Exploring the function of mammalian Adenosine Deaminase 2, a novel anti-inflammatory glycoprotein of the macrophage lysosome

Ebba Straume Loennechen

Master Thesis in Medical Technology

Department of Chemistry / Biomedicine
University of Bergen

June, 2024

Scientific environment

This thesis is part of Professor Ole Kristian Greiner-Tollersrud's research on ADA2 and DADA2, primarily carried out at the Department of Biochemistry at The Arctic University of Norway (UiT). Greiner-Tollersrud is part of the lysosomal research group, focusing mainly on glycobiology.

The thesis was also carried out at the Department of Biomedicine at The University of Bergen (UiB).

Parts of the study were funded by the Grethe Harbitz Trust, a trust for research on rheumatic diseases.

UiT The Arctic University of Norway

UNIVERSITY OF BERGEN

Acknowledgements

Firstly, I would like to thank my supervisors, Knut Teigen and Ole Kristian Greiner-Tollersurd. I was fortunate to join Greiner-Tollersrud's research on DADA2 despite it taking place at The Arctic University of Norway in Tromsø. I was also fortunate to have a short stay in Tromsø, where all experiments were conducted. There, I got valuable guidance from my supervisor and the person responsible for the very important MS analysis, Jack-Ansgar Bruun.

In Bergen, the Department of Biomedicine at the University of Bergen welcomed me with open arms. From my supervisor, Knut Teigen, I received valuable feedback throughout the whole writing process. I also want to thank both of my supervisors for having patience and taking the time over the past 1.5 years to have consistent meetings and guide me through small and large obstacles. I also thank the Department of Chemistry for allowing me to write my thesis with UiT and the Department of Biomedicine here in Bergen, which allowed me to contribute to research on such a disease.

Lastly, thank you to the other students at Medical Technology for your continued motivation and support during these past five years at UiB. And, thank you to my family and friends for your support the past five years, it has been invaluable!

Ebba Straume Loennechen
Bergen, June 2024

Abstract

Deficiency of Adenosine Deaminase 2 (DADA2) is an autosomal recessive autoinflammatory disease currently standing without sufficient treatment, except in a few mild cases where bone marrow transplantation was possible. The main reason for the lack of targeted therapy is the lack of understanding the protein's function, resulting in only symptomatic, anti-inflammatory treatments. Despite great efforts from all researchers, existing theories on the function have all fallen short. This thesis aims to contribute to the research, which can, ultimately, be used to develop a therapy for DADA2 patients.

The predominant hypothesis on the molecular function of ADA2 is that it is a secreted N-glycosylated protein produced in monocytes/macrophages and hydrolyzes plasmic Adenosine (Ado) to Inosine (Ino). As Ado is a proinflammatory molecule, ADA2 would function as an anti-inflammatory protein. The low affinity of ADA2 for Ado raises the question of whether ADA2 hydrolyzes plasmic Ado or if there is some other proinflammatory molecule ADA2 regulates. A novel theory considering this, is that ADA2 is a lysosomal protein regulating DNA at acidic pH. DNA, like Ado, is highly pro-inflammatory and could also cause the symptoms seen in DADA2 patients.

Large evolutionary differences are visible, such as a complete gene deletion from rodents and large variations in affinity for Ado. For example, where chicken has a low K_m=0.1mM, human ADA2 has a very high K_m=2.5mM. To gain more insight into the differences between species, ADA2 from chicken liver and sheep spleen was partially purified and analyzed with MS/MS. Enzyme assays testing heat stability and specific activity were also conducted. The MS/MS analysis revealed that sheep spleen had significantly less ADA2 than previously purified pig spleen, and there were also surprisingly low amounts in chicken liver. Unfortunately, the MS/MS did not provide results to determine any glycan structures. Yet, the results indicate that sheep ADA2 no longer has the monocyte/macrophage-specific promoter seen in pigs and humans. The enzyme assays showed that both chicken and sheep ADA2 are highly heat-stable proteins, indicating a stable protein aligning with a lysosomal function. However, where

sheep ADA2's glycosylation near the active site might hinder any DNA-binding activity, chicken ADA2 might have this property. On the other hand, chicken ADA2 displays a high affinity for free Ado, suggesting that hydrolysis of free Ado might be its main function.

In conclusion, this thesis' results suggest an evolutionary drift in ADA2's function, from hydrolysis of free Ado in invertebrates to the larger substrate DNA in mammals. Some mammals, such as pigs and humans, where ADA2 is promoted in monocytes/-macrophages, might be attributed to ADA2 functioning as a regulator of lysosomal DNA in highly stressed macrophages and avoiding critical DNA concentrations causing inflammation. More research is required to understand how an ADA2 deficiency in humans results in inflammation due to the accumulation of either Ado or DNA. Overall, the research aims to improve the lives of patients with DADA2, and understanding the function and pathomechanism of DADA2 will allow for the development of a therapy.

Contents

Abbreviations

0.1 Abbreviations

ADA1	Adenosine deaminase 1
ADA2	Adenosine deaminase 2
ADGF	Adenosine deaminase growth factor
Ado	Adenosine
AID	Autoinflammatory disease
cDNA	Complementary DNA
CECR1	Cat eye syndrome critical region candidate 1
cGAMP	Cyclic GMAMP
cGAS	Cyclic GMP-AMP synthase
CID	Collision-induced ionization
CNS	Central nervous system
CM	Carboxymethyl
Con A	Concanavalin A
DADA1	Deficiency of adenosine deaminase 1
DADA2	Deficiency of adenosine deaminase 2
dAdo	2'-deoxyadenosine
DAMP	Damage-associated molecular patterns
dATP	Deoxyadenosine triphosphate
dIno	2'-deoxyinosine
DPP IV	Dipeptidyl peptidase IV
Dol-P	Dolichol-phosphate
dsDNA	Double-stranded DNA

DTT	Dithiothreitol
EHNA	Erythro-9-(2-hydroxy-3-nonyl)-adenine hydrochloride
EI	Electron impact ionization
EndoH	Endoglycosidase H
ER	Endoplasmatic Reticulum
ESI	Electrospray ionization
Fuc	L-fucose
FUC8	α1-6-fycosyltransferase
Gal	D-galactose
GANAB	α-glucosaide II
GBP	Glycan-binding protein
GC	Gas chromatography
Glc	D-glucose
GlcA	D-glucuroic acid
GlcNAc	N-acetyl-D-glucoseamine
GPI	Glycosylphosphatidylinositol
Hyp	Hypoxanthine
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
Ino	Inosine
IRF7	Interferon regulatory factor 7
LacNAc	N-acetyllactosamine
LMAN	Lysosomal α-mannosidase
LRR	Leucine-rich repeats
LSD	Lysosomal storage disease
M6P	Mannose-6-phosphate
MAN	α-mannosidase
Man	D-mannose
MANB1	ER α-mannosidase I
MGAT	N-acetylglucosaminyltransferase
MOGS	α-glucosaide I
MS-MS	Tandem mass spectroscopy
MyD88	Myeloid differentiation primary-response protein 88

NAFLD	Nonalcoholic fatty liver disease
NAGPA	Phosphodiester N-acetylglucoseamindase
NBT	Nitro blue tetrazolium
NET	Neutrophil extracellular traps
OST	Oligosaccharyltransferase
PAMP	Pathogen-associated molecular patterns
PAN	Polyarteritics nodosa
PNGaseF	Peptide N-glycosidase F
PNP	Purine nucleoside phosphorylase
PRB	Putative receptor-binding
UiB	University of Bergen
UiT	The Arctic University of Norway
RD	Rheumatic disease
SAID	Systemic autoinflammatory disease
SCID	Severe combined immunodeficiency
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
ssDNA	Single-stranded DNA
TIR	Toll-interleukin-1 receptor
TLR	Toll-like receptor
TNF	Tumor necrosis factor
TNFRS1A	Tumor necrosis factor receptor superfamily member 1A
TRAPS	Tumor necrosis factor receptor-associated periodic syndrome
Xa	Xanthine
XO	Xanthine oxidase
Xyl	D-xylose

0.2 Amino acids

Name	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	R
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Aim of the study

To this day, patients with a deficiency of the protein Adenosine Deaminase 2 (ADA2) are left without sufficient targeted treatment. For the past five decades, since the discovery of ADA2, researchers have tried to find the protein's function without success. Not knowing the protein's function is the main reason for the lack of targeted treatment, which is the obvious goal as the medicinal world strives for personalized, specific treatments.

Deficiency of ADA2 (DADA2) is an inherited, monogenic, autoinflammatory disease causing systemic symptoms in the patients. Specifically, this includes recurring fevers, early stroke, vasculitis, and bone marrow failure (1). The severity of the symptoms varies between patients, but all have the foundation of systemic inflammation. There are only around 600 confirmed patients, but as the protein's function is unknown, it is difficult to diagnose, and it is speculated that the actual number of patients is up to 35 000 (1).

The underlying theory of research on ADA2 is that it is a secreted glycoprotein that hydrolyzes extracellular Adenosine (Ado) and deoxyadenosine (dAdo). However, all presented theories have some aspects that cannot be explained, debunking the theory. This thesis will investigate the protein using a different model, namely that mammalian ADA2 is mainly sorted to the lysosomes of monocytes/macrophages, where it has an anti-inflammatory role by interacting with DNA (2).

This theory will be investigated through MS/MS analysis on ADA2 partially purified from chicken liver and sheep spleen. The model may only apply to some mammals

since rodents lack the ADA2 gene. Chicken, a non-mammalian vertebrate, and sheep were chosen as sheep ADA2 have a predicted N-glycan in the DNA-binding groove, which may prevent interaction between DNA and ADA2. The focus of the results from this analysis will be on the glycan structures present in the protein, as well as the abundance. In addition, an extensive literature search will be done to find similarities and dissimilarities to other lysosomal enzymes and endonucleases and their respective pathological conditions when deficient.

The long-term goal of this work is to find the optimal treatment for DADA2 patients. To do this, one first has to know the physiological function of ADA2 to determine its pathomechanism and, from there, find the optimal way to tackle the disease. This is the goal of every researcher investigating ADA2, and hopefully, this thesis will provide valuable insight.

It is also worth mentioning the strong driving force of the The DADA2 Foundation, an organization formed by DADA2 patients and their family members (3). The organization has played an important role in keeping a focus on the ADA2 research and creating a collaborative platform for families, academia, pharma companies, and non-profit organizations.

1.2 Biological background

1.2.1 Inflammation

Inflammation is the physiological response in vascular tissue to infections and cell or tissue damage (4), and it is a crucial part of a functional immune system. Common causes of inflammation are viral, bacterial, or fungal infections, tissue necrosis, foreign bodies, and immune reactions. Immune reactions include all kinds of hypersensitivity, e.g. allergies, but also autoimmune and autoinflammatory diseases.

Inflammation is divided into acute and chronic inflammation; the two types have different characteristics. Acute inflammation is the first-line response to any stress and begins minutes or hours after the initial stress. At a cellular level, neutrophils are the main cell type. Many "visible" changes also occur, such as edema and redness. Chronic inflammation, on the other hand, is characterized by monocytes/macrophages and lymphocytes and has fewer "visible" signs. Chronic inflammation also causes more severe tissue fibrosis due to an equilibrium of constant tissue damage and tissue repair (4).

There are multiple steps in an inflammatory reaction, and different inflammation pathways are activated as the initiating step varies. The first step can be a cell receptor or antibody recognizing, e.g., a foreign body. One example is the cell receptor family of Toll-like receptors (TLRs), which recognize motifs present in many microbes. These motifs are often referred to as PAMPs, pathogen-associated molecular patterns (4). Other similar receptor families also exist, though they have different pathways. TLRs, when activated, induce the production and secretion of cytokines, interferons, and other membrane proteins that cause acute inflammation and lymphocyte activation. Another initiating step is the activation of cytosolic sensors for cell damage. These sensors recognize molecules and ions related to damaging changes in the cell and endo- and exogenous DNA, to mention a few. Free DNA in the cytosol is highly pro-inflammatory and causes the expression of interferons due to binding to cytosolic receptors, such as cGAS, further discussed later in this section. Interferon production is part of the anti-viral defense, which induces interferon-stimulatory genes. Concerning ADA2, endosomal TLR9, cGAS, and the potential leakage of DNA into the cytosol are discussed, as well as how the lack of ADA2's potential functions causes autoinflammation in humans.

Toll-like receptor 9

TLRs are transmembrane glycoproteins that recognize PAMPs and damage-associated molecular patterns (DAMPs) (5). The TLRs, 10, are confirmed in humans and are differentially expressed on cell and tissue types depending on what ligand they are specified for and which process they initiate (6). Nonetheless, common for all TLRs is that they initiate an immune response by recognizing a pattern inherently linked to something pathogenic.

