfsd8*$^{-}$ cells (Figure 2.4). After 8 hours of starvation, it was observed that there was a decrease in autophagic activity of 20±8% in *mfsd8*$^{-}$ cells compared to WT cells. This result shows that loss of *mfsd8* alters the normal functioning of the autophagic pathways during the early stages of multicellular development.

Table 2.4. List of genes associated with autophagy observed in 8-hour starvation transcriptomic dataset

Gene ID	**UniProt ID**	**Gene name**	**Gene**	**Protein name**	**log2 Fold Change**
DDB_G0290491	*Q54G11*	*autophagy protein 8b*	*atg8b*	*Autophagy-related protein 8-like protein DDB_G0290491*	*-0.95*
DDB_G0286191	*Q86CR8*	*autophagy protein 8a*	*atg8a*	*Autophagy-related protein 8 (Autophagy-related ubiquitin-like modifier atg8)*	*-0.548*
DDB_G0273553	*Q557H7*	*autophagy protein 4*	*atg4-2*	*Cysteine protease atg4, EC 3.4.22.- (Autophagy-related protein 4)*	*-0.451*
DDB_G0273443	*Q557H7*	*autophagy protein 4*	*atg4-1*	*Cysteine protease atg4, EC 3.4.22.- (Autophagy-related protein 4)*	*-0.389*
DDB_G0269162	*Q55BY0*	*autophagy protein 13*	*atg13*	*ATG13 domain-containing protein*	*0.352*
DDB_G0295673	*B0G140*	*autophagy -related protein 17*	*atg17*	*Probable autophagy-related protein 17*	*0.359*
DDB_G0289881	*Q54GT9*	*autophagy protein 5*	*atg5*	*Autophagy protein 5*	*0.496*
DDB_G0285323	*Q54NA3*	*autophagy protein 9*	*atg9*	*Autophagy-related protein 9*	*0.603*
DDB_G0271096	*Q86CR9*	*autophagy protein 7*	*atg7*	*Ubiquitin-like modifier-activating enzyme atg7 (ATG12-activating enzyme E1 atg7) (Autophagy-related protein 7)*	*0.825*

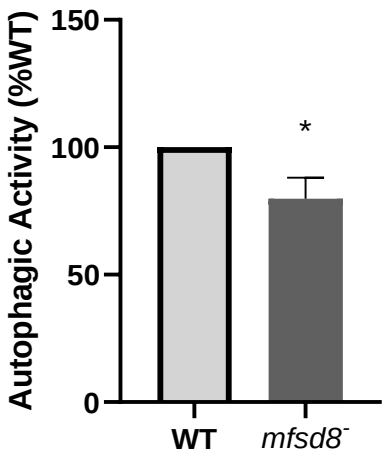

Figure 2.4. Effect of *mfsd8* loss on autophagic activity after 8 hours of starvation. Loss of *mfsd8*$^{-}$ affects the autophagy pathway during starvation. WT and *mfsd8*-deficient cells during 8-hour KK2 starvation were lysed and autophagic activity was quantified as described in the Materials and Methods. All autophagic activity values were standardized against those of the WT cells Statistical significance was determined using the one-sample t-test. Data obtained from plate reader represented as mean autophagic activity relative to WT +/- SEM (n=7 Biological replicates) *p<0.05.

2.3.5 RT-PCR

Our work produced a list of differentially expressed genes to further support the results of the transcriptomic dataset and to gain further insight of the transcriptomic differences between WT and *mfsd8*$^{-}$cells, RT-PCR was employed (Figure 2.5). During growth *mbtps1*, was examined while during multicellular development, *sibC* was examined. During growth the loss of *mfsd8* resulted in a non-significant increase in *mbtps1*. There was a significant increase in the level of *sibC observed during the early stages of development.* The results for our RT-PCR specifically the analysis of *sibC,*

supports the finding of our Comparative transcriptomics as the result of RT-PCR are consistent with the finding of the comparative transcriptomic analysis.

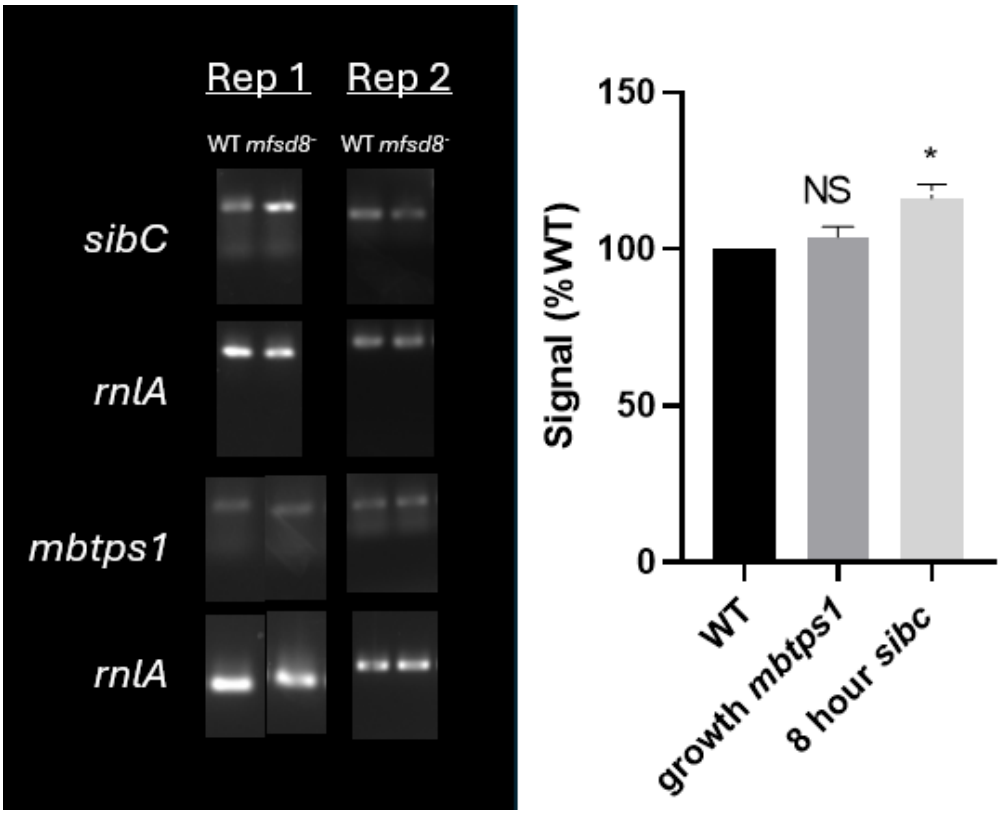

Figure 2.5. RT-PCR to corroborate the differentially expressed genes in *mfsd8*$^{-}$ cells during growth and after 8 hours of starvation. Gel electrophoresis was carried out on prepared RT-PCR samples and images gels were obtained. Fiji/ImageJ was used to capture signal of the resulting bands. Statistical significance was determined using the one-sample *t*-test. Obtained signal data presented as mean signal relative to WT +/- SEM (*mbtps1* and *sibc* n=3 biological replicates) *p<0.05

2.4 Discussion

2.4.1 Overview of Effect of *mfsd8*-deficiency on protein Transcriptome during the early stages of *D. discoideum*

The deletion of *mfsd8* in *D. discoideum* causes the appearance of several phenotypic differences from WT cells. During growth, there is aberrant cytokinesis, an increase in both cell size and proliferation, while in early multicellular development there is a delay in aggregation and aberrant protein secretion. In this study, comparative transcriptomics was used to gain insight regarding the molecular mechanisms underlying the observed phenotypical differences. The resulting expression data was then used to generate follow up work further examining transcriptomic differences, enzyme activity, autophagic activity.

This work created a catalogue of genes differentially expressed during growth and early multicellular development. There were 445 and 4954 genes aberrantly expressed during growth and the early stages of multicellular development, respectively. For starvation, 47% of the genes observed after the filtration were down regulated while 54% of differentially expressed genes were down regulated during growth. 60% of the genes observed within the growth dataset were also observed in the starvation dataset. The expression of several upregulated and downregulated genes were analyzed by RT-PCR. A multitude of genes observed to be differentially expressed fit into the previously observed phenotypes produced by *mfsd8* loss in *D. discoideum* (Yap et al., 2022).

2.4.2 *mfsd8*$^{-}$ transcriptome reveals possible reasons for observed phenotypes in *D. discoideum*

Our comparative transcriptomics analysis revealed genes associated with aggregation, chemotaxis, and cell substrate adhesion in *mfsd8*$^{-}$ cells which is consistent with the observed phenotypes seen in previous work (Yap et al., 2022). Regarding cell substrate adhesion, our work has revealed several proteins differentially expressed including vasP (down) *ampa* (up) *paxb* (up) *dscE* (up) (Han et al., 2002, Noratel et al., 2012, Duran et al., 2009, Springer et al.,1984). All these proteins have roles in normal cell substrate adhesion, the overexpression of paxb and decreases in expression of *vasP* have been directly shown to decrease cell substrate adhesion. There are several lysosomal enzymes that through their secretion, negatively impact cell substrate adhesion by hydrolysing surface glycoproteins such as alpha-mannosidase and N-Acetyl glucosaminidase (Loomis et al., 2012; Tarantola et al., 2014). Previous research has shown that intracellular and extracellular alpha-mannosidase and N-Acetyl glucosaminidase activity are increased by the loss of *mfsd8* (Yap 2021). Our findings mirror this on the RNA level, as alpha-mannosidase and several known N-Acetyl glucosaminidases (*nagB, nagD and nagE*) all had increased expression in our *mfsd8* knockout cell line compared to WT. Countin, a known regulator of aggregate size was previously shown to be affected by the deletion of *mfsd8*. Huber *(2022)* found that there was a decrease in intracellular CtnA and an increase in extracellular CtnA. In this study we found that *ctnA, ctnB* and the associated gene *cfaA* to all have decreased expression.

D discoideum suffers a significant delay in aggregation due to the loss of m*fsd8* and the process of aggregation is governed by cAMP-mediated chemotaxis. Utilizing aggregation assays, previous studies have found that cAMP-mediated chemotaxis is not significantly impacted by the loss of *mfsd8* (Yap et al., 2022) however, our Differential gene analysis observed several differentially expressed genes were associated with cAMP mediated chemotaxis. c*mfA and cmfB* were both observed to have increased expression levels after the loss of *mfsd8*. The proteins they produce CmfA and CmfB function together to manipulate the process of cAMP mediated chemotaxis and the proper expression of several genes necessary for proper differentiation during early multicellular development. These shifts in gene expression of *cmfA* and *cmfB* and delayed aggregation are similar to those observed in knockout models of CLN5 disease in *D. discoideum* (*cln5*$^{-}$*)* (Kim and Huber, 2022) as a delay in aggregation and increases in *cmfA* and *cmfB* expression were also observed. These findings further support a theory of a network for the NCL genes (Huber 2020).

Several other genes related to *cAMP* mediated chemotaxis are impacted by the loss of *mfsd8.* The Phosphodiesterase Inhibitor, *pdiA,* was identified to be overexpressed in our *mfsd8 knockout line.* This would imply that there could be aberrant decreased activity of *pdsA the* extracellular cAMP phosphodiesterase that *pdiA* inhibits. The differential gene expression analysis did not observe any shift in the expression of *pdsA.* Experiments done on *pdiA* overexpression lines have shown delays in aggregation (Wu et al 1995) while null mutations have shown aberrant oscillatory cAMP signaling. The implications of increased activity of *pdsA* could be theorised by studies observing *pdsA*$^{-/+}$

mutants. In null strains there are observed abolishment of aggregation, aberrant chemotaxis to cAMP, aberrant sorocarp development and aberrant stream morphology (Sucgang et al., 1997). Overexpression lines result in precocious development, aberrant sarcocarp development and multicellular development halting during the mound stage (Hall et al., 1993).

Several genes relating to Scar (Suppressor of Cyclic AMP Receptor) and Wasp (Wiskott-Aldrich Syndrome Protein) proteins, genes crucial to actin-facilitated pseudopod-based motility, also had aberrant gene expression (Litschko et al., 2017, Myers et al., 2005). Proper actin dynamics are required for the proper aggregation, *scrA* (Suppressor of cAMP Receptor mutation) which stimulate actin polymerization, *abiA*, a component of the scar complex, and *wasB* (Chung et al., 2016) which regulates both f-actin polymerization and regulates pseudopod formation were all shown to have a significant decrease in expression. Null mutation *scrA* and *abiA* decreased cellular motility (Pollitt & Insall 2008). Actins 12 and 21 had an increased expression while actins 8, 26-29 along with actin-related protein 2/3 complex, subunit 5 and actin related protein (arp) 5 all had reduced expression. The known Mfsd8 protein interactors *aprB aprC mhcA* all had modified expression levels and *cln3,* which has a known role in regulating the actin cytoskeleton through its interactions with *myosin-II-b,* was also observed to have increased expression levels (Getty et al., 2011, Cotman & Lefrancois 2021). Together these findings support a possible perturbation in actin dynamics as a result of the loss of *mfsd8* which could be a possible reason for delays in aggregation.

2.4.3 *mfsd8*$^{-}$ transcriptome reveals possible reasons for perturbations in autophagy

Perturbations in the normal functioning of the autophagic pathways have been observed in several subtypes of NCL (Kim et al., 2022) including CLN7 disease. Several DEGs in the early multicellular development dataset were related to autophagy, these include atg4 1 (down), atg4-2 (down), atg5 (up), atg7 (up), atg8a (down), atg8b (up), atg9 (up), atg13 (up), atg17 (up) and fip200 (up) pkbA (down), pia1 (up) and piks B, D, E, F, and H (down). As this was the case, the autophagic activity of *mfsd8*$^{-}$ cells were assessed. It was found that there was significant decrease in autophagic activity. This impairment of autophagic function is consistent finding in mice models (Brandenstein et al., 2016). One possible reason for the reduction in autophagic activity is the lowered expression of both atg4s and atg8a. The reduction of the protein products of the two genes would result in the perturbation of autophagy as there would be lowered amounts of LC3 due to lowered Atg8a and lowered LC3-1 as Atg4 is reduced (Kim et al., 2022). Intriguingly there is the reduction in pkbA and increases in atg 13 and fip200 imply an increase in effects on the ULK1 complex (Kim et al., 2022). Together these results show that autophagy is negatively impacted by the loss of *mfsd8,* but it reveals a possible mechanism for why this occurs which need to be examined.