The TLRs are able to recognize PAMPs and DAMPs as pathogenic sequences that have been conserved in the receptors' genes through evolution. This also shows that the TLRs are a part of the *innate* (born) immune system as they do not have any adaptability in their recognition mechanisms. This will be further discussed (5).

All TLRs consist of three domains: A "recognition" domain, also known as the ectodomain, a transmembrane domain, and an interacting domain responsible for downstream signaling, known as the Toll-interleukin-1 receptor (TIR) (6). Crystallization studies have shown that they form dimers, hetero- or homodimers, depending on

the TLR, through the ectodomain. As stated, they are type I transmembrane glycoproteins with multiple leucine-rich repeats (LRR), usually pictured as a "hook." The presence of an LRR domain is another indication that TLRs are, in fact, pattern-recognizing receptors, and it is this domain that interacts with PAMPs and DAMPs. The different TLRs are located on the cell surface, with an extracellular LRR domain, or an intracellular LRR domain facing the inside of the organelle (Figure 1.1) (5). The TIR domain, or endodomain, binds an adapter protein when a TLR binds a ligand that initiates the downstream signaling. The different TLRs utilize multiple adapter proteins, and their purpose is to signal further and activate the intended response (6).

From an evolutionary perspective, TLRs are mainly meant to participate in the innate immune system and protect against foreign substances. It is important to know that some TLRs also initiate this response to endogenous PAMPs and DAMPs. Specifically, TLRs 2, 3, 4, 7, 8, and 9 have ligands originating from the self (5). For example, TLR2 and 4 are specific to DAMPs related to heat damage, i.e., motifs in heat-shock proteins. In the case of DADA2, TLR9 is the relevant receptor as it has self-double-stranded DNA (dsDNA) as its primary endogenous ligand. However, it also recognizes CpG motifs in microbial DNA (5; 6). TLR9 has some ability to differentiate between self- and nonself-ligands as it favors binding to unmethylated CpG motifs. As human DNA is partially methylated, while microbial and bacterial DNA is unmethylated, there is some differenatiation.

TLR9 functions as a homodimer and is located on the endosome, lysosome, and endoplasmatic reticulum (ER) membrane in cells that initiate an immune response through type I IFN and cytokine production (5). TLR3, 7, and 8 are also in this group, but while TLR9 recognizes viral and bacterial CpG DNA, the other three recognize RNA. In the event that TLR9 recognizes its ligand, it activates the IFN regulatory factor to induce type I IFN production. The regulatory factor activation completely depends on the adapter protein *Myeloid differentiation primary-response protein 88* (MyD88). MyD88 induces a cascade reaction leading to the translocation of interferon regulatory factor 7 (IRF7) to the nucleus, which induces type I IFN production (7).

The concept of TLR9 recognizing self-dsDNA and initiating an immune response is the basis for some theories regarding the pathogenesis of DADA2, which is presented in Section 1.3.5.

[1] Illustration by Sameer and Nissar for "Toll-Like Receptors (TLRs): Structure, Functions, Signaling, and Role of Their Polymorphisms in Colorectal Cancer Susceptibility", licensed under a Creative Commons Attribution License (accessed January 22nd, 2024) (6)

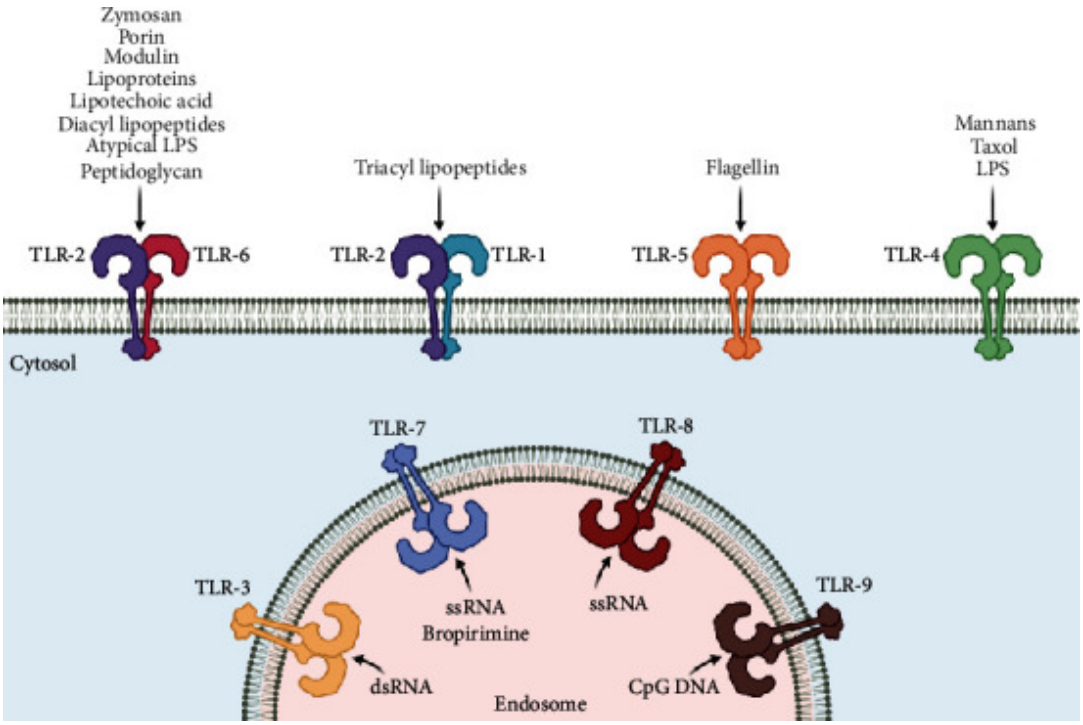


Figure 1.1: Location of nine of the Toll-like receptors. Credit[1]

The STING-pathway

Cyclic GMP-AMP synthase (cGAS) is a cytosolic sensor for dsDNA. When it binds dsDNA, it forms a 2:2 complex. The complex formation causes a conformational change in the active seat, catalyzing cyclic guanosine monophosphate–adenosine monophosphate (cGAMP) synthesis. cGAMP is a second messenger that activates STING and is bound to the ER membrane. Activated STING is transferred from the ER membrane to an ER-Golgi intermediate compartment on the Golgi apparatus. During this mobilization, STING recruits TBK1, which phosphorylates and activates IRF3. IRF3 is a transcription factor for type I IFN, and when phosphorylated, it dimerizes and enters the nucleus, stimulating type I IFN production. STING also activates kinase IKK, inhibiting the transcription factor NF-κB and inhibiting interferon production. (8). In other words, both STING and TLR9 activation induce type I IFN production.

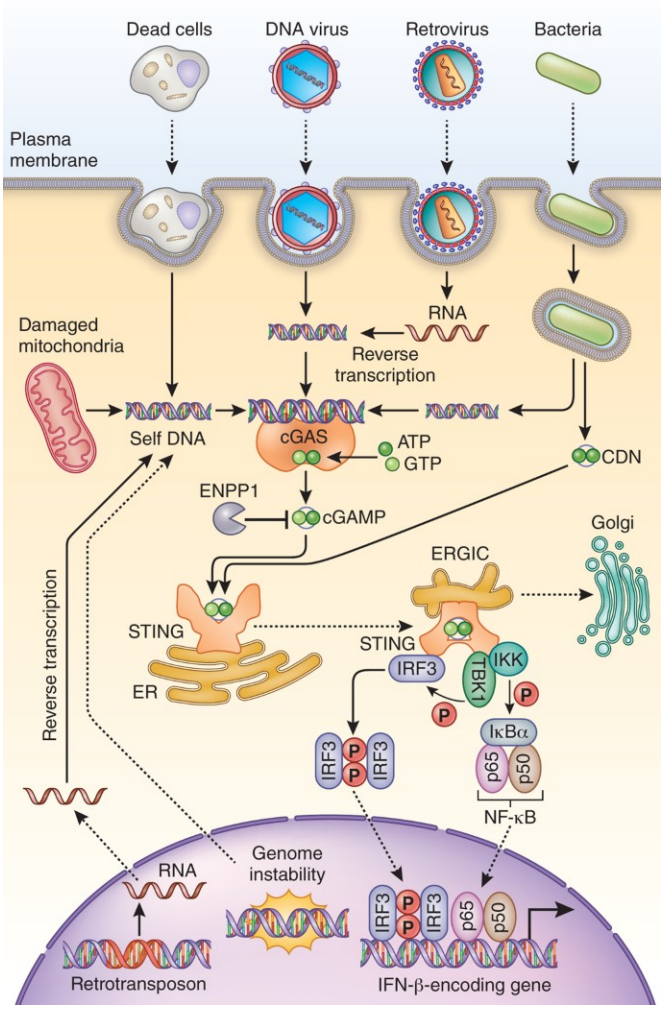


Figure 1.2: Activation of cGAS-STING pathway by self, viral, and bacterial dsDNA. Credit[2]

DNA-binding occurs between the sugar-phosphate bond in dsDNA and cGAS, not the base residues. This is why there is no differentiation between self, viral, and bacterial dsDNA. Nonetheless, there is some selectivity towards structures where the backbone of dsDNA is more available as the bases can be steric hindrances (8).

The cGAS-STING pathway is a part of the innate immune system, as TLRs, and activates the transcription and production of pro-inflammatory type I IFN. Unlike other innate immune pathways activated by PAMPs or other molecular patterns, e.g., TLRs, the cGAS-STING pathway is activated by dsDNA, independent of the sequence. As stated, there is no differentiation between pathogenic dsDNA and self-dsDNA. Suppose there is an accumulation of self-dsDNA due to some pathological condition, it may cause an autoinflammatory condition due to activation of type I IFN, unless the accumulated DNA is sufficiently dealt with by another compensating mechanism (8).

[2]Illustration by Debbie Maizels for "Regulation and function of the cGAS–STING pathway of cytosolic DNA sensing", licensed under a Springer Nature License (accessed November 10th, 2023) (8)

This is seen in some autoimmune diseases, such as lupus, where leakage of dsDNA from the nucleus and mitochondria causes an over-transcription of type I IFN and other inflammatory molecules (9). However, there is also evidence of a steady state activation of interferons due to mutations in STING (8).

A protein that deals with the accumulation of DNA in the cytosol and regulates cytosolic DNA sensors, such as cGAS, is TREX1 (DNaseIII) (10; 11). TREX1 regulates the cGAS-STING pathway through the metabolism of cytosolic DNA, removing the substrate that activates the pathway and reducing type I IFN production. Hence, dysfunction of TREX1 leading to an accumulation of DNA causes an overactivation of cGAS, resulting in inflammation and autoimmune disease (11). TREX1 is not specific for a type of DNA, meaning it metabolizes complementary DNA (cDNA), single-stranded DNA (ssDNA), and dsDNA from, e.g., dead cells, cancer cells, damaged mitochondria, or the nucleus (Figure 1.2) (12). dsDNA is TREX1's main substrate, though there is evidence it degrades cDNA and ssDNA as well (12; 13).

Another possible source for cytosolic DNA is leakage from the lysosomal membrane (13). Kawane et al. (2014) proposed a mechanism for the leakage of dsDNA, resulting in cGAS-activation due to TREX1 being oversaturated (Figure 1.3). TREX1, like most enzymes, has a limit to how much substrate it can degrade before it is unable to properly regulate the substrate concentration. The mechanism is related to DNase type II (DNaseII) deficient macrophages, a lysosomal DNase (further discussed in Section 1.2.3). Still, deficiencies of other enzymes related to lysosomal DNA metabolism are expected to have a similar outcome. The theory behind the leaking membrane is rather new and is based on spontaneous pore formation (14). The pores are first formed as a hydrophobic pre-pore, then a hydrophilic pore as the hydrophilic "head" of the phospholipid translocates. The topic of spontaneous pores and membrane leakage is still new and not yet fully understood. However, it likely explains the over-activation of cGAs and autoinflammation. This theory is not the only one looking at "leaking" membranes. Additional theories are looking at how an accumulation of some macromolecules induces more frequent membrane pores and/or breakage, leading to content leakage. One example is Terje Espevik's, from the Norwegian University of Science and Technology (NTNU), research on atherosclerosis and on how cholesterol crystals cause lysosomal membrane damage and rupture, leading to inflammation and release of interleukin-1 β (15; 16).

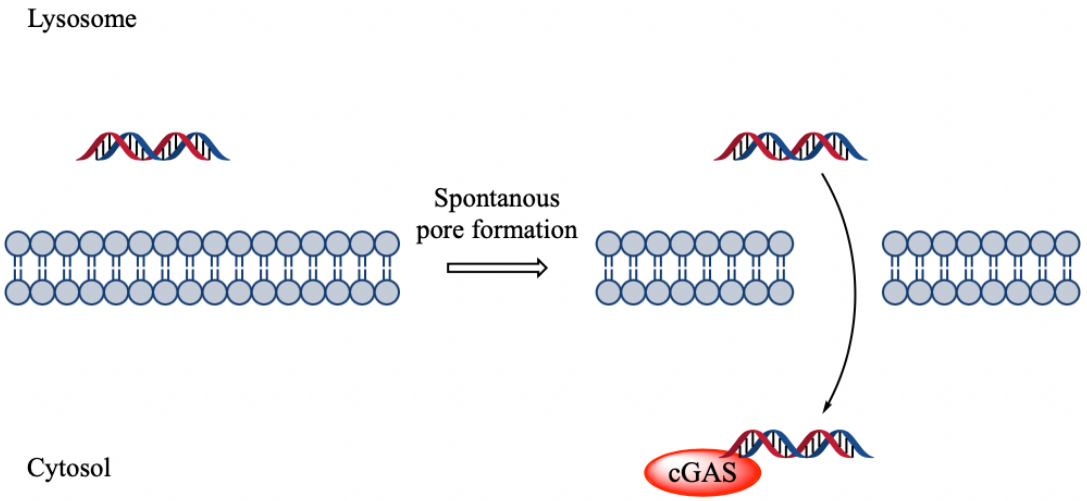


Figure 1.3: Spontaneous pore formation in the lysosomal membrane leads to dsDNA leakage, activating cGAS-STING. Credit[3]

1.2.2 Autoinflammatory diseases

The National Institute of Arthritis and Musculoskeletal and Skin Diseases defines autoinflammatory diseases (AIDs) as "problems with the innate immune system's reaction" (17). Even though this definition can seem rather vague, it brings the main characteristic to light: the innate immune system based on leukocytes, not antibodies (which are part of the adaptive immune system). Dysfunction or dysregulation of the innate immune system leads to systemic inflammation (18). Systemic inflammation refers to inflammation throughout the body rather than a specific organ. The cells associated with AIDs are monocytes, macrophages, and neutrophils. The systemic inflammation is mediated by these cell types through interferons, tumor necrosis factor (TNF), and interleukins - the inflammatory cytokines induced with TLR9 or cGAS-STING activation, along with other mechanisms not discussed in this thesis.

Systemic AIDs (SAIDs) are divided into two groups based on the genetic background of the disease: Monogenic and polygenic (18). Most monogenic diseases follow a clear pattern of dominant inheritance of the gene, making it easier to pinpoint the exact gene that causes the disease and determine the pathogenic mechanism. On the other hand, polygenic SAIDs arise from a combination of mutations and normal gene types, where a single gene or mutation by itself is not enough to be pathogenic. This means that the genetic pattern of two patients with the same polygenic SAID can differ. Hence, determining the pathogenesis of these diseases can be difficult. It is also important to be aware of the environmental factors that affect the presentation and severity of

[3]Illustration made by the candidate in ChemDraw, inspired by (13; 14)

polygenic SAIDs (18).

Additionally, there is a distinction amongst monogenic SAIDs, the initial effect of a mutation. A mutation can cause 1) increased activity in pattern-recognizing receptors or their adapter molecules, e.g., TLR9 and MyD88, 2) decreased activity in cells maintaining cell homeostasis, 3) decreased downregulation of proinflammatory responses, or 4) increased immune receptor signaling (18). The second group is the most relevant in this thesis as DADA2 is monogenic and refers to the total or partial loss of ADA2 activity. An accumulation of adenosine, ADA2s substrate, or any other substrate that stimulates and activates intracellular sensors/receptors upregulates the production of inflammatory mediators. Which mediator is produced depends on the mutated gene and substrate. *Tumor Necrosis Factor Receptor-Associated periodic syndrome* (TRAPS) is a good example of such a SAID. It is monogenic and caused by a mutation in the gene *tumor necrosis factor receptor superfamily member 1A* (TNFRS1A). The mutation causes dysregulation of cysteine-cysteine disulfide bonds, leading to misfolded receptors that accumulate in the cytoplasm. Accumulation leads to enhanced NF-κB activation (19).

Since different proinflammatory mediators are present in different SAIDs, the symptoms can vary. However, some general characteristic symptoms are recurrent fevers, abdominal pain, arthritis (rheumatism), vasculitis, and cutaneous signs (20). Vasculitis, sometimes called vasculopathy, is the inflammation of blood and lymph vessels and is part of causing the more visible cutaneous signs (21; 22). Livedo reticularis and polyarteritis nodosa-like symptoms (PAN) are common conditions, especially in DADA2 cases (22).

Some SAIDs also fall into the category of *rheumatic diseases* (RDs). RD, commonly known as arthritis, is a general term for any condition that affects joints and soft tissue (23; 24). These symptoms can also be caused by autoimmune diseases, but for this thesis, autoinflammatory rheumatic diseases will be discussed. Previously mentioned TRAPS is also an example of a monogenic autoinflammatory RD (25). DADA2 is another example where systemic inflammation also affects the joints and causes rheumatic symptoms (Section 1.4).

1.2.3 The lysosome and lysosomal storage disorders

The lysosome is, in many ways, the cell's garbage disposal, but it has an important role in immune responses, cell signaling, and metabolism (26). The organelle consists of enzymes contributing to the degradation of macromolecules, such as hydrolases, but also membrane proteins which are critical for the lysosome's stability and acidity. The composition of the different hydrolases and membrane proteins dictates the specific function of the lysosome and varies between cell types. Ions are also a critical component as they regulate the pH, which is crucial for the lysosomal function (26).

The lysosomal enzymes responsible for the degradation are mainly hydrolases, e.g., proteases, nucleases, and glycosidases. A common factor for these types of enzymes is having an acidic pH optimum of approximately 5 (26). Proteases process, modify, and degrade proteins and peptides by cleaving peptide bonds through hydrolysis (27). Nucleases are separated into DNases and RNases, as they degrade either DNA or RNA through hydrolysis of the phosphodiester bond between nucleotides (28; 29). Glycosidases degrade glycans through hydrolysis of the glycosidic bonds, and these are important in modifying and trimming glycan structures (Section 1.2.4). There is also an important cooperation between, e.g., glucosidases and proteases. Glycan structures can be steric hindrances for the proteases, thereby inhibiting the enzymes' ability to cleave the peptide bonds (30). When glycosidases trim the glycan structures, they allow the glycoproteins to be degraded by removing or reducing the hindrance eventually. Hence, a dysfunction or deficiency of a glycosidase can cause glycoprotein and glycan accumulation in the lysosomes and cause a *lysosomal storage disorder* (LSD). Some examples of this are discussed later in this section.

Other than degrading through hydrolysis, lysosomal enzymes also share a low pH optimum of approximately 5 (26). The lysosomes have an acidic environment regulated by ions, including Ca^{2+}, Na^{+}, K^{+}, Zn^{2+}, H^{+}, Fe^{2+}, and Cl^{-}. Some also have other purposes in the lysosomes in addition to pH regulation. For example, the presence of Zn^{2+} and Fe^{2+} are important as they are crucial for the function of some enzymes, e.g., ADA2 has Zn^{2+} as a cofactor (Section 1.3). The Ca^{2+}-concentration is significantly higher in the lysosomes than in the cytosol and is important for signal transduction, lysosomal homeostasis, and acidification. Ca^{2+} and H^{+} also function as a pair as Ca^{2+}/H^{+}-exchangers to keep the Ca^{2+}-gradient. By itself, H^{+} is kept at a steady concentration through an ATP-driven pump to keep the luminal pH at 4.6 (26). To balance the H^{+}-concentration and avoid a too high positive charge in the lysosomes, a complementary Cl^{-}-pump is also present. Additionally, pumps for the outflux of K^{+} and Na^{+}

will have the same purpose as a the influx of Cl^-, i.e., keeping the charge of the lysosomes while upholding a low pH.

Lysosomal Storage Disorders

A dysfunction or deficiency of any of the lysosomal enzymes can easily lead to an accumulation of macromolecules due to the vast amounts that reach the lysosomes. An accumulation can, in turn, lead to cancer, cardiovascular disease, metabolic disorder, or chronic inflammation (26). These diseases have the "umbrella" term LSDs and are genetic. As with SAIDs, the pathogenesis of LSDs can vary significantly depending on which lysosomal enzyme is dysfunctional or deficient and which macromolecule accumulates. The phenotype of LSDs is mainly physical, meaning inflammation, respiratory failure, or anemia, but it can cause neurological symptoms as seen in Parkinson's and Alzheimer's disease (26).

DNaseII, mentioned in Section 1.2.1, is an endonuclease that fragments and degrades exogenous and endogenous DNA originating from phagocytosis and apoptosis (programmed cell death) (31). As seen in mice and humans, DNaseII deficiency can result in DNA accumulation, leading to autoinflammatory disorders. "Free" self-DNA functions as nucleic antigens and causes an immune reaction, resulting in the accumulated DNA activating unwanted immune responses (31; 32). A double knock-out of DNaseII in mice lung cells indicated that the accumulated DNA activates the cGAS-STING pathway (Section 1.2.1), which induces type I IFN production and inflammation (32; 33). In other words, DNase II deficiency falls into the category of SAID (Section 1.2.2), and the clinical phenotype aligns with this.

Another example of an LSD is a deficiency of the exoglycosidase α-*mannosidase* (MANB), α-mannosidosis (34). It is important to specify that it is a deficiency of the lysosomal MAN (MANB), as there are also other MAN in the ER/Golgi-Network (35). There are two kinds of MANB, where the major type can cleave any mannose (Man) in the core glycan (Man9-1) (Section 1.2.4), starting at the terminal end, including α-1,2, α-1,3, and α-1,6 linkages (35). The second MANB is specified for the α-1,6 linkage between Man3-2. There are no cases of a deficiency of the specified MANB, but a deficiency of the "general" MANB can result in an accumulation of partially degraded oligosaccharides and glycopeptides. Patients have different symptoms, both types and severity, ranging from skeletal abnormalities and poor muscle control to large effects on the central nervous system (CNS), causing early death. Different severity of mental

retardation is also seen (36). The enzyme replacement therapy (ERT), *Lamzede*, developed by Chiesi Farmaceutici S.p.A. was approved by the European Medicines Agency in 2018 (37). Related to DADA2, a similar therapy is likely suitable.

1.2.4 Glycobiology

Glycans are monosaccharides and oligosaccharides, commonly referred to as sugars and sugar chains. They are crucial in supporting the assembly of complex organisms and interactions in these organisms. Glycans form bonds with proteins and lipids, creating a vast diversity in proteins and lipids despite originating from only a few genes. Monosaccharides, the fundamental building blocks of all glycans, can form complex structures due to the inherent properties of carbohydrates. Unlike amino acids, which can only form linear peptides, the possible glycan structures are determined by the different monosaccharides and how they can bind (38).

There are different groups of glycans, including N-linked, O-linked, glycosaminoglycans, glycosylation of lipids, and glycosylphosphatidylinositol (GPI) (39). As there are different types of glycans, it is necessary to specify that from here on when it is referred to glycans, it is N-linked glycans which are meant. N-glycans is also a term which will be used. Single glycans will referred to as monosaccharides or glycans, depending on the context. N-glycans in mammals are assembled by only four types of monosaccharides (38; 40):

- Hexoses: Six-carbon neutral sugars - D-glucose (Glc), D-galactose (Gal), and D-mannose (Man)
- Hexosamines: Hexoses with an amino group at the 2-position, which is commonly acetylated - N-acetyl-D-glucosamine (GlcNAc)
- 6-Deoxyhexoses: L-fucose (Fuc)
- Sialic acids: N-Acetyl-Neuraminic Acid (Neu5Ac) and N-Glycolyl-Neuraminic Acid (Neu5Gc)

Many other monosaccharides not mentioned here are also part of these groups and are present in mammalian N-glycans. Additionally, uronic acids can be found in mammalian N-glycans, though they are rare. Lastly, pentoses are found N-glycans in vertebrates (41).

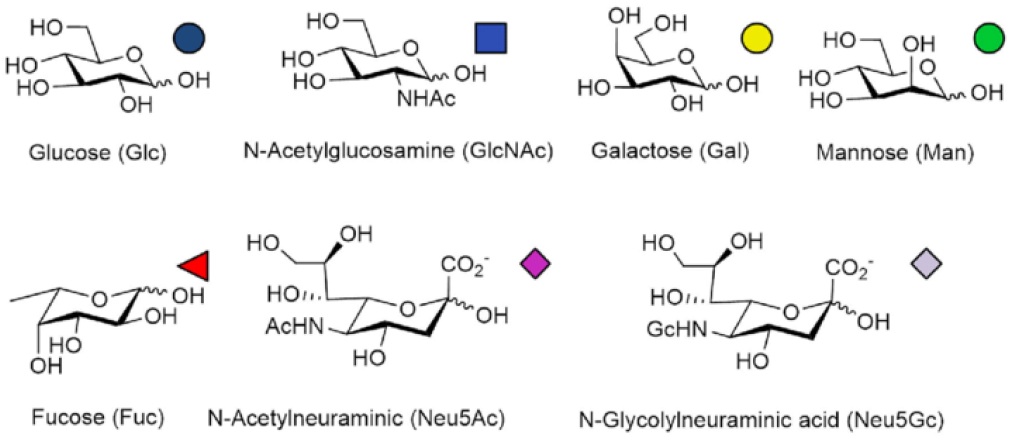


Figure 1.4: The most common monosaccharides in N-glycans, with generalized color coding for structure visualization. Credit[4].