Autophagy is dependent on the functioning of lysosomal enzymes as the autophagosome needs to fuse with the lysosome for proper autophagy activity (Brandenstein et al., 2015). Previous work showed that 58% of known Mfsd8-interactors observed during early multicellular development have catalytic activity (Yap et al., 2020). Studies using *mfsd8*-deficient mouse embryonic fibroblasts also showed

alterations in the amounts of lysosomal proteins (Danyukova et al., 2018). As this is the case, our examinations in the early multicellular development RNA-sequencing dataset unsurprisingly found several DEGs that encode lysosomal enzymes including β-glucosidase and α-glucosidase and β-galactosidase *nagB,E and D, manA-D F, G cprA and B tpp1E-D ctsD.* These finding were followed up by examining activities of some of the observed differentially expressed genes in early multicellular development. The three observed proteins, alpha glucosidase, beta-galactosidase, and ctsd showed increases in their activity in response to the loss of *mfsd8* which provides support for the assertation that *mfsd8 plays* a role in the regulation of these lysosomal enzymes. Previous work also explained the activities of two of the three enzymes β-Galactosidase and Ctsd which found non-significant increases in the activities of both enzymes (Yap et al., 2022).

2.4.4 GO terms relate to 8 hour observed phenotypes

Go term enrichment analyses identified differentially expressed genes associated with autophagy, catabolic processes and metabolic processes. This is consistent with previous studies examining Mfsd8 localization in both *D. discoideum* and mammals, as GO term analyses found enrichment of differentially expressed genes coding for proteins localized to the endocytic compartments (Yap et al., 2022, Huber et al 2020).

2.4.5 Conclusion

To conclude, this study not only provides insight on the mechanisms underlying the phenotypic differences between WT and our *mfsd8* knockout line, but it also

catalogues the differential gene expression seen. This both supports older finding and presents a scaffold in which new techniques, methods and examinations can be generated for *D. discoideum* and mammalian models of CLN7 disease. Our examination of lysosomal proteins delays in aggregation and protein secretion unveiled a continuing link between the genes and proteins affected by CLN7 disease and those affected in other CNL diseases further cementing that CLN proteins function in shared or convergent biological pathways. Further work examining protein level and activity changes in the products of the genes elucidated in the project will allow for greater understanding of the mechanisms underlying CLN7 disease, but it also would provide new avenues for therapeutic interventions for CLN7 disease in humans.

Chapter 3- Effect of *mfsd8*-deficiency on protein secretion during the early stages of *D. discoideum* development

The data contained within this chapter were previously published in the *European Journal of Cell Biology*: Huber RJ, Gray J, Kim WD. Loss of *mfsd8* alters the secretome during *Dictyostelium* aggregation. Eur J Cell Biol. 2023;102(4):151361. For this paper, I collected and analyzed all the wet lab data and revised the manuscript prior to submission. Robert Huber analyzed the mass spectrometry data and wrote the first draft of the paper. William Kim **assisted in the experimental design and the revision of the manuscript prior to submission.**

3.1 Introduction

Current research into the protein major facilitator superfamily domain containing 8 (MFSD8), also known as Ceroid Lipofuscinosis neuronal 7 (CLN7), has identified its function as a lysosomal transmembrane chloride channel (Siintola et al., 2007; Sharifi et al., 2010; Steenhuis et al., 2010; Steenhuis et al., 2012; Wang et al., 2021). It is predominantly found in the endocytic pathway including cytoplasmic puncta and vesicles, macropinosomes, acidic vesicles, and secretory lysosomes where it allows chloride ions to cross membranes (Rot et al., 2009; Journet et al., 2012; Huber et al., 2020). The *MFSD8* gene is present within all eukaryotes and has been shown to play roles in autophagy (Aiello et al., 2009). In humans, mutations in *MFSD8* result in the late infantile onset form of neuronal ceroid lipofuscinosis (NCL) or CLN7 disease. CLN7 disease is one of the 13 different subtypes of Batten disease. Each subtype is caused by

mutations in one of the 13 ceroid lipofuscinosis neuronal (*CLN*) genes (*PPT1/CLN1, TPP1/CLN2, CLN3, DNAJC5/CLN4, CLN5, CLN6, MFSD8/CLN7, CLN8, CTSD/CLN10, GRN/CLN11, ATP13A2/CLN12, CTSF/CLN13, KCTD7/CLN14*) (Butz et al., 2020). The clinical symptoms of the disease include vision loss, seizures, loss of motor/cognitive function and a reduction in lifespan. It is not currently understood how MFSD8 dysfunction leads to NCL (Schulz et al., 2013). The protein has been linked to an assortment of processes in mammals including autophagy, neuronal cell survival, and regulation of the lysosome including pH size and trafficking (Danyukova et al., 2018; Brandenstein et al., 2016; von Kleist et al., 2019; Wang et al., 2021).

The biomedical model used in this study is the social amoeba *Dictyostelium discoideum*. *D. discoideum* has been used to examine conserved cellular and developmental functions of several of the CLN proteins including Mfsd8 (Huber et al., 2020; Remtulla and Huber, 2023). *D. discoideum* undergoes a series of morphological changes during nutrient poor conditions from a haploid single celled growth phase to a multicellular fruiting body supporting masses of viable spores (Mathavarajah et al., 2017). Due to its life cycle, *D. discoideum* is used to examine fundamental cellular and developmental processes.

Recent work examining the *D. discoideum* homolog of MFSD8, Mfsd8, has revealed that Mfsd8 has a role in regulating conserved processes during growth and early multicellular development such as proliferation, pinocytosis, cytokinesis, lysosomal enzyme activity, cell-substrate adhesion and aggregation. It was also found that the loss of *mfsd8* modulates the extracellular amounts of Cln5, the homolog for human CLN5,

cathepsin D (CtsD), the homolog of human CTSD, and countin, a protein that regulates aggregation (Huber, 2017; Huber, 2020; Huber, 2021; Yap et al., 2022). Altered protein secretion has been linked to several types of NCL in *D. discoideum*. Loss of *cln3* has also been shown to affect the secretome during aggregation. In regard to MFSD8, beta-hexosaminidase activity has been reported in conditioned buffer from cerebellar granule neuron precursors isolated from *Mfsd8*-deficient mice (von Kleist et al., 2019) and decreases in tripeptidyl peptidase 1 amounts have been observed in urine from human patients with CLN7 Disease (Iwan et al., 2020).

This study uses *D. discoideum* and mass spectrometry to elucidate the role of Mfsd8 in the regulation of protein release by analyzing the secretome of *mfsd8*-deficient cells during the early stages of multicellular development. It is known that during the early stages of multicellular development, *D. discoideum* secretes proteins into its environment to facilitate cAMP-mediated chemotaxis and aggregation (Rossomando et al., 1978; Dimond et al., 1981; Gomer et al., 1991). As such, this study will show the effects of *mfsd8* loss on the secretome, which will provide the foundation for work in *D. discoideum* and mammalian models to uncover the role of altered protein secretion in CLN7 disease.

3.2 Materials and Methods

3.2.1 Cell lines

AX4 (DBS0237637, hereafter referred to as wild type, WT) and *hnm*$^{-}$ (GWDI_448_B_4) cell lines were purchased from the Genome Wide *Dictyostelium*

Insertion (GWDI) bank via the Dicty Stock Center (https://remi-seq.org) (Fey et al., 2019, Gruenheit et al., 2021). Cell lines were maintained at 22˚C on SM/2 plates and *D. discoideum* fruiting bodies were grown alongside Klebsiella aerogenes (Fey et al., 2007, Fey et al., 2019). Prior to experimentation, cells were grown axenically in HL5 medium purchased from Formedium (Hunstanton, Norfolk, United Kingdom). HL5 medium was supplemented with Ampicillin 100ug/ml and streptomycin sulfate 300ug/ml to *prevent* bacterial growth (Bioshop Canada Incorporated, Burlington, Ontario, Canada). The growth media for *mfsd8*$^{-}$ cells carrying a blasticidin S resistance (bsr) cassette was additionally supplemented by blasticidin S hydrochloride (10 μg/ml) as a selective (Bioshop Canada Incorporated, Burlington, Ontario, Canada). Before experimentation, media was changed to HL5 medium supplemented with Ampicillin 100ug/ml and streptomycin sulfate 300ug/ml and 2 doubling times were allowed to take place. Starvation assays were conducted using KK2 buffer (2.2 g/L KH2PO4 and 0.7 g/L K2HPO4, pH 6.5).

3.2.2 Antibodies

Rabbit polyclonal anti-conditioned medium factor A (CmfA) (Jain et al., 1992) and rabbit polyclonal anti-counting factor-associated protein D (CfaD) (Bakthavatsalam et al., 2008) were gifts from Dr. Richard Gomer (Texas A&M University, Texas, United States). Mouse monoclonal anti-calreticulin (252−234−2) (Müller-Taubenberger et al., 2001), mouse monoclonal anti-fimbrin (210−183−1) (Prassler et al., 1997), and mouse monoclonal anti-α-actinin (47−18−9) (Schleicher et al., 1988) mouse monoclonal anti-β-

actin (224−236−1) (Westphal et al., 1997), mouse monoclonal anti-α-tubulin (12G10) (Jerka-Dziadosz et al., 1995), were purchased from the Developmental Studies Hybridoma Bank (University of Iowa, Iowa City, Iowa, USA). Horse anti-mouse IgG HRP-linked antibody (7076) and goat anti-rabbit IgG HRP-linked antibody (7074) were purchased from New England Biolabs (Whitby, Ontario, Canada)

3.2.3 Starvation assay

WT and *mfsd8*$^{-}$ cells in the mid-log phase of growth (1-5x10^{6} cells /ml) were deposited into 100mmx15mm petri dishes containing HL5 (Fisher Scientific Company, Ottawa, Ontario, Canada) and allowed to adhere to the dishes for 1h after which dishes where examined for confluence. Adherent cells were then washed two times and starved in fresh KK2 buffer for 4- and 8-hours. Cells were imaged immediately after induction of starvation and at the two time points using a Nikon Ts2R-FL inverted microscope equipped with a Nikon 10 Digital Sight Qi2 monochrome camera (Nikon Canada Incorporated Instruments Division, Mississauga, Ontario, Canada). After 4 and 8 hours of starvation, the condition buffer was removed from the dishes and cells were lysed with buffer containing 0.5% NP-40 (v/v), 50 mM Tris–HCl (pH 8.0), 150 mM NaCl, and a protease inhibitor tablet (PIA32965) (Fisher Scientific Company, Ottawa, Ontario, Canada) whole cell lysate was added to a 1.5ml tube and stored at -80 for future use. The conditioned buffer was centrifuged for 10 minutes at 2500rpm and 4C to remove free floating cells. The resulting supernatant was then centrifuged for 25 minutes at 4200 rpm at4°C in an Amicon Ultra-4 centrifugal filter unit (UFC801024) (Fisher Scientific

Company, Ottawa, Ontario, Canada). The concentrated conditioned buffer was then added to 1.5ml microfuge tubes and stored at -80C for future usage.

3.2.4 Cell viability assay

The viability of the cell lines was examined using CellTiter-Glo 2.0 Cell Viability Assay following the manufacturer's guidelines (Promega, Madison, WI, USA). WT and AX4 cells in the mid log phase of growth were washed and resuspended in KK2 Buffer at 4×10^6 cells/ml. The resuspended cells were added to flasks and incubated at 22˚C on an orbital shaker rotor set to 150rpm. After 4 and 8 hours of incubation $5x10^4$ cells were harvested from each flask and added to 1.5ml tubes. Each tube had its volume adjusted with KK2 buffer to achieve a total volume of 50ul. 50 µl of CellTiter-Glo 2.0 Cell Viability Assay working solution was added to each tube. Tubes were then incubated for 10 minutes at 22◦C. 50 µl of sample was then added to separate wells of an opaque bottom white 96 well plate. Luminescence values were captured using a BioTek Synergy HTX microplate reader (BioTek Instruments Incorporated, Winooski, Vermont, USA). For each biological replicate, raw luminescence values from *mfsd8*$^-$ samples were standardized against the values obtained for WT.

3.2.5 Mass spectrometry

Samples of concentrated conditioned buffer from four independent experiments were submitted to SPARC Biocentre at the Hospital for Sick Children (Toronto, Ontario, Canada) for analysis. The samples were reduced through the use of 20 mM DTT for 10

min at 95◦C and alkylated with 40 mM iodoacetamide for 30 min in the dark at 22◦C. 2ug of trypsin (Fisher Scientific Company, Ottawa, Ontario, Canada) was used to digest samples for 2 h at 47◦C in 50 mM ammonium bicarbonate (pH 8.0) and then eluted with 50 mM ammonium bicarbonate (pH 8.0), 0.2% formic acid, 50% acetonitrile/0.2% formic acid, and 80% acetonitrile/0.2% formic acid. 2 h at 47◦C in 50 mM ammonium bicarbonate (pH 8.0) and eluted with 50 mM ammonium bicarbonate (pH 8.0), 0.2% formic acid, 50% acetonitrile/0.2% formic acid, and 80% acetonitrile/0.2% formic acid. Samples were then lyophilized with a Speedvac and resuspended in 2% acn/0.1% formic acid. Samples were analyzed using a Thermo Scientific Exploris 480 mass spectrometer (Thermo Fisher Scientific, Whitby, Ontario, Canada) and the EASY-nLC 1000 nano-LC system (75 µm x 50 cm PepMax RSLC EASY-Spray column filled with 2 µM C18 beads, pressure 900 Bar, 60 ◦ 2◦C, Buffer A: 0.1% formic acid (v/v), Buffer B: 80% acetonitrile, 0.1% formic acid (v/v)). Identification and quantification of proteins were made using Proteome Discoverer version 2.5.0.400 and the uniport 3AUP000002195_Dictyostelium_discoideum_08022022.fasta database. The following search parameters were used: parent mass error tolerance: 50 ppm, fragment mass error tolerance: 0.02 Da, enzyme: trypsin, max missed cleavages: 3, fixed modifications: carbamidomethylation: C (+57.02), variable modifications: oxidation: M (+15.99), deamidation: N, Q (+ 0.98), acetylation: peptide N-term (+42.01). The program Scaffold (V 5.1.2) (Proteome Software Incorporated, Portland, Oregon, USA) was used to observe peptide and protein identities which were accepted if they, established >95% probability and contained 2 or more peptides for peptide and protein identity respectively. Scaffold

Q+ was used to quantify normalized total precursor intensity data and Statistical significance was examined using the Mann-Witney test with a <0.05 p-value. All data regarding the mass spectrometry proteomics have been added to the ProteomeXchange Consortium using the PRIDE partnership repository (Perez-Riverol et al., 2022) which can be found using the dataset identifier PXD042734.