Unlike a protein sequence determined by the genome, glycans are post-translational modifications. There are specific enzymes for synthesizing and trimming glycans, but these do not directly dictate which structures will be synthesized. Therefore, the glycans present on a glycoprotein are not as easily predicted despite being able to predict the possible structures. In addition, a glycoprotein with, e.g., five glycosylation sites might only have four of these glycosylated, and these four glycans can differ from protein to protein. In other words, there is a heterogeneity between glycoproteins, making it difficult to have a complete qualitative overview of a glycoprotein's glycans. The quantification of a glycoprotein is also significantly more complex than with an unglycosylated protein (38).

The biosynthesis of N-glycans

The complete biosynthesis is visualized in Figure 1.6.

In eukaryote cells, glycans are synthesized and assembled in the ER and Golgi apparatus. In general, secreted proteins are glycosylated, and their glycosylation follows the final stages of folding, which also occurs in the ER and Golgi apparatus. The proteins are transported from rough ER to intermediate compartments in the Golgi apparatus. The glycosylation processes differ between the types of glycans, but only N-glycans will be discussed. In the case of N-glycans, a precursor structure is assembled before it is attached to the protein. The precursor is then bound to the protein in rough ER before the complex is transported to the Golgi apparatus. There, further modifications

[4]Illustration by Chao et al. in 2020 for "Recent Progress in Chemo-Enzymatic Methods for the Synthesis of N-Glycans," licensed under a CC-BY License (accessed January 26th, 2024) (42)

take place, both addition and trimming. Even though the glycans are matured inside the ER and Golgi, the glycan structures withstand even further modification at their location of function. N-glycosylated proteins either have an extracellular or lysosomal function. The vastly different conditions in different extracellular spaces and the acidic lysosomes cause high diversity in glycan structures (39).

O-linked glycans, glycosaminoglycans, and glycosylation of lipids differ from N-linked glycans in assembly, while the GPI is similar. While N-linked glycans and GPI anchors are assembled before they are bound to the proteins, O-linked glycans, glycosaminoglycans, and glycans bound to lipids have a more step-like assembly. Despite these differences, there are some similarities regarding modifications (39).

N-glycans can covalently bind to an asparagine (N) in a protein when it is followed by the sequence *asparagine - X - serine/threonine* (NXT/S). This short motif, where X is any given amino acid, is called a glycosylation site. GlcNAc forms a covalent N-glycosidic bond to the residue on N. Following the first GlcNAc, all N-glycans share the same core of five glycans: GlcNAc-GlcNAc-Man-Man-Man (GlcNAc2Man3), where the two final Man bind to the first Man. The core is the basis for all kinds of N-glycans, and there are three main groups: 1) High-mannose, 2) hybrid, and 3) complex. How the different ones arise will be further explained (41).

The synthesis of the precursor oligosaccharide begins in the cytoplasm, outside the ER. *Dolichol phosphate* (Dol-P), which is bound in the ER membrane, binds to a P-GlcNAc provided by the complex UDP-GlcNAc. A second GlcNAc is attached, making Dol-P-P-GlcNAc2. Five Man are attached, provided by GDP-Man, to the terminal GlcNAc, giving Dol-P-P-GlcNAc2-Man5. An enzyme of unknown mechanism, commonly known as "flippase," flips the structure from the cytoplasm to the inside of the ER lumen. There, an additional four Man and three Glc are added, finalizing the precursor: Dol-P-P-GlcNAc2-Man9-Glc3 (Figure 1.5) (41).

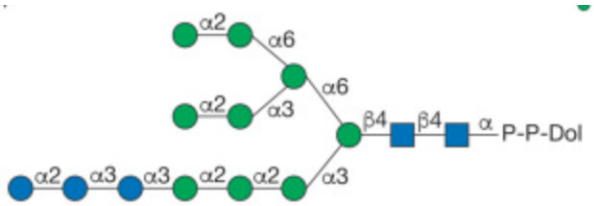


Figure 1.5: The precursor oligosaccharide before attachment to a polypeptide. Credit[5].

[5]Figure adapted from illustration 9.3 in "Essentials of Glycobiology", licensed under a Creative Commons

Inside the ER lumen, *oligosaccharyltransferase 9* (OST9) recognizes an NXT/S motif on a nascent polypeptide about to be folded. The energy released from cleaving the P-GlcNAc bond - as the monosaccharides are provided as Dol-P-Monosaccharadie - is then used to bind the precursor to N. Following the binding, the glycan is processed at the same time as the polypeptide is being folded. In other words, the glycan processing functions as a quality control of the folding and will, ideally, be ongoing until the protein is correctly folded. The processing consists of removing the three Glc by α*-glucosaide* I (MOGS) and II (GANAB). MOGS removes the terminal Glc, while GANAB removes the inner two. The inner Glc is removed and re-added until the protein is correctly folded, and its final removal signals that the protein is "ready." Before a transfer from the ER lumen to the *cis*-Golgi, α*-mannosidase I* (MANB1) removes the terminal α1-2Man from the central arm (41).

In *cis*-Golgi, the degree of processing dictates whether a glycan matures as a high-mannose or will be further processed to a hybrid or complex glycan. MANIA, IB, and IC remove Man yielding GlcNAc2-Man5. This smaller structure is the basis for hybrid and complex N-glycans. Some glycans escape the processing of MANIA, IB, and IC and, consequently, further modifications, leaving mature glycans with a simple GlcNAc2-Man5-9. The future hybrid and complex glycans are translocated to the *medial*-Golgi (41).

N-acetylglucosaminyltranferase 1 (MGAT1) adds a GlcNAc to C-2 on α1-3Man. Only if MGAT1 first added the GlcNAc can MANIIA and IIB trim the two terminal Man. Further, MGAT2 adds a second GlcNAc to α1-6Man, giving the base structure for biantennary, complex N-glycans: GlcNAc2-Man3-GlcNAc2. In the absence of MANIIA and IB processing, only MGAT1 processing, the base for hybrid N-glycans, is made: GlcNAc2-Man5-GlcNAc. For both hybrid and complex, further processing by the addition of a Fuc, by α*1-6-fycosyltransferase* (FUC8), to the N-linked GlcNAc can occur (41).

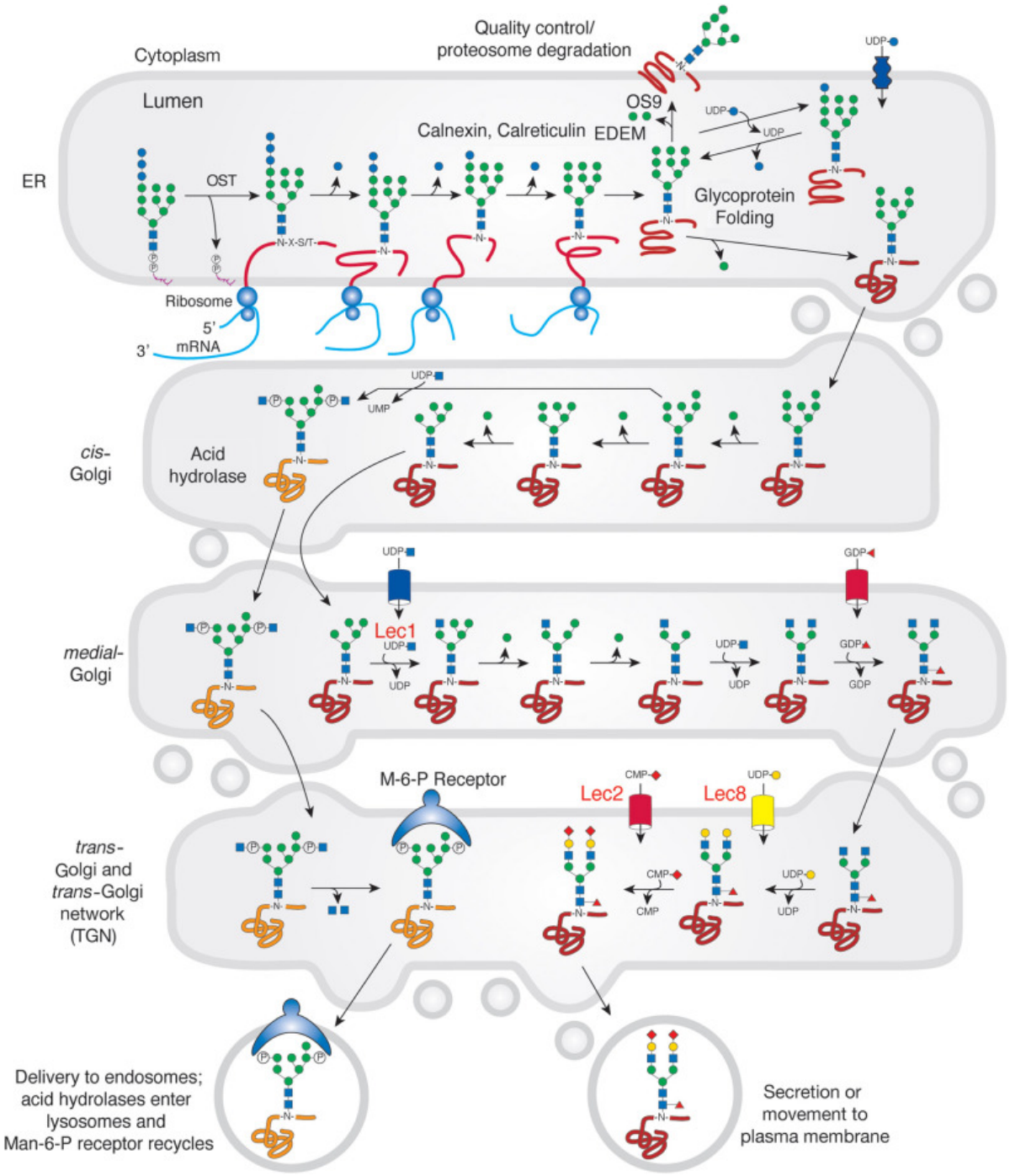


Figure 1.6: The biosynthesis of N-linked glycans, starting at the precursor oligosaccharide being flipped to the ER lumen, to the final maturation in the *trans*-Golgi. Lastly, the transportation of phosphorylated glycans to the lysosome via the M6P receptor, or secretion Credit[6].

Complex N-glycans can be extended to tri-, tetra-, or polyantennary by MGAT3-5, creating a vast amount of different complex N-glycans. MGAT3 binds a GlcNAc to βMan attached to the inner two GlcNAcs, creating the substrate for MGAT4A and 4B, and MGAT5 and 5B. MGAT4A/B and 5/5B attach GlcNAc to, respectively, α1-4Man

[6]Unedited illustration 9.4 from "Essentials of Glycobiology", licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Unported license (accessed January 16th, 2024) (41)

and α1-6Man. These two enzymes can also process the MGAT2 product directly before MGAT3 processing, creating substrates for one another. Finally, in birds and fish, not humans, glycans processed by MGAT3, 4A/B, and 5/5B can be processed by MGAT6 (41).

All branches added in the *medial*-Golgi can be extended, mainly through three "methods": 1) addition of single monosaccharides, 2) extension of branching GlcNAc, and 3) "capping" of branches. The first method applies to the core glycan and is the addition of Fuc to the N-linked GlcNAc, as previously explained. The Fuc is not further extended. The second method, extension of the branches, includes the addition of monosaccharides, e.g., Gal, to the terminal GlcNAc. The addition of Gal yields the unit *N-acetyllactosamine* (LacNAc) = Galβ1-4GlcNAc, commonly seen repeated on branches. The third and final method is "capping." Adding Sia, Fuc, Gal, GlcNAc, or sulfate groups, usually α-linked, elongates the branches. These α-linked monosaccharides are facing away from any repeating LacNAc-units, as the Gal creating the unit is β-linked. Figure 1.7 below shows some examples of extended complex N-glycans (41).

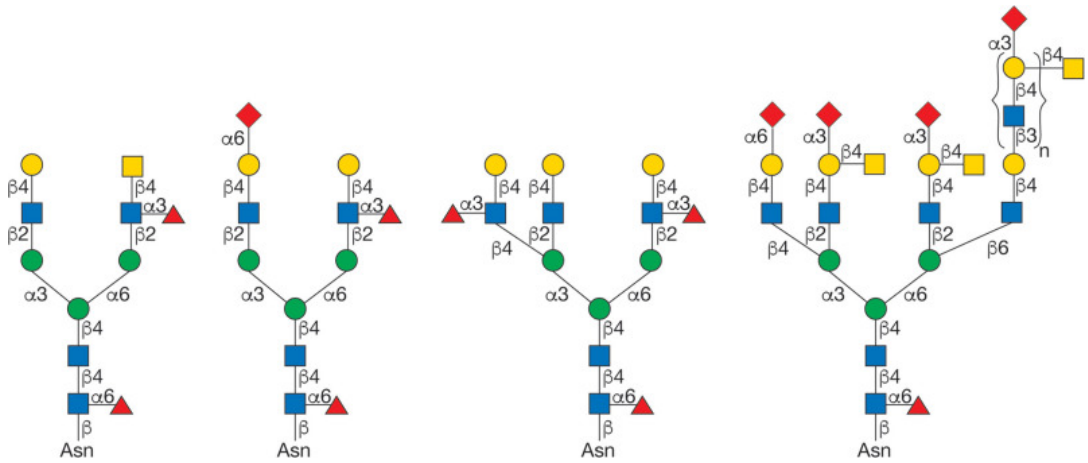


Figure 1.7: Examples of extended hybrid and complex N-glycans. Credit[7].

As seen in Figure 1.6, an alternative maturation process starts in the *cis*-Golgi. This process is applied to lysosomal hydrolases. Lysosomal hydrolases are transported to the lysosome via a pathway that requires phosphorylated high-mannose N-glycans. Before translocation to the *medial*-golgi, a single Man is removed as normal, but instead of continued removal of Man, phosphorylated GlcNAc is added to the C-6 of Man. No further processing occurs until it reaches the *trans*-Golgi, where the two newly added GlcNAcs are removed, leaving the phosphate group, by a *phosphodiester*

[7]Unedited illustration 9.6 from "Essentials of Glycobiology", licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Unported license (accessed January 16th, 2024) (41)

N-acetylglucosamindase (NAGPA). The Mannose-6-phosphate (M6P) receptor recognizes the structure, GlcNAc2-Man6-2P, and transports the glycoprotein to an acidic endosome which fuses with the lysosome (41; 43).

The function and purpose of N-glycans

At the beginning of this section, it was stated that glycans are important for assembling complex organisms and the interaction between cells. More specifically, N-glycans are important in keeping growth factors and cytokine receptors at cell surfaces and mediate cell interactions to transport leukocytes to sites of inflammation and tissue damage (41). In addition to these "complex" functions, the N-glycans are also crucial for initial protein folding in the ER. While the focus of the previous section was the synthesis of the glycans themselves, the synthesis occurs in parallel to the folding of the protein it is bound to, and any mistakes in the glycan synthesis can cause misfolded peptides (44).

N-glycans and other glycans have regulatory functions and apply these through interaction with *glycan-binding proteins* (GBPs). The GBPs are how the glycans communicate their desired effect and outcome (45). This feedback applies to many human processes, including inflammation processes initiated by both endogenous and exogenous pathogens. The C-type lectin family, proteins that recognize glycans, are an example of how glycans regulate and initiate inflammation. Dectin-1 is a transmembrane, C-type lectin that recognizes β-linked glucose polymers and recruits Syk-kinase, which leads to NF-κB activation and cytokine and chemokine production (similar to that of the STING/cGAS pathway, Section 1.2.1 (45). Dectin-2 is a similar C-type lectin, but unlike Dectin-1, which has a cytoplasmic immunoreceptor for signaling, Dectin-2 uses an adapter protein. Dectin-2 is a better example of differentiation between pathogenic glycans and self N-glycans, as N-glycans typically are not large structures. Pathogenic glycans tend to be larger structures, which Dectin-2 recognizes and initiates an immune response (45). As with other inflammation-related proteins, Dectin-1 and -2 are mainly in monocytes/macrophages, neutrophils, and other myeloid cells. It is important to state that there is a vast number of proteins and glycan-receptors related to inflammation that will not be discussed in this thesis.

Tomi N-glycans

A review article by Greiner-Tollersrud, Venkatakrishnan, Wuhrer, and G. Karlsson (2024), currently submitted in for review, introduces a new term, *tomi* ("cut" in Greek),

for N-glycans that do not "fit" into the existing groups (hybrid, complex, high-mannose) and their composition cannot be explained by the synthesis route (46). The biosynthesis explained in this section shows that there is a specific way the monosaccharides can be assembled. Even though this results in a vast number of different N-glycans, some N-glycans that are found can only occur from post-synthesis catabolism. The article presents three routes for the catabolism of N-glycans but also specifies that this new group of *tomi* N-glycans does not include every kind of N-glycan present in humans. The three possible pathways introduced are:

- Further trimming in the lysosome (Figure 1.8A)
- Trimming by extracellular exoglycosidase (proposed mechanism for select cancers) (Figure 1.8B)
- Trimming during granulopoiesis of neutrophils, sorting N-glycans into mannosidase-containing azurophilic granules (Figure 1.8C)

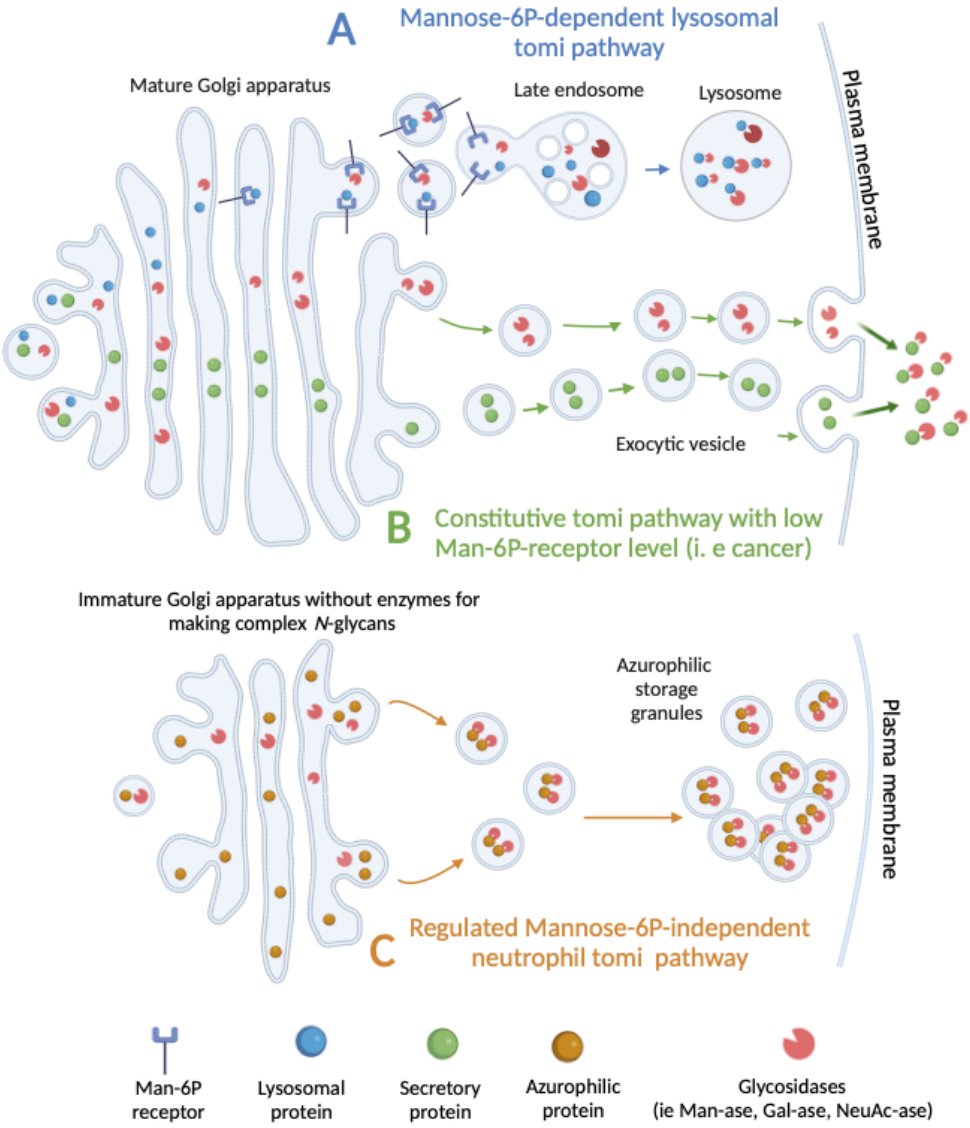


Figure 1.8: The three pathways for creation of *tomi* N-glycans. Credit[8].

Pathway A is the most relevant in this thesis on ADA2 and DADA2. Acidic exoglycosidases in the lysosomes are thought to metabolize bound and free N-glycans. During the initial discovery of these types of N-glycans, GlcNAc2-Fuc1-Man2 being a prevalent structure, they were thought to be a result of incomplete synthesis. Over time, it has been found that they originate from trimming in the lysosomes. Patients with a deficiency of the glycosidase β-galactosidase had higher levels of lysosomal glycoproteins with terminal β-galactose's than a control group, proving the presence of galactose trimming (47). Additionally, α-mannosidose is proven to have lysosomal trimming through mouse models. Healthy mice have GlcNAc2-Man2-Fuc1 as the main structure, while mice with α-mannosidosis (Section 1.2.3) have GlcNAc2-Man3-Fuc1

[8]Illustration by Greiner-Tollersrud et al. in 2024 for submitted review article (46)

as the main structure (48).

The origin of GlcNAc2-Fuc1-Man2 and other high-mannose N-glycans has been argued to be the processing of FUC8 on GlcNAc2-Man5. Studies have shown that GlcNAc2-Man5 is not a substrate for FUC8, rather, its substrates are complex and hybrid N-glycans. This shows that, e.g., GlcNAc2-Fuc1-Man2 is a product of trimming and is a *tomi* N-glycan.

1.3 Adenosine deaminase 2

Adenosine deaminase 2 (ADA2), historically referred to as *adenosine deaminase growth factor* (ADGF), is presumed to catalyze the extracellular hydrolysis of the signaling molecule *adenosine* (Ado) to *inosine* (Ino). ADA2 is mainly secreted from monocytes/macrophages and encoded by the gene *cat eye syndrome critical region candidate 1* (CECR1), part of an ADGF family, which is coupled to a macrophage/monocyte-specific promoter. It is secreted as a 110 kDa homodimer, and dimerization is critical for full enzymatic activity. As a monomer, it weighs approximately 57 kDa (49; 50).

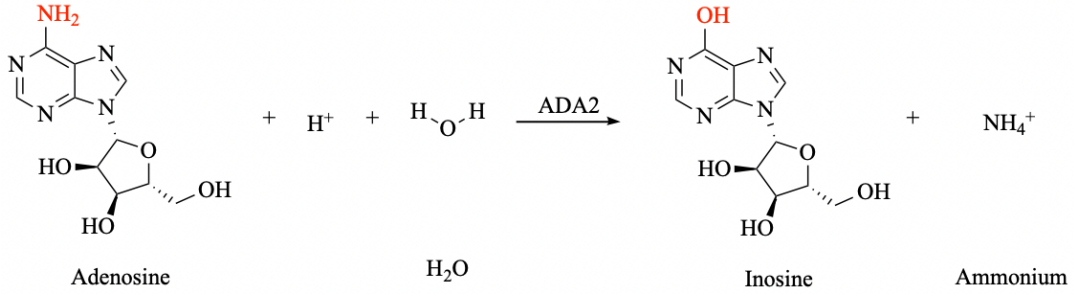


Figure 1.9: Catalytic activity of ADA2, hydrolysis of adenosine. Credit[9]

Adenosine deaminase 1 (ADA1), another ADA isoenzyme of 41 kDa, has the same catalytic activity as presumed for ADA2, hydrolysis of Ado to Ino, though intracellularly. Unlike ADA2, which is mainly in monocytes/macrophages, ADA1 is present in all cell types. ADA1 can function as an intracellular monomer but also as a membrane-bound heterodimer, dimerizing with either *dipeptidyl peptidase IV* (DPP IV) or CD26 (50).

[9]Illustration made by the candidate in ChemDraw

This section will review the significant differences between ADA1 and ADA2, the present theories for ADA2's function, and how ADA2 differs between species.