3.2.6 Gene ontology (GO) term enrichment analyses

LAGO was used for all GO term enrichment analyses (biological processes, molecular function and cellular components) (A Logically Accelerated GO Term Finder, https://go.princeton. edu/cgi-bin/LAGO) (Boyle et al., 2004). Uncharacterized genes, those annotated but without associated GO terms, were removed from the dataset during these analyses. At 4 hours, 51, 50 and 50 proteins were associated with biological processes, molecular function and cellular components respectively. At 8 hours, 53, 54, and 53 proteins were used for biological processes, molecular function and cellular components respectively. Statistical Significance was a p-value <0.05 and the Bonferroni correction was applied.

3.2.7 SDS-PAGE and western blotting

A Qubit 2.0 fluorometer (Fisher Scientific Company, Ottawa, Ontario, Canada) was used to quantify the protein concentrations of Whole cell lysate and conditioned buffer. An equalized amount of protein 10 ug for whole cell lysate and 0.4ug for conditioned buffer was used. Samples were then separated using SDS-PAGE and then

observed using Western blotting by standard methods. Membranes were incubated with primary antibodies in 5% (w/v) milk/TBST for 2 h at 22 °C for whole cell lysate and overnight at 4°C. A secondary antibody linked to horseradish peroxidase (HRP) in a 5% (w/v) milk/TBST solution was then used for 1 hour. The primary antibodies were used; anti-calreticulin (1:500), anti-fimbrin (1:500), anti-α-actinin (1:500), anti-CfaD (1:1000), anti-CmfA (1:1000), anti-β-actin (1:2000), anti-α-tubulin (1:2000). The secondary antibodies were anti-mouse IgG-HRP (1:1000) and anti-rabbit IgG-HRP (1:1000). The detected protein banding was consistent with previous studies (Müller-Taubenberger et al., 2001; Prassler et al., 1997; Schleicher et al., 1988; Westphal et al.,1997; Jerka-Dziadosz et al., 1995). Immunoblots were imaged using a ChemiDoc Imaging System (Bio-Rad Laboratories Canada, Mississauga, Ontario, Canada) and quantified using fiji and ImageJ (Schindelin et al., 2012). Band values were obtained by subtracting background from the band of image. The band values were than standardized against α-actinin and the values of *mfsd8*$^{-}$values were further normalized to WT for whole cell samples and conditioned buffer samples. Statistical analysis was conducted using the one-sample t-test in which a p-value of < 0.05 was considered significant.

3.2.8 Proteasome activity assay

4 hour and 8 hour starved WT and *mfsd8*$^{-}$ cells were lysed using a 0.1% NP40 in 0.05 M 2-(N-morpholino) ethanesulfonicacid (MES) (pH 6.5) buffer. A Proteasome 20S Activity Assay Kit (MAK172) (Sigma Aldrich Canada, Oakville, Ontario, Canada) was then used to measure proteasome 20S activity following the manufacturer's instructions. A

Qubit 2.0 fluorometer (Fisher Scientific Company, Ottawa, Ontario, Canada) was used to quantify the protein concentrations of Whole cell lysate and an equal 25ug of protein was used. All assays were performed in duplicate and measured using a Synergy HTX multi-mode plate reader (Agilent Technologies Canada, Mississauga, Ontario, Canada) (485/ 20 nm excitation, 528/20 nm emission). All activity values were corrected for using a blank and *mfsd8*$^{-}$ activity was standardized by the activity of the WT. a one sample t-test was used for statistical analysis and p-values < 0.05 were considered significant

3.3 Results

3.3.1 The effect of *mfsd8*-deficiency on the secretome during starvation

WT and *mfsd8*$^{-}$ cells were grown in HL5 and starved for 4 and 8 hours in KK2 buffer. Whole cell lysates and samples of conditioned buffer were then harvested from these cells. *mfsd8*$^{-}$ cells had delayed aggregation when compared to WT cells (Figure 3.1A). This aligns with previous research (Yap et al., 2022). *mfsd8* loss had no statistically significant effect on cell viability during the time points analyzed (Figure 3.1B).

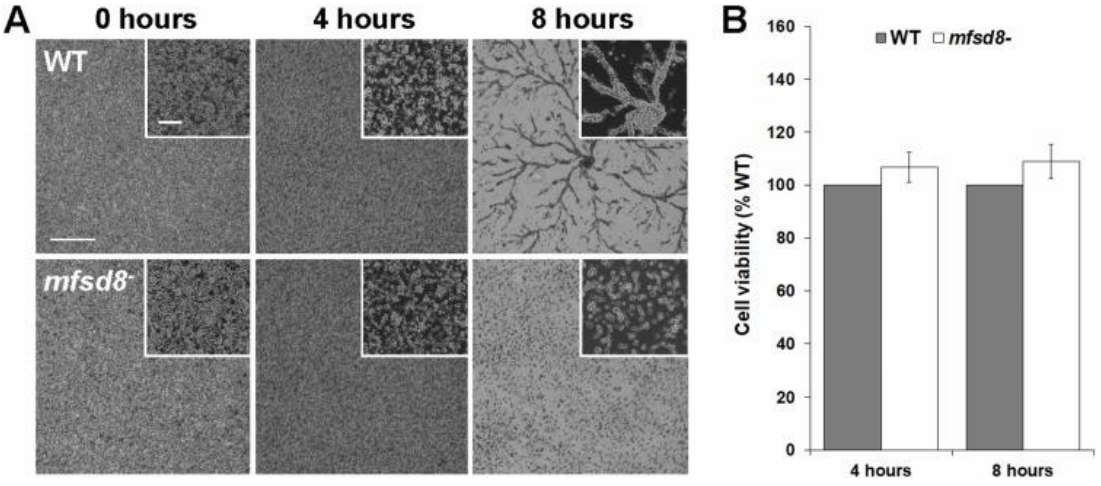

Figure 3.1. A) Representative images of starvation of WT and *mfsd8*- cells Scale bar = 500 µm (inset 100 µm). B) The effect of *mfsd8*-deficiency on cell viability. The viability of WT and *mfsd8*- cells after 4 and 8 h of starvation was assessed using the CellTiter-Glo 2.0. Luminescence data was captured and raw luminescence values from *mfsd8*$^{-}$ samples were standardized against the values obtained for WT. Data presented as mean cell viability (% WT) ± SEM (n = 6). Figure taken from Huber et al (2023).

The protein within the conditioned buffer of 4 biological replicates of both 4 and 8 hours, were examined through the use of liquid chromatography with tandem mass spectrometry (LC-MS/MS). After 4 hours of starvation, an average of 355 ± 17 (306–381) proteins were seen in WT (Table 3.1), while in *mfsd8*$^{-}$ 348 ± 20 proteins (309–404) were detected (based on total spectrum count, 2 identified peptides, and peptide/protein identifications established at > 95% probability). Regarding the replicates starved for 8 hours, 476 ± 55 (319–577) proteins were identified in WT conditioned buffer while 507 ± 54 (350–598) proteins were found within *mfsd8*$^{-}$ samples (Table 3.2). Total number of proteins secreted into conditioned buffer was consistent with the total number of proteins released in prior studies using AX2 (349) and AX3 (435), two other wild type lines regularly used for *D. discoideum* research (Bakthavatsalam and Gomer, 2010; Huber, 2017).

mfsd8-deficiency resulted in differential protein secretion for 62 proteins between the 4- and 8-hour starvation time points. A majority of the differential secretion

showed increased secretion, at 4 hours 87% of proteins were observed as increased as 87% and at 8 hours 74% of the proteins with altered secretion were increased. Of the proteins with altered secretion, seven were shared between the 4- and 8-hour time periods.

Western blotting was employed to provide additional support for the proteins detected by mass spectrometry. Through the use of commercially available antibodies, an increase of extracellular calreticulin, a previously known endoplasmic reticulum chaperone after 4 hours (Figure 3.2 A) and fimbrin, a calcium-regulated actin bundling protein after 8 h of starvation was confirmed (Figure 3.2 B) (Muller-Taubenberger, 2001; Prassler et al., 1997). This mirrors the mass spectrometry results. The western results are supported by the fact that the intracellular amounts of calreticulin and fimbrin were not altered and there was no effect on counting factor-associated protein D (CfaD) and conditioned medium factor A (CmfA), two known secreted proteins (Jain et al., 1992; Bakthavatsalam et al., 2008). The observed alteration of secretion was shown not to be a result of cell lysis as markers for cell lysis (β-actin, α-tubulin, α-actinin) were probed and observed to be unaltered.

Previous work in *D. discoideum* elucidated the secretome of AX2 and AX3 cells (Bakthavatsalam and Gomer, 2010). Our work shows that in the 4-hour starved condition buffer 56% (34/61) of the proteins aberrantly released by *mfsd8*$^{-}$ cells were not released by AX2 cells and 33% (20/61) of the proteins identified in altered amounts

in *mfsd8*$^{-}$ conditioned buffer were not released by AX3 cells. After 8 hours of starvation. 66% (40/61) of proteins aberrantly released by *mfsd8*$^{-}$ *cells* were not present in the AX2 secretome. 75% (12/16) proteins present in reduced amounts in *mfsd8*$^{-}$ *conditioned* buffer were identified in the AX2 secretome. 49% (30/61) of proteins affected by the loss of *mfsd8* were not detected in the AX3 secretome. As seen in AX2 a majority of the proteins released in lower amounts were seen in the AX3 secretome 63%. This indicates that due to the loss of *mfsd8* there is a release of proteins not normally seen within the wild type during early multicellular development.

Table 3.1. Summary of proteins aberrantly released by *mfsd8*$^{-}$ cells after 4 and 8 hours of starvation determined though mass spectrometry

Time of starvation	Extracellular amount	Number of proteins	% total
4 h	Decreased	7	11
	Increased	54	89
8 h	Decreased	16	26
	Increased	45	74

Table 3.2. List of common proteins aberrantly released by *mfsd8*$^{-}$ cells after 4 and 8 hours of starvation

Uniprot ID	dictyBase gene ID	dictyBase protein ID	Protein name	Gene name	Time	p-value*	Log2 FC**
B0G141	DDB_G0282979	DDB0234059	Glutathione S-transferase domain-containing protein	*efa1G*	4 h	0.0051	0.9
					8 h	0.011	1.36
Q54X73	DDB_G0279159	DDB0229908	Probable cytoplasmic aconitate hydratase	*aco1*	4 h	0.05	0.69
					8 h	0.04	0.55
Q869W9	DDB_G0275069	DDB0230068	Probable polyketide synthase 16	*pks16*	4 h	0.043	0.53
					8 h	0.011	0.84
P54681	DDB_G0271916	DDB0185120	Protein rtoA	*rtoA*	4 h	0.0045	2.04
					8 h	0.021	1.43
Q8T2T8	DDB_G0275111	DDB0307363	Uncharacterized protein (contains a predicted signal peptide; conserved in *Dictyostelium* and *Polysphondylium*)	*DDB_G0275111*	4 h	0.04	1.18
					8 h	0.0062	1.28
Q86KC7	DDB_G0274171	DDB0230011	Uncharacterized protein DDB_G0274171 (similar to Fraser syndrome protein, von Willebrand factor, and kielin; expressed in pstAO cells and in upper cup during culmination)	*stcC*	4 h	0.014	0.95
					8 h	0.001	1.14
Q54C24	DDB_G0293218	DDB0234193	Vacuolar protein sorting-associated protein 35	*vps35*	4 h	0.05	1.31
					8 h	0.025	1.38

*Mann-Whitney test
**Log2 FC (Mfsd8/WT)

3.3.2 Analysis of proteins aberrantly released by *mfsd8*$^{-}$ cells during starvation

In eukaryotes proteins can be released via conventional and unconventional pathways. The conventional pathway shuttles proteins from the endoplasmic reticulum to the Golgi apparatus in which proteins can be packaged for extracellular release in secretory vesicles (Viotti, 2016). Signal peptides are required for most proteins that are secreted by the conventional pathways. SignalP 6.0, which scans proteins for the presence of one of the five known types of signals peptides, was used to examine the proteins apparently released by *mfsd8*$^{-}$ cells (Teufel et al., 2022). During 4 hours of starvation 40/61 proteins had signal peptides while after 8 hours 22/61 proteins had signal peptides for secretion. All 16 proteins detected in a reduced amount during 8

hours of starvation had signaling peptides while 6/39 proteins observed as increased at 8 hours had signaling peptides. Theses results suggest that the conventional pathway and unconventional pathway of secretion are perturbed by the loss of *mfsd8,* and cells are secreting proteins not normally secreted.

Previous work elucidated 119 direct and indirect Mfsd8 interactors after 4 hours of starvation (Huber et al., 2020). Six of the known Mfsd8 interactors had abnormal secretion in *mfsd8*$^{-}$ cells after 4 hours of starvation including CtsD, which was previously shown to be increased extracellular via western blotting (Huber et al., 2020). At 8 hours 12 of the known interactors were abnormally secreted. All of the known Mfsd8 interactors observed to be apparently secreted after the loss of *mfsd8* were seen in increased amounts extracellular which suggests that Mfsd8 plays a role in regulating their release (Table 3.3).

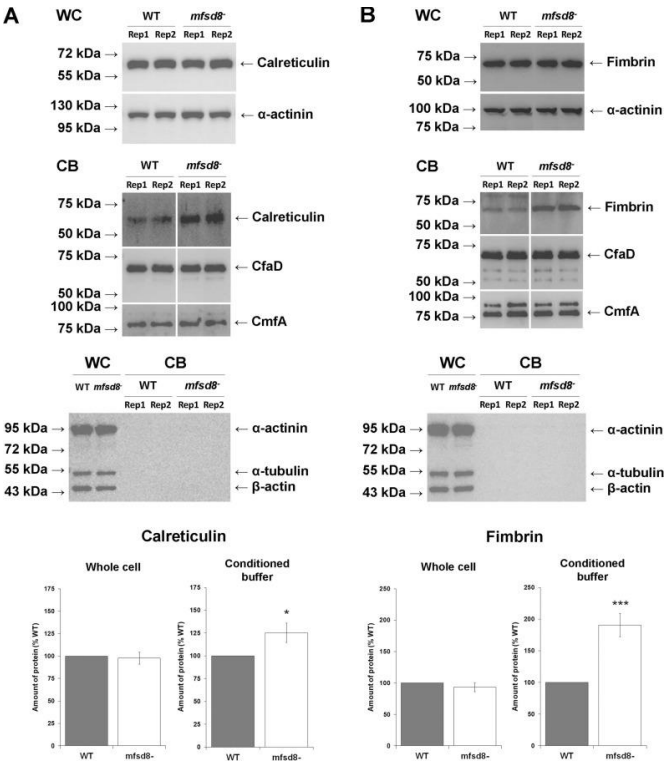

Figure 3.2. Effects of mfsd8-deficiency on intracellular and extracellular amounts of (A) calreticulin and (B) fimbrin. A) WT and *mfsd8*$^{-}$ cells were starved for 4 hours in KK2. 10 ug of Whole cell lysates (WC) and 0.4 conditioned buffer (CB) were separated through SDS page and analysed through western blotting with anti-calreticulin and anti-α-actinin. 2 representative band are shown. B) WT and *mfsd8*$^{-}$ cells were starved for 4 hours in KK2. 10 ug of Whole cell lysates (WC) and 0.4 conditioned buffer (CB) were separated through SDS page and analysed through western blotting with anti-fimbrin and anti-α-actinin. 2 representative band are shown. For both A and B a ll bands were quantified and standardized against α-actinin levels. CB blots were additionally probed with both anti-CfaD and anti-CmfA, to confirm lack of modification in CfaD and CmfA levels seen in mass spectrometry results, and antibodies against markers of cell lysis (β-actin, α-tubulin, α-actinin), to show that the differences seen were not due to cell lysis and

subsequent release of proteins into CB. *p < 0.05 and ***p < 0.001 (one sample t-test). Figure taken from Huber et al (2023).