1.3.1 Adenosine

During cell stress and injury, e.g., inflammation, ischemia, or hypoxia, Ado is released, and extracellular levels increase. A review by Haskó et al. (2008) on the function of Ado found that it functions as a signaling molecule for maintaining homeostasis and initiating immune responses, both directly and indirectly. Directly, Ado binds to four transmembrane receptors: A_1, A_{2A}, A_{2B}, and A_3 which, through varying mechanisms, induce interleukin production and decrease inflammation. The abundance of the receptors varies between cell types, receptor A_{2A} being the most abundant in macrophages. Appropriate A_{2A} activation leads to increased interleukin production and reduced production of pro-inflammatory TNF-α. It is important to note that the EC_{50} (half maximal effective concentration) for A_1, A_{2A}, and A_3 is 0.01-1 μM, while A_{2B} has an EC_{50} of 24 μM. A_{2B} is therefore only activated under pathological conditions, as the physiological concentration of adenosine is below 1 μM. Activation of A_{2B} has a much more pro-inflammatory effect than A_{2A} activation, as A_{2B} does not down-regulate TNF-α, but increases interleukin production in a variety of cells, including IL6 (induces inflammation through T-cells++) and IL10 (stimulates cytokine production in macrophages) from macrophages (51).

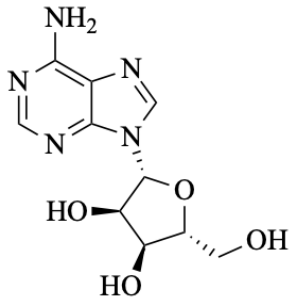


Figure 1.10: Adenosine. Credit[10]

Even though adenosine has a crucial anti-inflammatory function, it can have a pro-inflammatory effect in high concentrations. Borea (2017) found that the accumulation of adenosine can inhibit macrophages and other immune-related cells, as well as up-regulation of the A_{2B}-receptor. Increased activation of A_{2B} has been associated with increased production of IL17, causing pulmonary fibrosis and chronic lung injury, to

[10]Illustration made by the candidate in ChemDraw

name a few (52). In addition, Liu (2015) found that increased activation of A_{2B} and A_{2A} causes both dermal and hepatic fibrosis and brain damage (53). In sum, the intended anti-inflammatory effect of Ado can be highly inflammatory if the Ado concentration is too high.

1.3.2 The structure and properties of ADA2

Table 1.1 presents the most significant differences between ADA1 and ADA2. One is the inhibition by EHNA. Multiple studies have found that EHNA inhibits ADA1, while ADA2 is resistant, and this being consistent across species(22; 50; 54). Another key difference is the localization, where ADA1 is intracellular, and ADA2 is (assumed) extracellular and secreted. As ADA1 is present in all cells, ADA2 is in and secreted from activated monocytes/macrophages (50; 55). This might suggest that ADA2 has an activity related to inflammation, which is consistent with the elevated ADA2 levels found in some diseases (56).

Another large difference is their K_m-value for free Ado. K_m, Michaelis constant, is the required substrate concentration for the reaction rate to be half of its maximum rate, and it is a measure of an enzyme's affinity to its substrate (57). In other words, a high K_m means low affinity to the substrate, and a low K_m means high affinity. Human ADA2 has a K_m of 2-3mM, while ADA1 has 20-50μM (2; 50). ADA2 is, as initially stated, presumed to be an extracellular protein. This means ADA2 has a K_m 2000 to 20 000 times higher than the extracellular concentration of Ado (<1μM), resulting in nearly no extracellular hydrolysis of adenosine.

The isoelectric point (pI) and pH optimum for the two isoenzymes are also very different, indicating different locations for function. ADA2 has a pI of 7.75 and a pH optimum of 6.5, while ADA1 has 5.63 and 7.5 (50; 58; 59). All theoretical pI's are found using Expasy Compute pI/Mw. ADA2 has an acidic pH optimum, which does not correlate with the physiologic extracellular pH of 7.4 (60). Looking at ADA2's pI, the enzyme will have a slight negative charge in serum. In the lysosome, with a pH of 5, ADA2 will have a positive charge, making it possible for ADA2 to bind to adenosine on negatively charged DNA. In addition, local acidosis at inflammation sites and possible secretion of ADA2 to these sites could lead to a more efficient ADA2 function (4).

[11] Table adapted and combined after Tarrant et al. "Elucidating the pathogenesis of adenosine deaminase 2 deficiency: current status and unmet needs" (accessed May 9th, 2023) and Zavialov et al. "Structural Basis for the Growth Factor Activity of Human Adenosine Deaminase ADA2" (accessed May 11th, 2023), both licensed

Table 1.1: Properties and main differences between ADA1 and ADA2. Credit[11].

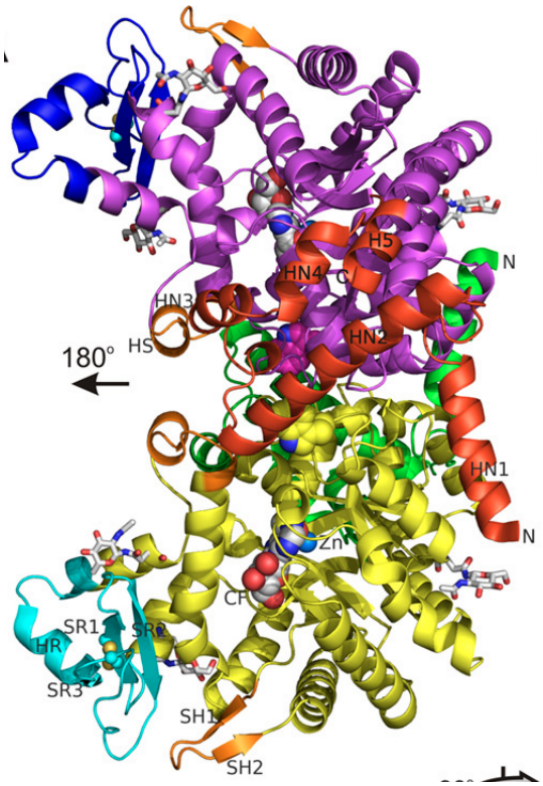
	ADA1	**ADA2**
Gene	ADA	ADA2 (CECR1)
Monomer	41 kDa	57 kDa
Active conformation	Monomer or heterodimer with CD26 or DPP IV	Homodimer
Inhibition by EHNA	High	Low/Resistant
Location	Intracellular (cytoplasmic). 'Ecto-ADA' when dimerized with CD26/DPP IV	Extracellular (secreted), lysosome?
Tissue distribution	All cells, highest in lymphoid cells (T-, B-, NK-)	Secreted by monocytes/macrophages, dendritic cells.(Plasma ADA2 activity elevated in some inflammatory conditions)
K_m	20-50μM	2-3mM
Isoelectric point	5.63	7.75
pH optimum	7.5	6.5

The structure of ADA2 was determined by Zavialov et al. (2009) using crystallization techniques. The dimer consists of two ADA domains responsible for the catalytic function, including the active site, two "putative receptor-binding" (PRB) domains responsible for receptor binding and signal transmittance, and lastly, two dimerization domains connecting the two monomers (Figure 1.11a). The ADA domains are the largest and structurally quite similar to ADA1 despite the highly dissimilar sequences. The PRB domains are located on the same side of the dimer and are highly exposed, fitting for receptor binding and signaling. The dimerization domains are the most conformationally stable, having a hydrophobic pocket for a tryptophan-residue (Trp) from the other monomer to bind to and, thereby, the connection of the two highly hydrophobic helices (49).

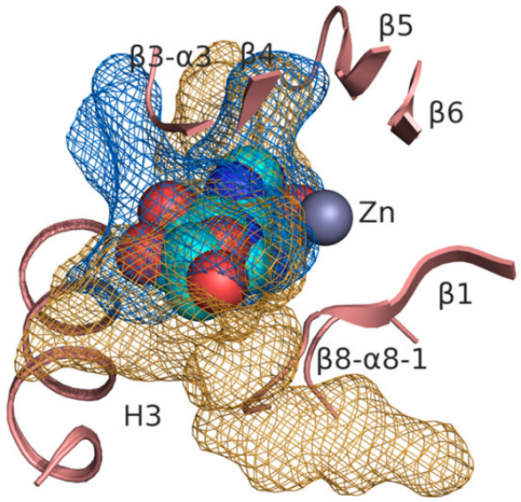
The active site depends on a Zn^{2+} ion for the catalytic function. The ion is the electrophile in the reaction and is oriented in the active site by His-86, -88, -330, and Asp-415. The opening of the active site is highly hydrophilic compared to ADA1, having Glu-182, His-267, and Ser-265 in the opening. The binding "pocket" itself is considerably more open in ADA2 than ADA1, which has been discussed as one of the factors for the high K_m-value and low specificity of ADA2 compared to ADA1 (Figure 1.11b). The two major differences are a hydrophobic chain, which is located further away from the purine ring of adenosine in ADA2, and a hydrophobic coil that forms a wall-like structure that is significantly shorter in ADA2 than ADA1 (49).

ADA2 also has multiple N-glycosylation sites (Figure 1.12). Zavialov (2010) found three glycosylated sites using a Fourier difference map: Asn-101, -159, and -352 (49).

[12]Figures adapted from illustrations by Zavialov et al. in "Structural Basis for the Growth Factor Activity of Human Adenosine Deaminase ADA2", licensed under Creative Commons Attribution-NonCommercial-NoDerivatives license (accessed May 11th, 2023) (49).



(a) The domains of ADA2. ADA, dimerization, and PRB domains of subunit 1 are yellow, green, and turquoise, respectively. Domains of subunit 2 are purple, red, and blue, respectively. Credit[12].



(b) Comparison of the active sites of ADA1 and ADA2. ADA2's and ADA1's binding sites are orange and blue "nets", respectively. The molecule in the center is coformycin. Credit[12].

Figure 1.11: Visulization of the domains and active site of ADA2.

However, the specific glycan structures were not determined. Glycosylation can reduce the proteolytic activity on ADA2 in extracellular spaces, which might suggest that ADA2 is an extracellular enzyme, but this alone is not definite proof.

Along with the glycosylation, a conserved disulfide bond is found in the PBR domain between Cys-111 and Cys-133. The bond is shown to be structurally important, as removing it stops the secretion of ADA2 (49; 61). Comparing the sequences of ADA2 from different species also shows that both the bond and glycosylation sites near the bond are conserved, confirming their importance for the protein's stability.

To summarize, ADA2 is active as a homodimer and consists of three domains: the catalytic ADA domain, the PBR domain with a conserved disulfide bond, and the dimerization domain. A Zn^{2+}-ion is critical for substrate binding and catalytic function, but the binding pocket of ADA2 is significantly larger and not as specific to Ado compared to ADA1.

1.3.3 N-glycans and their link to ADA2

The presence of glycosylation sites in ADA2 is preserved across species. Determination of these sites and potential glycan structures can tell something about their location and possible secretion pathway and function. Therefore, knowing the glycan structures is a crucial tool when investigating glycoproteins. The sequence of human ADA2 shows that there are, in fact, four possible sites, one more than Zavialov found in his studies.

MLVDGPSERP	ALCFLLLAVA	MSFFGSALSI	DETRAHLLLK	EKMMRLGGRL	VLNTKEELAN
ERLMTLKIAE	MKEAMRTLIF	PPSMHFFQAK	HLIERSQVFN	ILRMMPKGAA	LHLHDIGIVT
MDWLVRNVTY	RPHCHICFTP	RGIMQFRFAH	PTPRPSEKCS	KWILLEDYRK	RVQNVTEFDD
SLLRNVTLVT	QHPEVIYTNQ	NVVWSKFETI	FFTISGLIHY	APVFRDYVFR	SMQEFYEDNV
LYMEIRARLL	PVYELSGEHH	DEEWSVKTYQ	EVAQKFVETH	PEFIGIKIIY	SDHRSKDVAV
IAESIRMAMG	LRIKFPTVVA	GFDLVGHEDT	GHSLHDYKEA	LMIPAKDGVK	LPYFFHAGET
DWQGTSIDRN	ILDALMLNTT	RIGHGFALSK	HPAVRTYSWK	KDIPIEVCPI	SNQVLKLVSD
LRNHPVATLM	ATGHPMVISS	DDPAMFGAKG	LSYDFYEVFM	GIGGMKADLR	TLKQLAMNSI
KYSTLLESEK	NTFMEIWKKR	WDKFIADVAT	K		

Figure 1.12: Amino acid sequence of human ADA2 including possible glycosylation sites (red, underlined) and signal peptide (blue). Credit[13].

We find the sequence's theoretical glycosylation sites at Asn-127, -174, -185, and -378. The crystallographic analysis found Asn-101, -159, and -352 to be glycosylated, as they have removed the signal peptide of 26 amino acids when numbering the sequence (blue amino acids in Figure 1.12). In other words, the theoretical sites of Asn-127, -185, and -378 are found to be glycosylated. An N-glycan linked to Asn-174 may not have been detected due to its exposed location at the surface of the PRB domain. If the glycan has a flexible conformation and, therefore, not a set position, the crystallographic analysis may not detect the structure.