Table 3.3. List of Mfsd8-interactors aberrantly released by *mfsd8*$^{-}$ cells after 4 and 8 hours of starvation

4h						
Uniprot ID	***dictyBase***	***dictyBase protein ID***	***Protein name***	***Gene name***	***p-value****	***Log2 FC****
Q54MA6	*DDB_G0286075*	*DDB0230022*	*40 S ribosomal protein S5*	*rps5*	*0.0025*	*0.97*
Q54Z69	*DDB_G0277803*	*DDB0231241*	*60 S ribosomal protein L4*	*rpl4*	*0.02*	*1.04*
O76856	*DDB_G0279411*	*DDB0215012*	*Cathepsin D*	*ctsD*	*0.028*	*1.25*
B0G141	*DDB_G0282979*	*DDB0234059*	*Glutathione S-transferase domain-containing protein*	*efa1G*	*0.0051*	*0.9*
Q58A42	*DDB_G0283095*	*DDB0231683*	*Protein DD3–3 (similar to tunicate proteins; involved in O-glycosylation as identified by mRNA differential display)*	*DD3–3*	*0.015*	*1.03*
P34144	*DDB_G0277869*	*DDB0214822*	*Rho-related protein rac1A*	*rac1A*	*0.017*	*0.51*
8h						
Uniprot ID	***dictyBase***	***dictyBase protein ID***	***Protein name***	***Gene name***	***p-value****	***Log2 FC****
Q54X51	*DDB_G0279207*	*DDB0231059*	*40 S ribosomal protein S19*	*rps19*	*0.039*	*1.1*
P90526	*DDB_G0293000*	*DDB0201667*	*40 S ribosomal protein S3*	*rps3*	*0.018*	*1.18*
Q55BE6	*DDB_G0271298*	*DDB0230153*	*60 S ribosomal protein L27*	*rpl27*	*0.047*	*1.33*
P34113	*DDB_G0291862*	*DDB0191094*	*60 S ribosomal protein L3*	*rpl3*	*0.023*	*1.2*
P42528	*DDB_G0283755*	*DDB0219936*	*Actin-related protein 3*	*arpC*	*0.039*	*0.74*
P0CT31	*DDB_G0269134*	*DDB0191135*	*Elongation factor 1-alpha*	*eef1a1*	*0.0013*	*1.2*
Q9GRF8	*DDB_G0284035*	*DDB0191174*	*Elongation factor 1-beta*	*efa1B*	*0.038*	*0.91*
P15112	*DDB_G0288373*	*DDB0191363*	*Elongation factor 2*	*efbA*	*0.037*	*0.88*
B0G141	*DDB_G0282979*	*DDB0234059*	*Glutathione S-transferase domain-containing protein*	*efa1G*	*0.011*	*1.36*
P54651	*DDB_G0267400*	*DDB0191163*	*Heat shock cognate 90 kDa protein*	*hspD*	*0.0004 1*	*1.58*
Q54LP8	*DDB_G0286509*	*DDB0231622*	*Histone H2B.v3*	*H2Bv3*	*0.011*	*0.91*
Q54EW3	*DDB_G0291650*	*DDB0266483*	*Probable importin-5 homolog (importin subunit beta-3; conserved protein that serves as receptor for nuclear localization signals (NLS) in cargo substrates and functions as nuclear transport receptor)*	*DDB_G0291650*	*0.016*	*0.94*
*Mann-Whitney test **Log2 FC (Mfsd8/WT)						

3.3.3 GO term enrichment analysis of proteins aberrantly released by *mfsd8*- cells during starvation

The functions and localizations of the differentially secreted proteins were determined using LAGO term enrichment analysis (https://go.princeton.edu/cgi-

bin/LAGO) Examination of the proteins aberrantly secreted due to *mfsd8* knock out after 4 hours of starvation (Table 3.4) showed the following functions being impacted. Stimulus (30%), stress response (21%), development (21%; sorocarp, aggregation, slug), catabolism (21%), phagocytosis (11%), and organization of the actin cytoskeleton (11%). Forty-three percent of the aberrantly secreted protein have catalytic activities (e.g., hydrolase, peptidase) and 21% of the proteins participate in protein binding. A majority of proteins localised to the cytoplasm (54%). However, proteins also localized to the vesicles (34%; endocytic, phagocytic), extracellularly (26%), cell periphery (20%), vacuoles (15%), lysosomes (13%), and cytoskeleton (11%) in WT cells were observed.

Regarding the proteins aberrantly secreted after 8 hours of starvation (Table 3.5) metabolic (43%) and biosynthetic (34%) processes, translation (20%), responses to external stimuli (16%) were functions affected by the loss *of mfsd8*. Proteins that form structural components of ribosomes (11%), bind of actin (11%), regulate of translation (8%), and have hydrolase activity (5%) were also enriched. Again, as in 4 hours, a majority of the proteins localised to the cytoplasm (59%). Like the 4 hours proteins also localised to vesicles (31%; phagocytic, endocytic), cell periphery (30%), extracellularly (21%) however ribosomes (11%), and the cell cortex (10%) were protein localizations observed as enriched. The enrichment analysis provides understanding regarding the cellular processes and pathways affected by the loss of *mfsd8* on D discoideum.

Table 3.4. GO term enrichment analysis of aberrantly released proteins after 4 hours of starvation.

Biological process			
GO ID	*GO Term*	*Number of proteins*	*% total*
GO:0050896	*Response to stimulus*	*18*	*30*
GO:0032502	*Developmental process*	*13*	*21*
GO:0009056	*Catabolic process*	*13*	*21*
GO:0006950	*Response to stress*	*13*	*21*
GO:0030587	*Sorocarp development*	*11*	*18*
GO:0006909	*Phagocytosis*	*7*	*11*
GO:0030036	*Actin cytoskeleton organization*	*7*	*11*
GO:0031152	*Aggregation involved in sorocarp development*	*6*	*10*
GO:0031153	*Slug development involved in sorocarp development*	*3*	*5*
Molecular function			
GO ID	*GO Term*	*Number of proteins*	*% total*
GO:0003824	*Catalytic activity*	*26*	*43*
GO:0016787	*Hydrolase activity*	*16*	*26*
GO:0005515	*Protein binding*	*13*	*21*
GO:0008233	*Peptidase activity*	*6*	*10*
GO:0051015	*Actin filament binding*	*5*	*8*
GO:0004630	*Phospholipase D activity*	*2*	*3*
Cellular component			
GO ID	*GO Term*	*Number of proteins*	*% total*
GO:0005737	*Cytoplasm*	*33*	*54*
GO:0031982	*Vesicle*	*21*	*34*
GO:0030139	*Endocytic vesicle*	*18*	*30*
GO:0045335	*Phagocytic vesicle*	*17*	*28*
GO:0005576	*Extracellular region*	*16*	*26*
GO:0071944	*Cell periphery*	*12*	*20*

GO:0005773	*Vacuole*	*9*	*15*
GO:0005764	*Lysosome*	*8*	*13*
GO:0005856	*Cytoskeleton*	*7*	*11*

Table 3.5. GO term enrichment analysis of aberrantly released proteins after 8 hours of starvation.

Biological process			
GO ID	*GO Term*	*Number of proteins*	*% total*
GO:1901564	*Organonitrogen compound metabolic process*	*26*	*43*
GO:0009058	*Biosynthetic process*	*21*	*34*
GO:0006412	*Translation*	*12*	*20*
GO:0009605	*Response to external stimulus*	*10*	*16*
Molecular function			
GO ID	*GO Term*	*Number of proteins*	*% total*
GO:0003735	*Structural constituent of ribosome*	*7*	*11*
GO:0003779	*Actin binding*	*7*	*11*
GO:0045182	*Translation regulator activity*	*5*	*8*
GO:0008081	*Phosphoric diester hydrolase activity*	*3*	*5*
Cellular component			
GO ID	*GO Term*	*Number of proteins*	*% total*
GO:0005737	*Cytoplasm*	*36*	*59*
GO:0031982	*Vesicle*	*19*	*31*
GO:0071944	*Cell periphery*	*18*	*30*
GO:0045335	*Phagocytic vesicle*	*17*	*28*
GO:0030139	*Endocytic vesicle*	*17*	*28*
GO:0031012	*Extracellular matrix*	*13*	*21*
GO:0005840	*Ribosome*	*7*	*11*
GO:0005938	*Cell cortex*	*6*	*10*

3.3.4 Effect of *mfsd8*-deficiency on the proteasome

Mass spectrometry of the 8-hour starved condition buffer showed an increase in proteasome subunit beta type 6 (PsmB6) was aberrantly secreted. PsmB6 contributes to the formation the P20 proteasome as such *mfsd8* loss influencing proteasome activity was hypothesized. This hypothesis was supported by the go term analysis showing 43% of proteins aberrantly secreted in 8 hour starved conditioned buffer having roles in metabolic processes. With this foundation a commercially available kit for measuring p20s activity was used to assess the activity of proteasome of whole cell lysate samples of WT and *mfsd8* cells after 4 and 8 hours of starvation (Figure 3.3). The loss of *mfsd8* significantly affected proteasome 20S activity as after 4 and 8 hours of starvation the activity decreased by 61 ± 10% and 35 ± 9%, respectively. These results indicate that the loss of *mfsd8* causes a reduction in proteasome 20S activity after the secretion of the proteasomal subunit occurring after starvation.

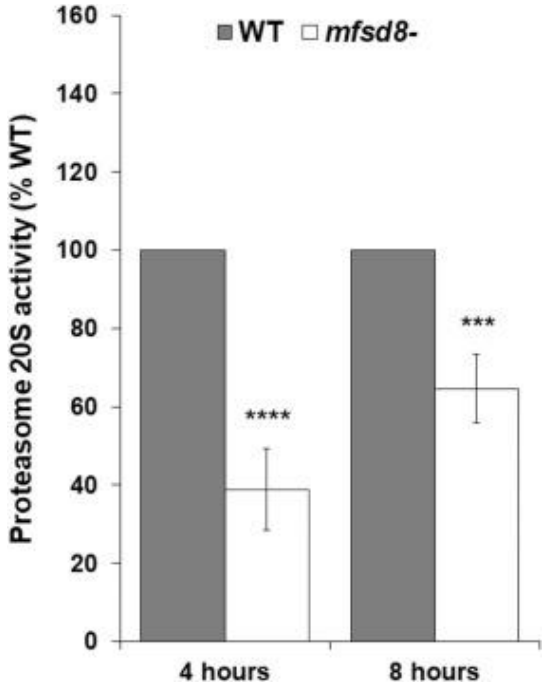

Figure 3.3. Effects of mfsd8-deficiency on proteosome 20S activity. P20S activity was assessed in both WT and mfsd8- cells 4 hour and 8 hours post starvation. Data presented as mean proteasome 20S activity (% WT) ± SEM (n = 8). ***p < 0.001 and ****p < 0.0001 (one sample t-test). Figure taken from Huber et al (2023).

3.3.5 Comparing proteins aberrantly secreted by *mfsd8*$^{-}$ and *cln3*$^{-}$ cells

In previous work, the secretome of early multicellular *cln3*$^{-}$ *D. discoideum* was reported (Huber, 2017) and evidence suggests that CLN genes have a converging and common biological pathway (Huber, 2020; Huber, 2023). The list of proteins released from *mfsd8*$^{-}$ *cells and cln3*$^{-}$ cells were compared and 7 proteins were shown to be common to both secretomes (Table 3.6). Included in the list of shared proteins were CtsD a known Mfsd8 interactor (Huber et al., 2020), the lysosomal protease cathepsin B (CtsB), which has been linked to several NCL subtypes (Huber, 2023). CtsB was shown to be increased in neuronal cell cultures of mice and increased in the urine of sheep with CLN6 disease (Iwan et al., 2020; Best et al., 2021). CtsB was the second most reduced protein after 8 hours of starvation (log_2FC -1.86). Loss of *mfsd8* also affected the extracellular amount of cysteine protease E (CprE) increasing it. CprE shares sequence similarity with human cathepsin F(CTSF) which has been linked to CLN13 Disease (Huber et al., 2020). *mfsd8*$^{-}$ condition buffer also contained cathepsin Z which was not secreted by *cln3*$^{-}$ cells. The release of CtsZ from *mfsd8*$^{-}$ cells aligns with mammalian work that found CTSZ in the cerebrospinal fluid of CLN1 and CLN2 diseased mice and the urine of

human with CLN2 Disease (Sleat et al., 2019; Iwan et al., 2020). These finding support the notion that CLN proteins function in converging pathways.