An important aspect of N-glycosylation is which N-glycans are present on the protein, as presented in Section 1.2.4. Whether the N-glycans are trimmed *tomi*-glycans or resemble the structures from the biosynthesis says something about where the protein is located: extracellularly as a secreted protein or in the lysosomes. The glycans serve different purposes depending on the localization of the protein; for example, in

[13]Illustration made by the candidate with sequence obtained from Uniprot.org, accession number Q9NZK5 (accessed February 2th, 2024)(58).

lysosomes, the purpose would be to protect the protein from proteases.

Also presented in Section 1.2.4 was the importance of N-glycans for the correct folding of glycoproteins. A study from 2022 by Ito, confirmed this is also the case for ADA2 (62). The study found that the initial binding of N-glycans to the peptide and modifications in the ER, except Asn-185, are crucial for the folding and secretion of active ADA2. A glycan profiling found a difference between intracellular and secreted ADA2. Intracellular ADA2 mainly had short, trimmed N-glycans, while secreted ADA2 mainly had long, complex-type N-glycans. Additionally, the activity of intracellular ADA2 was highly reduced when removing the N-glycans, while secreted ADA2 had unchanged activity. This implies that N-glycans are important for folding. Still, as deglycosylation of secreted ADA2 does not change the activity, it is unlikely they have any say in the catalytic activity of the protein. This aligns with the amino acid sequence as the glycosylation sites are not located near the active site and binding pocket but near the domains responsible for dimerization (62).

There is still insufficient knowledge of N-glycans' role in ADA2's function, but studies like the one by Ito give some insight. This topic should be further investigated, as it will give important insights into not only the function of ADA2 but also a better understanding of the function of N-glycans.

1.3.4 ADA2 across species

Part of what has complicated the research of ADA2 is the differences in the enzyme between species. This section will summarize the key differences between some species. For example, in chicken, the K_m-value and activity towards adenosine is significantly higher and comparable to ADA1. The same applies to pig ADA2. On the other hand, sheep have much lower activity, comparable to humans, despite having a K_m-value similar to chicken. However, sheep have not been investigated to the same extent as other species. In addition, ADA2 seems to be completely missing from rodents, i.e. mice and rats, only ADA1 remains (63). In comparison, ADA1 has a rather consistent function across the mentioned species, making one wonder why there are so large differences for ADA2.

Characterization studies of the total ADA activity in chicken and pig, meaning the sum of catalytic activity of ADA1 and ADA2, have been performed by EHNA studies (54; 64). EHNA is known to only be a potent inhibitor for ADA1 and is used to

find the activity of the specific enzymes (Section 1.3.2). Iwaki-Egawa et al. (2004) characterization of chicken found that ADA2 stands for 40% of the total ADA activity in serum (54). The characterization of pigs by Contreras-Aguilar et al. (2020) found that ADA2 accounted for approximately 13% and 0.2% of total ADA activity in serum and saliva, respectively (64). Iwaki-Egawa et al. used 3mM Ado was, significantly higher than the physiological concentration (>1μM), meaning the ADA2 activity is likely lower than 40%. Contreras-Aguilar et al. used a pre-made ADA assay, where the Ado-concentration is not specified. Additionally, note that the ADA2 was purified from chicken liver, while saliva and serum were used for pigs. ADA2 is assumed to have promoted expression in macrophages/monocytes, which are more abundant in the liver than in saliva and serum, reducing the comparability of the results.

One of the more obvious differences is the K_m-value. In humans, it is 2-3mM, while chicken and sheep have 0.1-0.3mM. The K_m-value of porcine ADA2 was not found by the author, but based on other findings, it is assumed to be similar to that of chicken and sheep. Finding the reason for this requires investigating the amino acid sequences to locate differences, especially at the active and binding sites. Already, it is determined that chicken has a more hydrophobic binding site, with Ala, whereas it is Ser in humans (49). The same trend is seen in the binding sites of human ADA1 and ADA2, where ADA1 has a much more hydrophobic binding site than ADA2 (Section 1.3.2). Another trend is that the K_m-value decreases with lower species, such as flies and fish, along with increased hydrophobic properties of the binding site. For example, insects' K_m-value is only 0.05-0.1mM, suggesting very high ADA2 activity. According to Zavialov, this might suggest that ADA2's function is more specialized with higher species (49). A comparison of the sequence of human and porcine ADA2 done by Greiner-Tollersrud et al. (2020).

The pI and pH optimum also differ between the species. The pH optimum for chicken ADA2 is 6.5, and the theoretical pI is 5.99, meaning the enzyme has a slight negative charge in its most optimal form (54). The pH optimum of sheep ADA2 has not been determined, but the theoretical pI is 8.64. Regardless of the pH optimum, sheep ADA2 will have a positive charge due to its high pI and the physiological extra- and intracellular pH never exceeding 8.6. The positive or negative effects the charge of the protein might have on its activity depends on what this activity is and whether it is lysosomal or extracellular.

The degree of glycosylation and position of the glycosylation sites also differ be-

tween the species. It is found that chicken ADA2 is indeed N-glycosylated, though the glycan structures are not determined (54). Chicken ADA2 has seven glycosylation sites (Figure 1.13). Sheep ADA2 has not yet been thoroughly researched regarding N-glycosylation, but one theoretical glycosylation site is of special interest. Asn-349 (corresponds to positions 293 and 291 in humans and chickens, respectively) is potentially N-glycosylated and is situated at the opening of the active site. Porcine ADA2 is also glycosylated, and the glycan structures in ADA2 from the brain were determined by Greiner-Tollersrud et al. (2020) (65). The sequences of humans and pigs were also compared in this article, though from a different perspective. Nonetheless, porcine ADA2 does not contain the glycosylation site near the active site and has only three glycosylation sites. Chicken and sheep ADA2 have more glycosylation sites than human ADA2: 8, 7, and 4, respectively.

```
tr|Q2VQV8|Q2VQV8_CHICK      ----------------------------------------------------MKPPSG  6
sp|Q9NZK5|ADA2_HUMAN        ----------------------------------------------------MLVDGPSE  8
tr|W5PV50|W5PV50_SHEEP      MCAHTDAAGHTPPPGCREVKPCPREAKDCEPRLPMGLCGNGTHSPLSLELSQGILASMSG  60
                                                                                  :*:

tr|Q2VQV8|Q2VQV8_CHICK      QGIRLACL--LGLASLC--LPTPLWEDRDSLMQEEEARLTGGNLVLSTQEQQLSMKLISL  62
sp|Q9NZK5|ADA2_HUMAN        RPALCFLLLAVAMSFFG--SALSIDETRAHLLLKEKMMRLGGRLVLNTKEELANERLMTL  66
tr|W5PV50|W5PV50_SHEEP      RLALPSLLLAVVVSSFCPATSTSGDETRNQLLTREKMMRVGGQLVLTREEELANQRLRAL  120
                            : ::  :*:::  :.:.     :  * *  *:  *:::: ** *** * *:::: :*  *

tr|Q2VQV8|Q2VQV8_CHICK      KKKEVAAALNTGQFPPSMHFFRAKSLIEQSAVFSILKRMPKGAVLHLHDYAMLSVEWLVY  122
sp|Q9NZK5|ADA2_HUMAN        KIAEMKEAMRTLIFPPSMHFFQAKHLIERSQVFNILRMMPKGAALHLHDIGIVTMDWLVR  126
tr|W5PV50|W5PV50_SHEEP      KEAEMREALRTGSILPSMHFFQAKSLIEKSKVFNILKKMPKGAALHVHDFSITSPDWLVK  180
                            * :*: :*.:*  ::******:**.*** * **:**.:*****:**.**  : . :***

tr|Q2VQV8|Q2VQV8_CHICK      NATYLPDCYICFTQTGTVRFRFSKPHPPHPVPAQCSQWVLLETYRKQLRNVTEFDHSLLR  182
sp|Q9NZK5|ADA2_HUMAN        NVTYRPHCHICFTPRGIMQFRFAHPTPRP--SEKCSKWILLEDYRKRVQNVTEFDDSLLR  184
tr|W5PV50|W5PV50_SHEEP      NATYRPHCYFCFTRRGTLQFRFAHPAPPTPKPAECPQWVLLEKFRKGLHNVSDFDNSLLR  240
                            *.**:*:*.:*** :*. :***::* *. . .. *:.*.*** :** . **::** ****

tr|Q2VQV8|Q2VQV8_CHICK      NLTLVTDAPELAYPSQAFIWKRFEEVFLIASGLICYAPVFKAYFYQGLLELYEDNIQYVE  242
sp|Q9NZK5|ADA2_HUMAN        NFTLVTQHPEVIYTNQNVVWSKFETIFFTISGLIHYAPVFRDYVFRSMQEFYEDNVLYME  244
tr|W5PV50|W5PV50_SHEEP      NFTLMTENPHVTYADQDAIWAKFKTIFFVLRGLVSYAPVFRDYLSQGLEEFYQDNVLYLE  300
                            *:**:*  *:: *  *  .* :*:::*:  :**: *****::*  ... *:*:**::* *

tr|Q2VQV8|Q2VQV8_CHICK      IRAILPAVYELDGTQHNKSWSVAAYQEVSRRFVQEHPDFLGAKVIFAAHRMLNVSQIKET  302
sp|Q9NZK5|ADA2_HUMAN        IRARLLPVYELSGEHHDEEWSVKTYQEVAQKFVETHPEFIGIKIIYSDHRSKDVAVIAES  304
tr|W5PV50|W5PV50_SHEEP      IRASLYPVYELNGTVYSQEWLVRTYEEVAHNFAKAHPGFIGIKLIFSDNRTKNESLIKKS  360
                            *** * :**** .. :  :*:* :*:**:  *:  ** ***:* *.:::* :.:. *.::

tr|Q2VQV8|Q2VQV8_CHICK      IITAMELRAHFPDTLAGFDLVGNEDEGHSLWEFKDVLTIPYSLRANLPYFFHAGETDWQG  362
sp|Q9NZK5|ADA2_HUMAN        IRMAMGLRIKFPTVVAGFDLVGHEDTGHSLHDYKEALMIPAKDGVKLPYFFHAGETDWQG  364
tr|W5PV50|W5PV50_SHEEP      IQTAMGLRVKFPGIVAGFDLVGREDTGHTLYDYREALMIPASQGVHLPYFFHAGETDLEG  420
                            * .**:** ::*  :******* **:**:* :::::*:**:. :: ***********::*

tr|Q2VQV8|Q2VQV8_CHICK      TSVDNNVLDALILNTSRIGHGLALTKHPVAKNLSLKRDIPVEVCPISNQVLLLVSDLRSH  422
sp|Q9NZK5|ADA2_HUMAN        TSIDRNILDALMLNTTRIGHGFALSKHPAVRTYSWKKDIPIEVCPISNQVLKLVSDLRNH  424
tr|W5PV50|W5PV50_SHEEP      TPIDRNLLDAVILNSTRIGHGFALSKHPAVLAEAWKKDIPIEVCPISNQVLKLVSDLRNH  480
                            *::*:* ***::**::*****:**:***::   ::*:***:**********:******:*

tr|Q2VQV8|Q2VQV8_CHICK      PAASLMAEGHPVVVSPDDPSIFGAKGVSYDFYEVFMAIGGMNADLRTLKQLALNSIKYSA  482
sp|Q9NZK5|ADA2_HUMAN        PVATLMATGHPMVISSDDPAMFGAKGLSYDFYEVFMGIGGMKADLRTLKQLAMNSIKYST  484
tr|W5PV50|W5PV50_SHEEP      PAAVMMAAGYPMVISSDDPSVFGARGLSYDFYEAFMGIGGMSADLRTLKQLAINSIRYST  540
                            *.* :** *:*:*:*:***. *****:******:******* ********** ***:**:

tr|Q2VQV8|Q2VQV8_CHICK      MLLDEKAKAREVWQKKWDQFVADLTDSSLMEL  514
sp|Q9NZK5|ADA2_HUMAN        LLESEKNTFMEIWKKRWDKFIADVATK-----  511
tr|W5PV50|W5PV50_SHEEP      LSETMKKSAMETWEERWQKFVAGLARGPK---  569
                            :*: :*  .:* * ::*::*.*:.:
```

Figure 1.13: Comparison of the amino acid sequences of human, chicken, and sheep ADA2. Double underline, signal peptide; "*", same amino acid; ":", Gg/Oa same as Hs; ".", Gg and Oa the same; Yellow highlight, binding site; Green highlight, active site; Bold red N, glycosylation site. Credit[14].

The effect of the pH value on the activity, the glycans bound, etc, on sheep and chicken will be discussed as ADA2 from these species are purified and analyzed.