Table 3.6. List of proteins aberrantly released by *mfsd8*$^{-}$ cells after 4 and 8 hours of starvation and *cln3*$^{-}$ cells after 6 h of starvation

4 h						
Uniprot ID	**dictyBase gene ID**	**dictyBase protein ID**	**Protein name**	**Gene name**	***cln3*$^{-}$**	***mfsd8*$^{-}$**
O96624	DDB_G0292804	DDB0201632	Actin-related protein 2/3 complex subunit 3	*arcC*	Increase	Increase
Q54RZ4	DDB_G0282815	DDB0214941	BB_PF domain-containing protein (similar to SmlA)	*orfSGP*	Increase	Decrease
O76856	DDB_G0279411	DDB0215012	Cathepsin D	*ctsD*	Increase	Increase
P54640	DDB_G0272815	DDB0185092	Cysteine proteinase 5	*cprE*	Increase	Increase
Q55EA3	DDB_G0269322	DDB0306569	Uncharacterized protein	*DDB_G0269322*	Increase	Decrease
8 h						
Uniprot ID	**dictyBase gene ID**	**dictyBase protein ID**	**Protein name**	**Gene name**	***cln3*$^{-}$**	***mfsd8*$^{-}$**
Q54QD9	DDB_G0283921	DDB0233997	Cathepsin B	*ctsB*	Increase	Decrease
Q54P69	DDB_G0284759	DDB0231729	Protein psiM	*psiM*	Decrease	Decrease

3.3.6 The effect of *mfsd8*-deficiency on the release of Cln5-interactors

As seen in previous research, there is an impact on extracellular *cln5* seen in *mfsd8*$^{-}$ cells (Huber et al., 2020). As this is the case, release of cln5 interactions during starvation was examined. After 4 hours of starvation mass spectrometry identified 10 Cln5 interactors. This included calreticulin which has had the increased secretion seen in

mass spectrometry validated by western blotting and other studies observing neuronal mouse cell cultures and sheep urine (Iwan et al., 2020, Best et al., 2021). Other interactors include Ctsd and the proteins DDB0252831 and CprE, which both share sequence similarity with human CTSF (Huber et al., 2020). These results indicate that the loss of mfds8 affects the release of Cln5 and its interactors**.**

In previous research, there is an impact on extracellular cln5 seen in *mfsd8*$^{-}$ cells (Huber et al., 2020). as loss of *mfsd8* changes the extracellular Cln5 amounts during starvation the affect on the known Cln5 interactors was examined (Table 3.7). During 4 hours of starvation loss of *mfsd8* affected both CtsD and calreticulin. The observed increase in calreticulin was also mirrored in western blotting results. Research showed that extracellular calreticulin was increased in neuronal cell cultures of mice and decreased in the urine of sheep (urine decrease) models of CLN6 disease. DDB0252831 and CprE, which both share sequence similarity to human CTSF (Huber et al., 2020) also, were seen in increased amounts extracellularly. During 8 hours of starvation 6 Cln5 interactors were seen, including two heat shock proteins. These results suggest that the loss of *mfsd8* affects the release of Cln5 interactors in addition to the release of Cln5.

Table 3.7. List of Cln5-interactors aberrantly released by mfsd8^{-} cells after 4 and 8 hours of starvation

4 h						
Uniprot ID	**dictyBase gene ID**	**dictyBase protein ID**	**Protein name**	**Gene name**	**p-value***	**Log2 FC****
Q23858	DDB_G0283539	DDB0191384	Calreticulin	*crtA*	0.027	1.36
P12019	DDB_G0285995	DDB0219974	cAMP/cGMP-dependent 3',5'-cAMP/cGMP phosphodiesterase A	*pdsA*	0.0044	2.74
O76856	DDB_G0279411	DDB0215012	Cathepsin D	*ctsD*	0.028	1.25
P22549	DDB_G0277863	DDB0214948	Cyclic nucleotide phosphodiesterase inhibitor	*pdiA*	0.0078	1.34
P54640	DDB_G0272815	DDB0185092	Cysteine proteinase 5	*cprE*	0.011	1.91
Q9NKX1	DDB_G0280057	DDB0215015	Endoplasmin homolog	*grp94*	0.0032	3.93
B0G141	DDB_G0282979	DDB0234059	Glutathione S-transferase domain-containing protein	*efa1G*	0.0051	0.9
Q23892	DDB_G0292810	DDB0215373	Lysosomal beta glucosidase	*gluA*	0.036	-1.99
Q54GG6	DDB_G0290177	DDB0238177	Uncharacterized protein (contains an N-terminal signal sequence and a C-terminal transmembrane domain; under expressed in gskA- and zakA- null mutants)	*DDB_G0290177*	0.05	0.97
Q54F16	DDB_G0291191	DDB0252831	Uncharacterized protein (highly similar to Dictyostelium cprA)	*DDB_G0291191*	0.023	2.04
8 h						
Uniprot ID	**dictyBase gene ID**	**dictyBase protein ID**	**Protein name**	**Gene name**	**p-value***	**Log2 FC****

P90532	DDB_G0288065	DDB0191154	Cell division cycle protein 48	*cdcD*	0.012	0.69
P0CT31	DDB_G0269134	DDB0191135	Elongation factor 1-alpha	*eef1a1*	0.0013	1.2
B0G141	DDB_G0282979	DDB0234059	Glutathione S-transferase domain-containing protein	*efa1G*	0.011	1.36
P36415	DDB_G0269144	DDB0191168	Heat shock cognate 70 kDa protein 1	*hspB*	0.025	0.62
P54651	DDB_G0267400	DDB0191163	Heat shock cognate 90 kDa protein	*hspD*	0.0004	1.58
Q95US4	DDB_G0279921	DDB0214937	Lipid-anchored plasma membrane glycoprotein 130	*gp130*	0.015	-1.16
*Mann-Whitney test **Log2 FC (Mfsd8/WT)						

3.3.7 Mass spectrometry identifies aberrantly released proteins linked to *mfsd8*-deficiency phenotypes during early development

Deletion of *mfsd8* was shown to compromise cell substrate adhesion in previous finding (Yap et al., 2022). Mass spectrometry results revealed the secretion of several protein linked to adhesion including disintegrin-like protein GdcA, the substrate adhesion protein SadA, and the multicellular aggregate-regulating protein SmlA which were all seen with increased extracellular amounts (Musial et al.,1990; Cesar et al., 2019; Fey et al., 2002; Roisin-Bouffay et al., 2000). Loss of *mfsd8* also altered the extracellular levels of proteins linked to cAMP signalling including coronin A which had decreased levels and, the cAMP phosphodiesterase PdsA, and the cAMP phosphodiesterase inhibitor PdiA which had increased levels (Vinet et al., 2014; Barra et al., 1980; Franke and Kessin, 1981). Vacuolar protein sorting-associated protein 35

(Vps35), which regulates intracellular trafficking in yeast and humans (retrograde transport from endosomes to the Golgi complex), was seen to have increased levels extracellularly during both starvation for 4 and 8 hours (Paravicini et al., 1992; Seaman et al., 1997; Edgar and Polak, 2000). Vps35 was not previously detected in the AX2 or AX3 secretomes which suggests that its not normally secreted by WT cells (Bakthavatsalam and Gomer, 2010; Huber, 2017). Grp94, The D. discoideum homolog of the human endoplasmic reticulum protein endoplasmin, which is a molecular chaperone that plays a role in the processing and transport of secreted proteins, was also observed to be at increased extracellular levels (Morita et al., 2000). Grp94 had the most significant increase in extracellular levels (log2FC 3.93). These findings identify released proteins that are associated with *mfsd8*$^{-}$ phenotypes observed during early multicellular development.

3.4 Discussion

3.4.1 Overview of Effect of *mfsd8*-deficiency on protein secretion during the early stages of *D. discoideum* development

This study showed that *mfsd8*-deficiency affects the secretome during early stages of *D. discoideum* multicellular development. GO term enrichment analyses revealed the processes and functions of the differentially released proteins. Several of the enriched terms were associated with the proteins and are linked to phenotypes observed in *mfsd8*$^{-}$ aggregation. This study also found the first evidence in any system linking MFSD8 to the proteasome. The findings in this study provides new avenues to

elucidate the function of Mfsd8 in *D. discoideum* and provides the foundation to future work in mammalian models which will provide a more comprehensive understanding of CLN7 disease.

3.4.2 *mfsd8*'s deletion perturbs protein secretion

61 proteins aberrantly released by *mfsd8*-deficient cells were identified after 4 and 8 hours of starvation with 54 unique proteins at each time point. The majority of proteins seen were in increased amounts extracellularly. Along with this, one third of the proteins identified during this study were not detected in previous studies examining the secretomes of AX2 or AX3 (Bakthavatsalam and Gomer,2010; Huber, 2017). This seems to indicate that the deletion of *mfsd8* perturbs protein secretion and results in the expulsion of protein not normally released from WT cells and retention of protein normally released. This is supported by the release of proteins associated with translation being released and detected in *mfsd8*-deficient condition buffer. This is also supported by proteins normally released into condition buffer not being present (Bakthavatsalam and Gomer, 2010; Huber, 2017). These findings are a main reason for the generation of *mfsd8*$^{-}$ phenotypes during the early stages of multicellular development.

Increased extracellular calreticulin and fibrin during 4 hours of starvation and 8 hours of starvation respectively were supported through the use of commercially available antibodies and western blotting. The western blotting analysis also showed that while extracellular amounts increased, the intracellular amount of the proteins were

unchanged. These findings indicate that the cells modulate the intracellular amounts of protein regardless of the release of protein from the cell. It was previously shown that *mfsd8 deletion* reduces the intracellular amount of CtsD but increases the conditioned buffer (Huber et al., 2020). This finding supports the active release of CtsD and is consistent with the amount observed in conditioned buffer via mass spectrometry. CfaD and CmfA results from mass spectrometry were also corroborated through the use of western blotting as both proteins were shown to not have variance between WT and *mfsd8*$^{-}$ cells.

3.4.3 Effect of *mfsd8* deletion on the conventional and unconventional pathways of protein secretion

Many of the proteins aberrantly secreted from *mfsd8*-deficient cells have putative signal peptides for secretion and as such it is possible that the deletion of *mfsd8* perturbs the conventional secretions pathway in which proteins transit from the endoplasmic reticulum to Golgi apparatus to secretory vesicles to be shuttled outside of the cell (Viotti, 2016). This notion is supported by data from GO term enrichment analysis as many of the differentially secreted proteins are located in regions and compartments of the endo-lysosomal system and secretory pathways. The notion that the conventional pathway is perturbed is also supported by the finding that all proteins released in reduced amounts after 8 hours had secretory punitive signal peptides. In studies it was shown that that Mfsd8 *is* localized to the cytoplasmic puncta and vesicles (Huber et al., 2020) as not all these structures were marked by vat the acidic vesicle

marker VatC (catalytic subunit of V-ATPase) or the secretory lysosome marker p80 (putative copper transporter) it is theorized that there is also a cytoplasmic, non-lysosomal pool of Mfsd8 in *D. discoideum*. As there are non lysosomal proteins being released lysosomal defects and or exocytosis cannot be solely responsible for the finding. These findings support the theory that Mfsd8 plays a role in the conventional pathway.

Unconventional pathways of protein secretion have also been revealed in *D. discoideum* and mammalian models which allow for the secretion of proteins lacking a signal peptide (Viotti, 2016; Filaquier et al., 2022). Some of these pathways have been shown to involve autophagy (Kinseth et al., 2007; Duran et al., 2010; Noh et al., 2022), which has been reported to behave aberrantly in mouse models of CLN7 disease (Brandenstein et al., 2016; Lopez-Fabuel et al., 2022) as such the loss of *mfsd8* could affect unconventional secretion pathways. This is supported by the fact that only 66% of the proteins at 4 hours and 36% of the proteins at 8 hours having signal peptides for secretion further implicating perturbations in the unconventional pathway. This aberration of the unconventional secretory pathways could explain the cytosolic proteins present within the conditioned buffer of *mfsd8*-deficient *D. discoideum.* The abnormal release of cell-derived extracellular vesicles or exosomes would also play a role in secretome alterations seen in *mfsd8*$^{-}$ *D. discoideum* (Tatischeff et al., 2012; Tatischeff, 2019). Reduced endocytosis or increased exocytosis could be a rational for proteins observed and studies in Cln7$^{-/-}$ mice have found impaired lysosomal exocytosis

(Danyukova et al., 2018) however, a reduction in endocytosis does not explain the reduced amount of some proteins present and previous research has shown that mfsd8-deficiency induces an increase in fluid phase exocytosis (Yap et al., 2022). This supports the reasoning that secretory pathways are what is being affected by the deletion of *mfsd8*. The altered secretory pathways combined with an increase in Mfsd8 interacting proteins suggest that the secretion of some proteins is regulated by mfsd8.

3.4.4 Deletion of *mfsd8* alters extracellular amounts of lysosomal proteins and CLN proteins

There has been work (Yap et al., 2022) showing that *mfsd8*$^{-}$ cells have altered activity of known lysosomal proteins. Our work showed that several lysosomal proteins including beta-glucosidase, beta-galactosidase, various cathepsins (CtsD, CtsB, CtsZ), and proteins that share sequence similarity with human CTSF (uncharacterized protein DDB0252831, CprE) have altered extracellular amounts. The detection of increased extracellular amounts of Ctsd is mirrored in studies which used western blotting to reveal this information (Huber et al., 2020). An increase of Cln5 was seen in the western blotting results of which is mirrored in our work by our mass spectrometry finding however our increase was not statistically significant (Huber,2020). Several known Cln5 interactors were aberrantly secreted in *mfsd8*$^{-}$ *conditioned buffer.* Previous work examining Cln3 revealed that it also effects extracellular levels of both CtsD and Cln5 (Huber, 2017), this suggests the possibility of a common pathway in the regulation of trafficking the two CLN proteins. *cln3* and *mfsd8*-deficiency also causes that aberrant

secretion of CtsB and CprE (Huber, 2017). CtsB is a known biomarker for CLN6 Disease (Huber, 2021). *mfsd8*$^{-}$ condition buffer also contained increased amounts of CtsZ which has been observed to be increased in Cerebrospinal fluid of patients with CLN2 disease (Sleat et al., 2019; Iwan et al., 2020). As a whole this supports the notion that the *CLN* genes and Proteins function as a Molecular network in observed models (Huber, 2020; Huber, 2023).

3.4.5 The secretome of *mfsd8*$^{-}$ cells contributes to observed phenotypes

Normal *D. discoideum* development involves the secretion of proteins into the environment to facilitate cAMP-mediated chemotaxis and aggregation (Rossomando et al., 1978; Dimond et al.,1981; Gomer et al., 1991). It has been shown that the loss of *mfsd8* generates a phenotype characterised by altered lysosomal enzyme activity, reduced cell-substrate adhesion, and delayed aggregation during the early stages of multicellular development (Yap et al.,2022). We show that *mfsd8*$^{-}$ cells have the abnormal release of several enzymes which act on aggregation and adhesion and are therefor the possible reasons for the phenotypical differences. GdcA, SadA, and SmlA, proteins associated with adhesion (Musial et al., 1990; Cesar et al., 2019; Fey et al., 2002; Roisin-Bouffay et al., 2000)., all had increased extracellular levels seen in the mass spectrometry data. Disintegrins prevent integrin-dependent cell adhesion in mammals as such the increased amount of GdcA in *mfsd8*$^{-}$ conditioned buffer could explain the aberrant adhesion of mfsd8- cells (Yap et al., 2022). Increases in SadA, which is thought to function as an integrin, could be due to cells trying to properly regulate

integrin/dintegrin activities. Loss of *mfsd8* also impacted the levels of several proteins linked to cAMP signaling. Coronin A was decreased while PdsA and PdiA were increased (Vinet et al., 2014; Barra et al., 1980; Franke and Kessin, 1981). Previous work showed that their cAMP mediated chemotaxis was not affected by the loss of *mfsd8 (Yap et al., 2022).* This work indicates that *mfsd8*$^{-}$ cells offset the loss from increased extracellular PdsA by increasing the secretion of PdiA to compensate and therefor allow the regulation of extracellular levels of cAMP. The trafficking and secretion-regulating proteins Vps35 and Grp94 are also affected by *mfsd8*$^{-}$ deletion which suggests a link between Mfsd8 and trafficking and secretion. Another possible reason is the abnormal release of lysosomal proteins preventing the normal degradation of autophagic elements which could contribute to abnormal adhesion and aggregation. Together these findings provide new understanding to the proteins behind the mechanism of the *mfsd8*-deficient phenotypes.

3.4.6 Support for previous works examining CLN7 disease and *D. discoideum*

This study supports work in mammalian models. A mouse model of CLN7 disease progression linked MFSD8 to roles in regulating the structure and function of the mitochondrion (Lopez-Fabuel et al., 2022). We found that 43% of the proteins aberrantly released by cells after 8 hours of starvation are linked to metabolic processes.

Our findings presented an abnormal amount of the proteosome subunit PsmB6 in *mfs8d*$^{-}$ conditioned buffer and consistent with that finding presented a significant reduction of Proteasomal 20s activity. This is the first finding linking MFSD8 to the

proteosome. As such this work, examining *D. discoideum* generated a scaffold for future researchers to examine the role of that organelle in the pathology of CLN7.

This study also adds to our collective knowledge of the WT *D. discoideum Secretome.* The strains AX2 and AX3 previously had their secretomes analysis (Bakthavatsalam and Gomer, 2010; Huber, 2017) leaving AX4, the strain with its genome sequenced as a part of the *D. discoideum* genome project (Eichinger et al., 2005) as an unknown. This work provides a catalog of the proteins secreted during 4 and 8 hours of starvation which compliments our understanding of the species secretome.

3.4.7 Conclusion

In conclusion, this series of experiments not only provide a catalogue of the proteins released during *D. discoideum* early multicellular development, but they also provide new possible insights into the role of MFSD8 in *D. discoideum* and in human. It provides new scaffolding by which new clinical insights can be found regarding the effect of protein secretion which can be used to elucidate novel diagnostic, prognostic and therapy response biomarkers for CLN7 disease.

Chapter 4- General Discussion

4.1 Overview of the effect of mfsd8 deletion on the secretome and transcriptome of *D discoideum*

As shown by the previous two articles the loss of mfsd8 affects the transcriptome and secretome in ways which relate to the observed phenotypical differences observed in mfsd8$^-$ cells. In respect to the secretomic work, 61 proteins were observed to be differentially secreted between 4 and 8 hours of starvation. Similarly, the transcription portion of this thesis revealed 445 genes aberrantly expressed during the growth stage and 4954 genes differentially expressed during starvation. Many of the observed proteins, and protein products of the genes are known interactors of mfsd8 and proteins which underlie phenotypical differences. In the secretomic dataset, it was observed that there were alterations in the release of several lysosomal enzymes which could be reducing substrate adhesion through their extracellular function or by the alterations preventing the normal degradation of intracellular autophagic material which could reduce adhesion and slow aggregation. Similarly, the transcriptomic dataset also includes several known lysosomal enzymes and genes relating to proper cell adhesion and processes by which it is known that *D. discoideum* cells utilize to aggregate. Several genes observed in the transcriptomic dataset as having altered transcription relate to cellular secretion and as such could contribute to the altered secretome and said altered secretome having most proteins showing increased extra cellular amounts suggesting that the loss of *mfsd8* played a role in the dysregulation of secretion.

4.2 The shared effects of mfsd8 deletion on the secretome and transcriptome of *D. discoideum*

To discuss the datasets of the two papers there is a crossover between the transcriptomic and secretomic data sets. Comparing the 8-hour transcriptomic and secretome datasets of the 16 proteins under secreted compared to WT 10 of those were observed within the 8-hour transcriptomic dataset and of the 45 proteins oversecreted 25 had corresponding genes with aberrant transcription observed in *mfsd8*$^{-}$ cells (Table 4.1). Fimbrin and several other proteins involved in actin polymerization (Prassler et al., 1997) were affected by the loss of *mfsd8* further compounding the relation of datasets. Interestingly sphingomyelinase B and D were affected by the loss of *mfsd8* in both datasets. Sphingomyelinases function in the sphingomyelin degradation and play roles in signal transduction and further in the transcriptomic dataset Sphingosine-1-phosphate lyase is also affected (increase) (Li et al., 2001). It is known that impacting sphingomyelin degradation can affect Ca^{2+} stores, actin polymerization, cell motility, chemotaxis, and cell differentiation (Li et al., 2001).

Table 4.1. Genes and proteins observed in both 8-hour transcriptomic and 8-hour secretomic dataset

Uniprot ID	dictyBase gene ID	dictyBase protein ID	Protein name	Gene name	Secretomic Log2 FC (Mfsd8/ AX4)	Transcriptomic Log2 FC (Mfsd8/ AX4)
B0G0Z4 (+1)	DDB_G0269616	DDB0237972, DDB0237973	Saposin B domain-containing protein (splice variant A contains a putative signal peptide)	*aplG*	-1.66	-0.434
Q54GI7	DDB_G0290139	DDB0237739	Peptidase S28 family protein (contains a putative signal peptide)	DDB_G0290139	-1.49	0.51
Q54HY0	DDB_G0289145	DDB0238626	cAMP/cGMP-dependent 3',5'-cAMP/cGMP phosphodiesterase 7	*pde7*	-1.26	0.322
Q54P69	DDB_G0284759	DDB0231729	Protein psiM	*psiM*	-1.18	0.876
Q55EM4	DDB_G0268828	DDB0238635	Uncharacterized protein (contains a predicted signal peptide; highly similar to neighboring genes DDB_G0269008)	DDB0190067	-1.05	-0.813
Q54XG9	DDB_G0278975	DDB0304438	G8 domain-containing protein DDB_G0278975 (similar to ComF)	DDB_G0278975	-1.03	0.501
Q54C16	DDB_G0293210	DDB0232049	Sphingomyelin phosphodiesterase B	*sgmB*	-0.78	1.061
Q8T2U1	DDB_G0275253	DDB0229926	Uncharacterized protein (expressed in pstO cells)	DDB0229926	-0.7	0.477
Q55GC7	DDB_G0268330	DDB0232051	Sphingomyelinase phosphodiesterase D	*sgmD*	-0.6	0.494
Q55EJ2	DDB_G0268870	DDB0306557	Uncharacterized protein	DDB0190094	-0.6	0.418
Q54X73	DDB_G0279159	DDB0229908	Probable cytoplasmic aconitate hydratase	*aco1*	0.55	0.746
P36415	DDB_G0269144	DDB0191168	Heat shock cognate 70 kDa protein 1	*hspB*	0.62	1.119
P90532	DDB_G0288065	DDB0191154	Cell division cycle protein 48	*cdcD*	0.69	0.766
P42528	DDB_G0283755	DDB0219936	Actin-related protein 3	*arpC*	0.74	0.468
P15112	DDB_G0288373	DDB0191363	Elongation factor 2	*efbA*	0.88	-0.379
Q54LP8	DDB_G0286509	DDB0231622	Histone H2B.v3	*H2Bv3*	0.91	-2.1
Q54RX6	DDB_G0282853	DDB0305046	Translationally controlled tumor protein homolog 1 (ortholog of human TPT1, a protein	*tpt1*	0.91	-0.634

			associated with tumors, and to S. cerevisiae TMA19, a protein associated with ribosomes)			
Q54VN6	DDB_G0280229	DDB0230151	60S ribosomal protein L24	*rpl24*	0.93	-1.037
P34122	DDB_G0277501	DDB0185023, DDB0233889	cAMP-binding protein 2	*capB*	0.93	-0.569
Q55GT3	DDB_G0267528	DDB0306483	Uncharacterized protein	DDB0189346	1.06	-0.977
P54680	DDB_G0277855	DDB0214994	Fimbrin	*fimA*	1.09	0.703
Q54X51	DDB_G0279207	DDB0231059	40S ribosomal protein S19	*rps19*	1.1	-7.233
Q54NG9	DDB_G0285267	DDB0238270	Probable saccharopine dehydrogenase [NADP(+), L-glutamate-forming]	*Sdh*	1.12	0.635
Q86KC7	DDB_G0274171	DDB0230011	Uncharacterized protein DDB_G0274171 (similar to Fraser syndrome protein, von Willebrand factor, and kielin; expressed in pstAO cells and in upper cup during culmination)	*stcC*	1.14	0.353
P90526	DDB_G0293000	DDB0201667	40S ribosomal protein S3	*rps3*	1.18	-0.515
P34113	DDB_G0291862	DDB0191094	60S ribosomal protein L3	*rpl3*	1.2	-0.533
P0CT31	DDB_G0269134	DDB0191135	Elongation factor 1-alpha	*eef1a1*	1.2	-0.533
Q552S7	DDB_G0275913	DDB0305472	Putative acetyltransferase DDB_G0275913 (similar to bacterial acetyltransferases; homolog of E. coli maa (maltose O-acetyltransferase); similar to D. purpureum protein)	DDB_G0275913	1.22	0.472
Q55BE6	DDB_G0271298	DDB0230153	60S ribosomal protein L27	*rpl27*	1.33	-0.755
Q54GK6	DDB_G0290091	DDB0230147	60S ribosomal protein L22 2	*rpl22a*	1.4	-0.456
Q75JY8	DDB_G0277615	DDB0232276	Uncharacterized protein (putative actin-binding protein; proteins containing the actin depolymerisation factor (ADF)/cofilin-like domain sever actin filaments and bind to actin monomers)	DDB_G0277615	1.45	0.47
P54654	DDB_G0288769	DDB0191139	Adenylyl cyclase-associated protein	*Cap*	1.46	0.879

P54651	DDB_G0267400	DDB0191163	Heat shock cognate 90 kDa protein	*hspD*	1.58	0.672
P25870	DDB_G0277221	DDB0185029	Clathrin heavy chain	*chcA*	1.74	0.564
Q55C77	DDB_G0270184	DDB0305155	GDP-L-fucose synthase	*Ger*	1.96	0.39

During the 4-hour starvation period there are several genes also observed in the 8-hour transcriptomic data set (Table 4.2) including Phosphodiesterase Inhibitor A and B. In the case of both data sets there was an increase in the transcription of both and an increase in the extracellular amount of the resulting protein cAMP phosphodiesterase inhibitor. Studies examining overexpression found that the overexpression of the gene in one mutant line led to delays in aggregation and in another line found that the mutation abolished culmination, slug development and caused the aberrant development of tipped mound (Wu et al., 1995), which align with the observed phenotypical characteristics of our *mfsd8* knockout cell line.

Table 4.2. Genes/Proteins observed in both 8-hour transcriptomic and 4-hour secretomic dataset

Uniprot ID	DictyBase gene ID	DictyBase protein ID	Protein name	Gene name	Log2 FC secretomic (Mfsd8/ AX4)	Log2 FC transcriptomic (Mfsd8/ AX4)
Q55EA3	DDB_G0269322	DDB0306569	Uncharacterized protein	DDB0190177	-2.2	0.538
Q556G3	DDB_G0274081	DDB0238181	Putative glutathione S-transferase alpha-2	*gsta2-1*	0.65	-2.251
Q54FW2	DDB_G0290659	DDB0238213	Uncharacterized protein (short-chain dehydrogenase/reductase (SDR) family protein)	*sdrA*	0.67	-2.076
Q54X73	DDB_G0279159	DDB0229908	Probable cytoplasmic aconitate hydratase	*aco1*	0.69	0.746
Q23919	DDB_G0288483	DDB0191348	Phosphoglucomutase-1	*pgmA*	0.9	1.099
Q94494	DDB_G0289393	DDB0191422	Protein psiH	*psiH*	0.91	0.9
Q86KC7	DDB_G0274171	DDB0230011	Uncharacterized protein DDB_G0274171 (similar to Fraser syndrome protein, von Willebrand factor, and kielin; expressed in pstAO cells and in upper cup during culmination)	*stcC*	0.95	0.353
Q54MA6	DDB_G0286075	DDB0230022	40S ribosomal protein S5	*rps5*	0.97	-0.914
Q869Q6	DDB_G0274107	DDB0185229	Uncharacterized protein	*sfbA*	1	0.355
Q54Z69	DDB_G0277803	DDB0231241	60S ribosomal protein L4	*rpl4*	1.04	-5.995
P0DJ26	DDB_G0277833	DDB0214987	Cofilin-1A	*cofA*	1.06	-4.73
Q54T09	DDB_G0282061	DDB0306992	Uncharacterized transmembrane protein DDB_G0282061	*DDB_G0282061*	1.09	-1.311
O76856	DDB_G0279411	DDB0215012	Cathepsin D	*ctsD*	1.25	0.699
P22549	DDB_G0277863	DDB0214948	Cyclic nucleotide phosphodiesterase inhibitor	*pdiA*	1.34	0.682
Q8I7T3	DDB_G0288511	DDB0191090	Substrate-adhesion molecule	*sadA*	1.36	0.451
Q54EN4	DDB_G0291434	DDB0231409	Protein disulfide-isomerase 2	*pdi2*	1.53	0.713
Q54VM2	DDB_G0280255	DDB0230187	Probable malate dehydrogenase 3	*mdhC*	1.82	0.329
Q54SA1	DDB_G0282579	DDB0231505	Phospholipase D Z	*pldZ*	1.95	0.337
Q54K50	DDB_G0287649	DDB0220113	Phospholipase D Y	*pldY*	2.12	0.437
Q54GE1	DDB_G0290217	DDB0266380	Beta-galactosidase 1	*glb1*	2.53	0.652
Q9NKX1	DDB_G0280057	DDB0215015	Endoplasmin homolog	*grp94*	3.93	1.128

4.3 Rational for and possible future experiments examining mitochondria

Fundamentally, our research has found several new and interesting findings as both the secretomic and transcriptomic portions of this work has unveiled possible mechanism behind the phenotypic differences generated by the loss of *mfsd8*. Interestingly, in both datasets there are data pointing to possible issues with mitochondria, as such, in future it would be imperative to examine some of the parameters of possible mitochondria disfunction. Batten disease has been linked to mitochondrial dysfunction, the accumulation of Subunit c of ATP synthase, which is a mitochondrial protein that is a hallmark of the disease (Palmer, 2006). In cln3 models in BD lymphoblast cells, several known mitochondrial related metabolic molecules had aberrant expression compared to wild type cells (Kang et al., 2013). Morphological differences have also been observed as enlarged mitochondria in GABAergic neurons of the neocortex, claustrum and basket cells of the cerebellum in the English setter model (Walkley et al., 1995). Enlarged cytochrome oxidase positive granules that were not present in other neurons were also observed (Walkley et al., 1995). Together both are theorised to be due to increases in mitochondrial stress or due to altered trafficking of subunit c of ATP synthase. This continues to cln7 disease as mouse models were observed to have structurally and bioenergetically impaired neuronal mitochondria along with increases in mitochondrial reactive oxygen species in said neurons (Lopez-Fabuel et al., 2022). CLN disease 1,2 and 3 been shown to have reduced basal ATP synthase activity (Das et al., 1999). Together this shows support for the need to examine

the mitochondrial function in *D. discoideum* to further understand the mitochondrial disfunction observed in CLN disease.

Dictyostelium, in addition to its uses in examination of neurological conditions also has a history as a model for mitochondrial diseases (Barth 2007). As such, there is a host of techniques that could be applied to solve this query. This could be achieved in several ways; electron microscopy (Chida et al.,2004) could be used to examine mitochondria for increased size or enlarged cytochrome oxidase positive granules which would be consistent with previously observed characterizations of mitochondrial disfunction seen in other models of CLN disease (Lopez-Fabuel et al., 2022). Variations of mitochondrial mass could be examined through the use of Mitotracker (Sen et al 2024). In addition to this, there are several commercially available kits which could be used to examine pertinent factors of mitochondrial health. MitoProbe™ JC-1 Assay Kit (Thermo Fisher Scientific, Whitby, Ontario, Canada) is a kit which is used to examine membrane potential by which any variation from wild type would be indicative of mitochondrial disfunction. Along with this there are several other kits which would allow us to examine mitochondrial ROS levels which again would inform us about possible mitochondrial stress (Zorov et al 2014).

4.4 Rational for and possible future experiments examining autophagic pathways

Another possible avenue of examination is autophagy pathways, since comparative transcriptomics uncovered a list of genes relating to autophagy that were

dysregulated by the loss of *mfsd8* and as shown by our autophagy kit results, there was a decrease in autophagic activity. As such, it would be imperative to examine the protein products of these affected genes. One way of examining this would be through the use of western blotting which would allow us to examine whether our RNA results translate to the protein level and if there is a bottle neck for the processes of autophagy. Continuing, we should employ immunofluorescent microscopy examining Atg8/LC3, FIP200 and Atg16/Wipi2 which would give an indication on whether there are issues impacting autolysosome formation (Geng et al., 2008). If this is the case, it will give an important indication regarding the autophagic dysregulation known to impact cln7 disease.

4.5 Rational for and possible future experiments examining the proteome

A third area that should be examined is the proteome. As shown in chapter 4.2 several genes that were observed to be upregulated in mfsd8^{-} had decreased secretion with the inverse also being true, this raises questions to what is happening on the protein level. If there is a decrease in transcription and a simultaneous an increase in protein, there are many possible reasons for this. It could be indicative of the knockout detecting a surplus in the specific protein and a reducing transcript production as a result or it could reduction in miRNA allowing for an increase in the rate of protein production in the mutant compared to WT. In the first case this could uncover disruptions in protein degradation pathways that Mfsd8 interacts with while the other could indicate new post Transcription regulation pathways Mfsd8 is involved in (Hinas et

al., 2007). As a result of these novel possibilities, a protein level assessment of mfsd8^{-} during the early development should be carried out.

4.6 Conclusion

Loss of mfsd8 affects *D. discoideum* in a variety of ways including delays in aggregation, reduced cell-substrate adhesion, and aberrant secretion during starvation along with increased proliferation and cell size during growth. While the alteration in the transcriptome did not provide direct reasons for the growth phase phenotypes, my hypothesis is that the secretome and transcriptome underlie the phenotypical observations seen in *mfsd8*$^{-}$ *D. discoideum.* There were observed shifts in both the transcriptome and secretome of *D. discoideum* cells in response to *mfsd8* loss. Regarding reduced substrate adhesion, there were changes in actin and myosin along with other genes known to play roles in cell motility. For cell substrate adhesion in both, there were changes in the production release and activity of several lysosomal enzymes along with the transcription of genes like *sadA* known to contribute to the modification of substrate adhesion. Finally, regarding secretion, there is not only the shift in many genes related to secretion but also a change in the release of protein related to secretion. These datasets not only catalog the effects of *loss of mfsd8* but also provide avenues for further research such as examining the genes related to not only the phenotypes observed but also autophagy, mitochondrial function and protein production and hopefully new opportunities for examination in mammalian models and therapeutics for CLN7 disease.

5.0 References

Barra, J., Barrand, P., Blondelet, M.-H., & Brachet, P. (1980). PDSA, a gene involved in the production of active phosphodiesterase during starvation of *Dictyostelium discoideum* amoebae. Molecular and General Genetics MGG, 177(4), 607–613.

Bakthavatsalam, D., Brock, D. A., Nikravan, N. N., Houston, K. D., Hatton, R. D., & Gomer, R. H. (2008). The secreted *Dictyostelium* protein CFAD is a chalone. Journal of Cell Science, 121(15), 2473–2480.

Bakthavatsalam, D., & Gomer, R. H. (2010). The secreted proteome profile of developing *Dictyostelium discoideum* cells. PROTEOMICS, 10(13), 2556–2559.

Barth, C., Le, P., & Fisher, P. R. (2007). Mitochondrial Biology and disease in *Dictyostelium*. International Review of Cytology, 207–252.

Best, H. L., Clare, A. J., McDonald, K. O., Wicky, H. E., & Hughes, S. M. (2021). An altered secretome is an early marker of the pathogenesis of CLN6 Batten disease. Journal of Neurochemistry, 157(3), 764–780.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114–2120.

Boyle, E. I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J. M., et al. (2004). GO::TermFinder – open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics 20, 3710–3715.

Brandenstein, L., Schweizer, M., Sedlacik, J., Fiehler, J., & Storch, S. (2015). Lysosomal dysfunction and impaired autophagy in a novel mouse model deficient for the lysosomal membrane protein CLN7. Human Molecular Genetics, 25(4), 777–791.

Bukharova, Bukahrova, Weijer, Bosgraaf, Dormann, van Haastert & Weijer (2005) 'Paxillin is required for cell-substrate adhesion, cell sorting and slug migration during *Dictyostelium* development.' J Cell Sci 118:4295-310

Chida, J., Yamaguchi, H., Amagai, A., & Maeda, Y. (2004). The necessity of mitochondrial genome DNA for normal development of Dictyostelium cells. Journal of Cell Science, 117(15), 3141–3152.

Butz, E. S., Chandrachud, U., Mole, S. E., & Cotman, S. L. (2020). Moving towards a new era of genomics in the neuronal ceroid lipofuscinoses. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 1866(9), 165571.

Cárcel-Trullols, J., Kovács, A. D., & Pearce, D. A. (2015). Cell biology of the NCL proteins: What they do and don't do. Biochimica et Biophysica Acta, 1852(10 Pt B), 2242–2255.

Cesar, P. H., Braga, M. A., Trento, M. V., Menaldo, D. L., & Marcussi, S. (2019). Snake Venom Disintegrins: An overview of their interaction with integrins. Current Drug Targets, 20(4), 465–477.

Cotman, S. L., & Lefrancois, S. (2021). Cln3, at the crossroads of endocytic trafficking. Neuroscience Letters, 762, 136117.

Coukell, B., Li, Y., Moniakis, J., & Cameron, A. (2004). The Ca 2+ /calcineurin-regulated cup gene family in *Dictyostelium discoideum* and its possible involvement in development. Eukaryotic Cell, 3(1), 61–71.

Danyukova, T., Ariunbat, K., Thelen, M., Brocke-Ahmadinejad, N., Mole, S. E., & Storch, S. (2018). Loss of CLN7 results in depletion of soluble lysosomal proteins and impaired mtor reactivation. Human Molecular Genetics, 27(10), 1711–1722.

Das, A. M., Jolly, R. D., & Kohlschütter, A. (1999). Anomalies of mitochondrial ATP synthase regulation in four different types of neuronal ceroid lipofuscinosis. Molecular Genetics and Metabolism, 66(4), 349–355.

Dimond, R. L., Burns, R. A., & Jordan, K. B. (1981). Secretion of lysosomal enzymes in the cellular slime mold, *Dictyostelium discoideum*. Journal of Biological Chemistry, 256(13), 6565–6572.

Duran, Rahman, Colten & Brazill (2009) '*Dictyostelium discoideum* paxillin regulates actin-based processes.' Protist 160:221-32

Duran, J. M., Anjard, C., Stefan, C., Loomis, W. F., & Malhotra, V. (2010). Unconventional secretion of ACB1 is mediated by autophagosomes. Journal of Cell Biology, 188(4), 527–536.

Edgar, A. J., & Polak, J. M. (2000). Human homologues of yeast vacuolar protein sorting 29 and 35. Biochemical and Biophysical Research Communications, 277(3), 622–630.

Elleder, M., Kousi, M., Lehesjoki, A.E., Mole, S.E., Siintola, E., & Topcu, M. (2011) Chapter 11: CLN7. In: S. Mole, R. Williams, H. Goebel, (Eds), The Neuronal Ceroid Lipofuscinoses (Batten Disease) 2nd ed. (pp. 176-188). Oxford: OUP Oxford

Fey, P., Stephens, S., Titus, M. A., & Chisholm, R. L. (2002). Sada, a novel adhesion receptor in Dictyostelium. The Journal of Cell Biology, 159(6), 1109–1119.

Fey, P., Kowal, A. S., Gaudet, P., Pilcher, K. E., & Chisholm, R. L. (2007). Protocols for growth and development of *Dictyostelium discoideum*. Nature Protocols, 2(6), 1307–1316.

Fey, P., Dodson, R. J., Basu, S., Hartline, E. C., & Chisholm, R. L. (2019). DictyBase and the DICTY Stock Center (version 2.0) - a progress report. The International Journal of Developmental Biology, 63(8-9–10), 563–572.

Filaquier, A., Marin, P., Parmentier, M.-L., & Villeneuve, J. (2022). Roads and hubs of unconventional protein secretion. Current Opinion in Cell Biology, 75, 102072.

Franke, J., & Kessin, R. H. (1981). The cyclic nucleotide phosphodiesterase inhibitory protein of *Dictyostelium discoideum*. purification and characterization. Journal of Biological Chemistry, 256(14), 7628–7637.

Gomer, R. H., Yuen, I. S., & Firtel, R. A. (1991). A secreted $80 \times 10^3 M_r$ protein mediates sensing of cell density and the onset of development in *Dictyostelium*. Development, 112(1), 269–278.

Gruenheit, N., Baldwin, A., Stewart, B., Jaques, S., Keller, T., Parkinson, K., Salvidge, W., Baines, R., Brimson, C., Wolf, J. B., Chisholm, R., Harwood, A. J., & Thompson, C. R.

(2021). Mutant Resources for Functional Genomics in *Dictyostelium discoideum* using REMI-Seq Technology. BMC Biology, 19(1).

Haltia M. (2006). The neuronal ceroid-lipofuscinoses: from past to present. Biochimica et Biophysica Acta, 1762(10), 850–856.

Han, Chung, Wessels, Stephens, Titus, Soll & Firtel (2002) 'Requirement of a vasodilator-stimulated phosphoprotein family member for cell adhesion, the formation of filopodia, and chemotaxis in *Dictyostelium*.' J Biol Chem 277:49877-87

Hinas, A., Reimegård, J., Wagner, E. G., Nellen, W., Ambros, V. R., & Söderbom, F. (2007). The small RNA repertoire of *Dictyostelium discoideum* and its regulation by components of the rnai pathway. *Nucleic Acids Research*, *35*(20), 6714–6726.

Huber RJ. (2017). Loss of Cln3 impacts protein secretion in the social amoeba *Dictyostelium*. Cellular Signalling 35, 61-72.

Huber RJ, Mathavarajah S. (2019). Comparative transcriptomics reveals mechanisms underlying cln3-deficiency phenotypes in *Dictyostelium*. Cellular Signalling 58, 79-90.

Huber, R. J. (2020). Molecular Networking in the Neuronal Ceroid Lipofuscinoses: Insights from Mammalian Models and the Social Amoeba *Dictyostelium discoideum*. J. Biomed. Sci. 27, 64.

Huber RJ, Mathavarajah S, Yap SQ. (2020). Mfsd8 localizes to endocytic compartments and influences the secretion of Cln5 and cathepsin D in *Dictyostelium*. Cellular Signalling 70, 109572.

Iwan, K., Clayton, R., Mills, P., Csanyi, B., Gissen, P., Mole, S. E., Palmer, D. N., Mills, K., & Heywood, W. E. (2021). Urine proteomics analysis of patients with neuronal ceroid lipofuscinoses. iScience, 24(2), 102020.

Jain, R., Yuen, I. S., Taphouse, C. R., & Gomer, R. H. (1992). A density-sensing factor controls development in *Dictyostelium*. Genes & Development, 6(3), 390–400.

Jerka-Dziadosz, M., Jenkins, L. M., Nelsen, E. M., Williams, N. E., Jaeckel-Williams, R., & Frankel, J. (1995). Cellular polarity in ciliates: Persistence of global polarity in a disorganized mutant of Tetrahymena thermophila that disrupts cytoskeletal organization. Developmental Biology, 169(2), 644–661.

Journet, A., Klein, G., Brugière, S., Vandenbrouck, Y., Chapel, A., Kieffer, S., Bruley, C., Masselon, C., & Aubry, L. (2011). Investigating the macropinocytic proteome of

Dictyostelium amoebae by high-resolution mass spectrometry. PROTEOMICS, 12(2), 241–245.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver. D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D. (2021). Highly accurate protein structure prediction with alphafold. *Nature*, *596*(7873), 583–589.

Kibler, K., Svetz, J., Nguyen, T.-L., Shaw, C., & Shaulsky, G. (2003). A cell-adhesion pathway regulates intercellular communication during *Dictyostelium* development. Developmental Biology, 264(2), 506–521.

Kinseth, M. A., Anjard, C., Fuller, D., Guizzunti, G., Loomis, W. F., & Malhotra, V. (2007). The golgi-associated protein grasp is required for unconventional protein secretion during development. Cell, 130(3), 524–534.

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with Hisat2 and HISAT-genotype. Nature Biotechnology, 37(8), 907–

Kim WD, Wilson-Smillie MLDM, Thanabalasingam A, Lefrancois S, Cotman SL, Huber RJ. (2022). Autophagy in the neuronal ceroid lipofuscinoses (Batten disease). Frontiers in Cell and Developmental Biology 10, 812728. In Research Topic: Defective Macroautophagy in Organelle Turnover: From Basic Mechanisms to Human Disease.

Kim WD, Huber RJ. (2022). An altered transcriptome underlies cln5-deficiency phenotypes in *Dictyostelium discoideum*. Frontiers in Genetics 13, 1045738. In Research Topic: Neuronal Ceroid Lipofuscinosis: Molecular Genetics & Epigenetics.

König, T., & McBride, H. M. (2024). Mitochondrial-derived vesicles in metabolism, disease, and aging. Cell Metabolism, 36(1), 21–35.

Kousi, M., Siintola, E., Dvorakova, L., Vlaskova, H., Turnbull, J., Topcu, M., Yuksel, D., Gokben, S., Minassian, B. A., Elleder, M., Mole, S. E., & Lehesjoki, A.-E. (2009). Mutations in CLN7/MFSD8 are a common cause of variant late-infantile neuronal ceroid lipofuscinosis. Brain, 132(3), 810–819.

Li, G., Foote, C., Alexander, S., & Alexander, H. (2001). Sphingosine-1-phosphate lyase has a central role in the development of *Dictyostelium discoideum*. Development, 128(18), 3473–3483.

Loomis, W. F., Fuller, D., Gutierrez, E., Groisman, A., & Rappel, W. J. (2012). Innate non-specific cell substratum adhesion. PloS One, 7(8), e42033.

Lopez-Fabuel, I., Garcia-Macia, M., Buondelmonte, C., Burmistrova, O., Bonora, N., Alonso-Batan, P., Morant-Ferrando, B., Vicente-Gutierrez, C., Jimenez-Blasco, D., Quintana-Cabrera, R., Fernandez, E., Llop, J., Ramos-Cabrer, P., Sharaireh, A., Guevara-Ferrer, M., Fitzpatrick, L., Thompton, C. D., McKay, T. R., Storch, S., Bolaños, J. P. (2022). Aberrant upregulation of the glycolytic enzyme PFKFB3 in CLN7 neuronal ceroid lipofuscinosis. Nature Communications, 13(1).

Geng, J., Baba, M., Nair, U., & Klionsky, D. J. (2008). Quantitative analysis of autophagy-related protein stoichiometry by fluorescence microscopy. The Journal of Cell Biology, 182(1), 129–140.

Maruhn, D. (1976). Rapid colorimetric assay of β-galactosidase and N-acetyl-β-glucosaminidase in human urine. Clin. Chim. Acta. 73, 453–461.

Mathavarajah S, Flores A, Huber RJ. (2017). *Dictyostelium discoideum*: A model system for cell and developmental biology. Current Protocols Essential Laboratory Techniques 15, 14.1.1-

Mesquita, A., Cardenal-Muñoz, E., Dominguez, E., Muñoz-Braceras, S., Nuñez-Corcuera, B., Phillips, B. A., Tábara, L. C., Xiong, Q., Coria, R., Eichinger, L., Golstein, P., King, J. S., Soldati, T., Vincent, O., & Escalante, R. (2016). Autophagy in Dictyostelium: Mechanisms, regulation and disease in a simple biomedical model. Autophagy, 13(1), 24–40.

Morita, T., Saitoh, K., Takagi, T., & Maeda, Y. (2000). Involvement of the glucose-regulated protein 94 (DD-GRP94) in starvation response of *Dictyostelium discoideum* cells. Biochemical and Biophysical Research Communications, 274(2), 323–331.

Mole, S. E., & Cotman, S. L. (2015). Genetics of the neuronal ceroid lipofuscinoses (Batten disease). Biochimica et Biophysica Acta, 1852(10 Pt B), 2237–2241.

Moore, S. J., Buckley, D. J., MacMillan, A., Marshall, H. D., Steele, L., Ray, P. N., Nawaz, Z., Baskin, B., Frecker, M., Carr, S. M., Ives, E., & Parfrey, P. S. (2008). The clinical and genetic epidemiology of neuronal ceroid lipofuscinosis in Newfoundland. Clinical Genetics, 74(3), 213–222.

Muller-Taubenberger, A. (2001). Calreticulin and calnexin in the endoplasmic reticulum are important for phagocytosis. The EMBO Journal, 20(23), 6772–6782.

Musial, J., Niewiarowski, S., Rucinski, B., Stewart, G. J., Cook, J. J., Williams, J. A., & Edmunds, L. H. (1990). Inhibition of platelet adhesion to surfaces of extracorporeal

circuits by disintegrins. RGD-containing peptides from Viper Venoms. Circulation, 82(1), 261–273.

Noh, S. H., Kim, Y. J., & Lee, M. G. (2022). Autophagy-related pathways in vesicular unconventional protein secretion. Frontiers in Cell and Developmental Biology, 10.

Palmer, D. N., Fearnley, I. M., Walker, J. E., Hall, N. A., Lake, B. D., Wolfe, L. S., Haltia, M., Martinus, R. D., & Jolly, R. D. (1992). Mitochondrial ATP synthase subunit c storage in the ceroid-lipofuscinoses (Batten disease). American Journal of Medical Genetics, 42(4), 561–567.

Paravicini, G., Horazdovsky, B. F., & Emr, S. D. (1992). Alternative pathways for the sorting of soluble vacuolar proteins in yeast: A vps35 null mutant missorts and secretes only a subset of vacuolar hydrolases. Molecular Biology of the Cell, 3(4), 415–427.

Pollitt, A. Y., & Insall, R. H. (2008). Abi Mutants in Dictyostelium reveal specific roles for the scar/wave complex in cytokinesis. Current Biology, 18(3), 203–210.

Prassler, J., Stocker, S., Marriott, G., Heidecker, M., Kellermann, J., & Gerisch, G. (1997). Interaction of a Dictyostelium member of the plastin/fimbrin family with actin filaments and actin-myosin complexes. Molecular Biology of the Cell, 8(1), 83–95.

Remtulla AAN, Huber RJ. (2023). The conserved cellular roles of CLN proteins: Novel insights from *Dictyostelium discoideum*. European Journal of Cell Biology 102(2), 151305. In Special Issue: Cell Biology's Leading Edge - Part 2.

Roisin-Bouffay, C. (2000). A precise group size in Dictyostelium is generated by a cell-counting factor modulating cell–cell adhesion. Molecular Cell, 6(4), 953–959.

Rosenberg, J. B., Chen, A., Kaminsky, S. M., Crystal, R. G., & Sondhi, D. (2019). Advances in the treatment of neuronal ceroid lipofuscinosis. Expert Opinion on Orphan Drugs, 7(11), 473–500.

Rossomando, E. F., Maldonado, B., Crean, E. V., & Kollar, E. J. (1978). Protease secretion during onset of development in *Dictyostelium discoideum*. Journal of Cell Science, 30(1), 305–318.

Rot, G., Parikh, A., Curk, T., Kuspa, A., Shaulsky, G., & Zupan, B. (2009). DictyExpress: A *Dictyostelium discoideum* gene expression database with an explorative data analysis web-based interface. BMC Bioinformatics, 10(1).

Seaman, M. N. J., Marcusson, E. G., Cereghino, J. L., & Emr, S. D. (1997). Endosome to golgi retrieval of the vacuolar protein sorting receptor, Vps10p, requires the function of the vps29, vps30, and vps35 gene products. The Journal of Cell Biology, 137(1), 79–92.

Sen, M. G., Sanislav, O., Fisher, P. R., & Annesley, S. J. (2024). The multifaceted interactions of Dictyostelium ATG1 with mitochondrial function, endocytosis, growth, and development. Cells, 13(14), 1191.

Schleicher, M., Noegel, A., Schwarz, T., Raff, E. W., Brink, M., Faix, J., Geriscii, G., & Isenberg, G. (1988). A *Dictyostelium* mutant with severe defects in α-actinin: Its characterization using cdna probes and monoclonal antibodies. Journal of Cell Science, 90(1), 59–71.

Schulz, A., Kohlschütter, A., Mink, J., Simonati, A., and Williams, R. (2013). NCL diseases — Clinical perspectives. Biochim. Biophys. Acta 1832, 1801–1806.

Schulz, A., Ajayi, T., Specchio, N., de Los Reyes, E., Gissen, P., Ballon, D., Dyke, J. P., Cahan, H., Slasor, P., Jacoby, D., Kohlschütter, A., & CLN2 Study Group (2018). Study of intraventricular cerliponase alfa for CLN2 Disease. The New England Journal of Medicine, 378(20), 1898–1907.

Siintola, E., Topcu, M., Aula, N., Lohi, H., Minassian, B. A., Paterson, A. D., Liu, X. Q., Wilson, C., Lahtinen, U., Anttonen, A. K., & Lehesjoki, A. E. (2007). The novel neuronal ceroid lipofuscinosis gene MFSD8 encodes a putative lysosomal transporter. American Journal of Human Genetics, 81(1), 136–146.

Springer, Cooper & Barondes (1984) 'Discoidin I is implicated in cell-substratum attachment and ordered cell migration of *Dictyostelium discoideum* and resembles fibronectin.' Cell 39:557-64

Stajdohar M, Rosengarten RD, Kokosar J, Jeran L, Blenkus D, Shaulsky G, Zupan B. (2017) "dictyExpress: a web-based platform for sequence data management and analytics in Dictyostelium and beyond." BMC Bioinformatics. 18(1):291

Steenhuis, P., Herder, S., Gelis, S., Braulke, T., & Storch, S. (2010). Lysosomal targeting of the CLN7 membrane glycoprotein and transport via the plasma membrane require a Dileucine motif. Traffic, 11(7), 987–1000.

Steenhuis, P., Froemming, J., Reinheckel, T., & Storch, S. (2012). Proteolytic cleavage of the disease-related lysosomal membrane glycoprotein CLN7. Biochimica et Biophysica Acta, 1822(10), 1617–1628.

Sucgang, Weijer, Siegert, Franke & Kessin (1997) 'Null mutations of the *Dictyostelium* cyclic nucleotide phosphodiesterase gene block chemotactic cell movement in developing aggregates.

Sharifi, A., Kousi, M., Sagné, C., Bellenchi, G. C., Morel, L., Darmon, M., Hůlková, H., Ruivo, R., Debacker, C., El Mestikawy, S., Elleder, M., Lehesjoki, A.-E., Jalanko, A., Gasnier,

B., & Kyttälä, A. (2010). Expression and lysosomal targeting of CLN7, a major facilitator superfamily transporter associated with variant late-infantile neuronal ceroid lipofuscinosis. Human Molecular Genetics, 19(22), 4497–4514.

Tang, L., Ammann, R., Gao, T., & Gomer, R. H. (2001). A cell number-counting factor regulates group size in Dictyostelium by differentially modulating cAMP-induced cAMP and cGMP pulse sizes. The Journal of biological chemistry, 276(29), 27663–27669.

Tarantola, M., Bae, A., Fuller, D., Bodenschatz, E., Rappel, W. J., & Loomis, W. F. (2014). Cell substratum adhesion during early development of *Dictyostelium discoideum*. PloS One, 9(9), e106574.

Tatischeff, I. (2019). *Dictyostelium*: A model for studying the extracellular vesicle messengers involved in human health and disease. Cells, 8(3), 225.

Tatischeff, I., Larquet, E., Falcón-Pérez, J. M., Turpin, P., & Kruglik, S. G. (2012). Fast characterisation of cell-derived extracellular vesicles by nanoparticles tracking analysis, Cryo-electron microscopy, and Raman tweezers microspectroscopy. Journal of Extracellular Vesicles, 1(1).

Varet, H., Brillet-Guéguen, L., Coppée, J.-Y., & Dillies, M.-A. (2016). SARTools: A deseq2- and edger-based R pipeline for comprehensive differential analysis of RNA-Seq Data. PLOS ONE, 11(6).

Varadi, M., Bertoni, D., Magana, P., Paramval, U., Pidruchna, I., Radhakrishnan, M., Tsenkov, M., Nair, S., Mirdita, M., Yeo, J., Kovalevskiy, O., Tunyasuvunakool, K., Laydon, A., Žídek, A., Tomlinson, H., Hariharan, D., Abrahamson, J., Green, T., Jumper, J., Birney, E., Steinegger, M., Hassabis, D., Velankar, S. (2023). Alphafold protein structure database in 2024: Providing structure coverage for over 214 million protein sequences. *Nucleic Acids Research*, *52*(D1).

Viotti, C. (2016). ER to golgi-dependent protein secretion: The conventional pathway. Methods in Molecular Biology, 3–29.

Vinet, A. F., Fiedler, T., Studer, V., Froquet, R., Dardel, A., Cosson, P., & Pieters, J. (2014). Initiation of multicellular differentiation in *Dictyostelium discoideum* regulated by Coronin A. Molecular Biology of the Cell, 25(5), 688–701.

von Kleist, L., Ariunbat, K., Braren, I., Stauber, T., Storch, S., & Danyukova, T. (2019). A newly generated neuronal cell model of cln7 disease reveals aberrant lysosome motility and impaired cell survival. Molecular Genetics and Metabolism, 126(2), 196–205.

Walkley, S. U., March, P. A., Schroeder, C. E., Wurzelmann, S., & Jolly, R. D. (1995). Pathogenesis of brain dysfunction in Batten disease. American Journal of Medical Genetics, 57(2), 196–203.

Wang, Y., Zeng, W., Lin, B., Yao, Y., Li, C., Hu, W., Wu, H., Huang, J., Zhang, M., Xue, T., Ren, D., Qu, L., & Cang, C. (2021). Cln7 is an organellar chloride channel regulating lysosomal function. Science Advances, 7(51).

Westphal, M., Jungbluth, A., Heidecker, M., Mühlbauer, B., Heizer, C., Schwartz, J.-M., Marriott, G., & Gerisch, G. (1997). Microfilament Dynamics during cell movement and Chemotaxis monitored using a GFP–actin fusion protein. Current Biology, 7(3), 176–183.

Wimmer, B., Lottspeich, F., Ritter, J., and Bronnenmeier, K. (1997). A novel type of thermostable alpha-D-glucosidase from Thermoanaerobacter thermohydrosulfuricus exhibiting maltodextrinohydrolase activity. Biochem. J. 328, 581–586.

Williams, R.E. (2011) NCL incidence and prevalence data. In: S. Mole, R. Williams, H, Goebel, (Eds), The Neuronal Ceroid Lipofuscinoses (Batten Disease) 2nd ed (pp. 361-365). Oxford: OUP Oxford

Wu, Franke, Blanton, Podgorski & Kessin (1995) 'The phosphodiesterase secreted by prestalk cells is necessary for *Dictyostelium* morphogenesis.

Yap SQ, Kim WD, Huber RJ. (2022). Mfsd8 modulates growth and the early stages of multicellular development in *Dictyostelium discoideum*. Frontiers in Cell and Developmental Biology 10, 930235. In Research Topic: Ion Transporters and Channels in Cellular Pathophysiology.

Zorov, D. B., Juhaszova, M., & Sollott, S. J. (2014). Mitochondrial reactive oxygen species (ROS) and Ros-induced Ros release. Physiological Reviews, 94(3), 909–950.