1.3.5 ADA2 and its possible functions - Current theories and models

One of the leading researchers in the field of ADA2 and DADA2, Michael S. Hershfield, wrote a summarizing article with Teresa K. Tarrant and Susan J. Kelly (2021) on the then-current theories on ADA2's function. To begin with, ADA2 was discovered as a possible growth factor in flesh flies, getting the name insect-derived growth factor (IDGF), and soon after renamed ADGF. It was later found that ADGF could degrade adenosine, which allowed for growth, rather than ADGF itself stimulating growth (56). The idea of human ADA2 being a secreted GF by regulating extracellular Ado was set when it was found that it was transcripted by the CECR1-gene. ADA2's GF function would come from the PBR domain (Section 1.3.2), which is missing in ADA1. Further investigating of ADA2 receptor-binding found that ADA2 did cause some proliferation of CD4+ T-cells, but additional findings seemed to disprove a GF function through Ado regulation (56; 68).

Despite the disproval of a GF function through Ado regulation, there are still some theories on the specific regulation of Ado. One theory by Carmoina-Rivera et al. is that elevated adenosine levels cause the formation of *neutrophil extracellular traps* (NETs). NETs stimulate M1 macrophages to produce TNF-α. The theory was based on DADA2 patients having NETs during active inflammation, as well as higher Ado levels (69). Tarrant, Kelly, and Hersfield disprove this theory as NETs are present in many autoinflammatory and autoimmune diseases, and it is unlikely that elevated adenosine levels are the sole reason for NET formation. Another assumption in this theory is that ADA2 is solely responsible for the extracellular regulation of Ado and ADA1 for intracellular regulation, i.e., a complete separation of the two isoenzymes. This idea is not likely correct as ADA1 has been proven to also have extracellular activity, and ADA1 has a significantly higher hydrolysis rate of adenosine than ADA2 (56).

A hypothesis on another disease by Dhanwani et al. (2020), *nonalcoholic fatty liver disease* (NAFLD), suggests that it is the products from ADA2's catalytic activity, rather than the substrate that induces pathogenic effects (70). Lacking functional ADA2 causes the translocation of *2'-deoxyadenosine* (dAdo) from plasma to endothe-

[14] Illustration made by the candidate with sequences obtained from Uniprot.org, accession numbers Q9NZK5, Q2VQV8, and W5PV50 (accessed February 2th and 11th, 2024) (58; 66; 67).

lial cells, where ADA1 will produce dIno, which accumulates and triggers type I IFN-β production. The theory shares some issues with the NETs theory of the complete separation of ADA1 and ADA2, ADA1 being intracellular and ADA2 being extracellular. In addition, the theory does not take into account that ADA2 is even less effective on dAdo than Ado and that the deamination of the substrate to *2'-deoxyinosine* (dIno) by ADA1 is an important reaction and it is unlikely for it to cause any pathologic effects, even with a higher substrate load. Other research on ADA2 and NAFLD has revealed a paradox regarding the inflammatory effects of ADA2 and DADA2 (71). It was found that ADA2 secreted from monocyte-derived macrophages increased the production of pro-inflammatory cytokines, or at least it occurred in parallel, as it was not investigated if the cytokine production was ADA2 dependent. The paradox is that a deficiency of ADA2 causes pro-inflammatory reactions, and the accumulation of both substrate and product increases inflammation.

As theories regarding ADA2's extracellular function seem to fall short, other theories regarding its location have become more present. A theory by Greiner-Tollersurd et al. (2020) is that ADA2 is a lysosomal DNase, binding to adenosine on DNA and regulates type I IFN response through the STING pathway (65). Other theories with an evolutionary perspective might suggest human ADA2 has evolved to be active towards another substrate or have a non-catalytic function. The possibility of ADA2 being a lysosomal DNase is what will be investigated in this thesis.

1.4 Deficiency of adenosine deaminase 2

One of the main motivators for determining the function of ADA2 is to ultimately find the pathogenic mechanism of DADA2 and treatment for the disease. As stated in Section 1.2.2, DADA2 is a monogenic, autoinflammatory disease and generally has an early onset. 77% of patients are diagnosed before the age of 10 and 25% before the age of 1 (1). Being a monogenic, autosomal recessive inherited disease, it is rare, with about 600 confirmed patients globally. However, it is thought that there are many undetected cases, making the number much higher.

The symptoms vary between patients, even patients from the same family. But, some common symptoms are vasculitis, PAN-like, livedo racemosa, early-onset stroke, bone marrow failure, and, lastly, systemic inflammation (72). Some patients only have mild symptoms and respond to treatment, while others have much more severe symp-

toms leading to early death. The mortality rate is 8% before the age of 30 (1).

Nanthapisal et al. (2016) attempted to summarize and characterize the symptoms of 15 DADA2 patients and an additional 39 patients from three other screenings by Navon Elkan et al. (2014), Zhou et al. (2014), and Batu et al. (2015). 10 out of the 15 patients in Nanthapisal's group showed symptoms, despite all having a mutation that may cause DADA2. The symptom-presenting patients all had an onset before the age of 10, the youngest at 2 years old, and the symptoms range from mild vasculitis of the palms to severe inflammation and neurological effects (72). A repeating trend for DADA2 patients is their PAN-like symptoms, and this disease being the first thought diagnosis. In Nanthapisal's study, 9 out of 10 had initial symptoms of PAN, which eventually led to a DADA2 diagnosis. As PAN is a vascular disease, and vasculitis is a common symptom of DADA2, these two diseases are similar and genetic screening is necessary for accurate diagnosis (72–75).

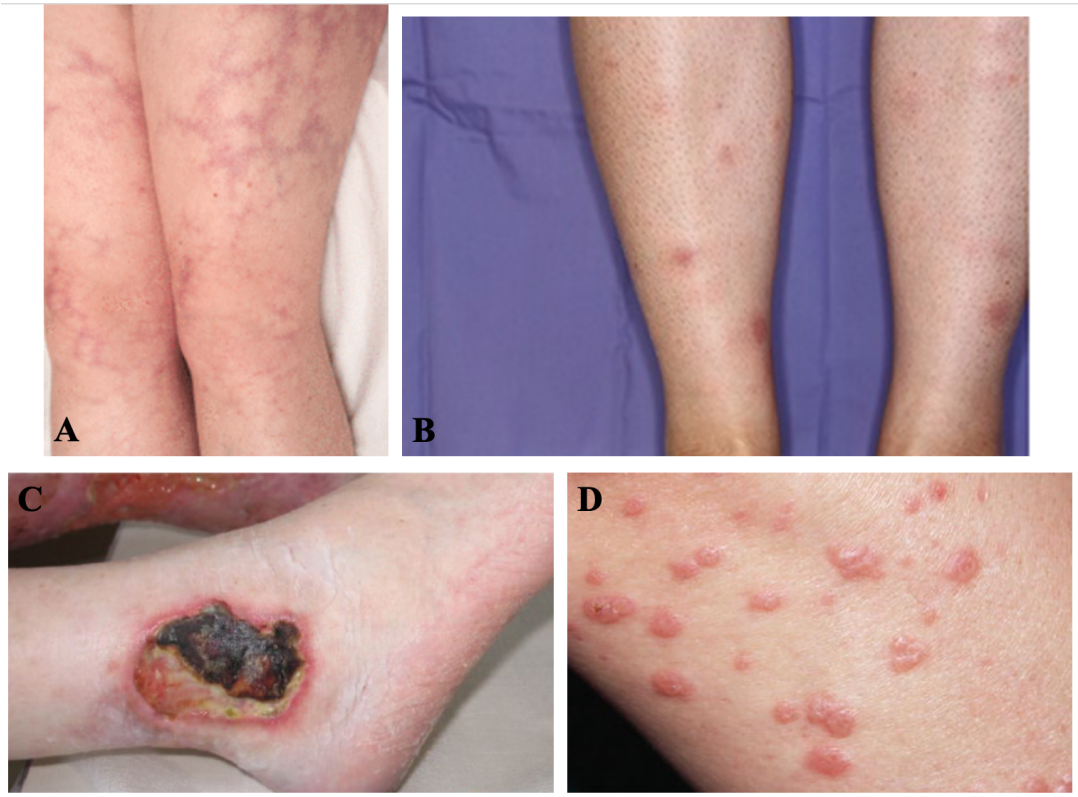


Figure 1.14: Cutaneous symptoms in DADA2 patients. A, livedo racemose; B, PAN; C, ulceration; D, erythema nodosum. Credit[15].

Through characterization of patients in all four screenings, a total of 54 individuals, it was observed that cutaneous symptoms were the most prevalent (Table 1.2). Cuta-

[15]Illustration is a combination of images obtained from dada2.org (Last updated June 23th 2023, accessed February 27th 2024) (76).

neous symptoms include anything related to the skin, e.g., livedo racemosa, peripheral ischemia, ulceration, PAN, and erythema nodosum (Figure 1.14). The latter was only observed in the three additional studies. The neurological symptoms described include brain, brainstem, and basal ganglia stroke, intracerebral hemorrhage, motor neuron defects, and axonal polyneuropathy. Bone marrow-related symptoms such as cytopenia, neutropenia, and deficiency of B cells are also seen. Ig levels were monitored regarding immunological symptoms, concluding some patients had low levels of either IgM or IgG, suggesting an immunodeficiency. In addition, systemic inflammation and recurrent fevers were also common symptoms. All patients in the three additional studies had recurring fevers, consistent with autoinflammatory diseases (Section 1.2.2).

Table 1.2: Summary of symptom groups of DADA2 patients from four screenings. Credit[16].

Symptom groups	**No. of patients (n = 54)**	
Cutaneous involvement	41	(75.9%)
Neurological involvement	35	(64.8%)
Central nervous system	23	(42.6%)
Peripheral nervous system	12	(22.2%) *
Gastrointestinal involvement	25	(46.3%)
Immunologic involvement	14	(25.9%) **
Renal involvement	13	(24.1%)
Ophthalmologic involvement	9	(16.7%) **
Asymptomatic	5	(9.2%)
Congenital heart disease	1	(1.9%) ***
Treatment with anti-TNF	28	(51.8%)

* (74; 75) not reported
** (73) not reported
*** (73–75) not reported

There is some consistency in the mutations in the patients (Table 1.3). The most notable mutation is the homozygous substitution of Gly47Arg (G47R), as Gly47 is conserved across species. Gly47 is likely to be important for the stability of the ADA2 homodimer, and a mutation causes failure in the dimerization (73). According to Batu (2015), this mutation may also be correlated to PAN-like symptoms but not central nervous system (CNS) disease. Patients with other mutations present more symptoms of the CNS, including stroke (75). Other prevalent mutations are the substitution of Arg169Gln (R169Q) and Pro251Leu (P251L). Arg169 and Pro251 are considered important for the PBR domain and active site, respectively (73).

[16]Table adapted from table by Nanthapisal et al. in "Deficiency of Adenosine Deaminase Type 2: A Description of Phenotype and Genotype in Fifteen Cases" (accessed October 20th, 2023) (72).

Table 1.3: Recurring mutations in DADA2 patients. Credit[17].

Mutation	No. of patients (n = 54)	
Hom. G47R	30	(55.6%)
Het. G47R/A/V	3	(5.6%)
Hom. P251L	7	(13.0%) *
Het. P251L	4	(7.4%)
Hom. R169Q	2	(3.7%)
Het. R169Q	5	(9.2%)

Hom, homozygous mutation; Het, heterozygous mutation
* 5 asymptomatic

Anti-TNF is the most common choice of treatment for DADA2 patients. Additional treatments have also been given for specific symptoms, e.g. corticosteroids, antibodies (rituximab, monoclonal antibody against CD20), and methotrexate/adalimumab (for rheumatic symptoms).

1.4.1 ADA-SCID and DADA2

Deficiency of ADA2 resembles deficiency of ADA1 in numerous ways, but there are some very important differences. Deficiency of ADA1 causes *severe combined immunodeficiency disease* (SCID), which is characterized by depletion of T-, B-, and NK lymphocytes (part of the adaptive immune system) and, specific for ADA-SCID, elevated levels of Ado and dAdo (56). In ADA-SCID, the accumulation of Ado and dAdo leads to apoptosis of lymphocytes. Especially dATP, when accumulated, functions as an inhibitor for dinucleotide reductase, inhibiting DNA synthesis (77). Since these symptoms are well known and part of the pathogenesis of ADA-SCID, they might be more easily diagnosed and treated. In addition, the function of ADA1 is well understood compared to ADA2, hence the pathomechanism of the disease, making it is easier to develop and implement a specific treatment.

ADA-SCID patients are initially treated with enzyme replacement therapy and later, if possible, with *stem cell transplantation* (SCT) from a family member. As of 2016, ADA-SCID is also being treated with *ex vivo* hematopoietic stem cells (HSC), a medicine called Strimvelis^{TM}, developed by the Telethon Foundation and San Raffaele Scientific Institute (78). DADA2 patients are treated with anti-TNF, but there have been cases of successful SCT in DADA2 patients with hematologic symptoms

[17]Table made by the candidate based on data from Nanthapisal et al., Navon et al., Zhou et al., and Batu et al (72–75).

(22). As ADA-SCID generally has more severe and lethal symptoms, enzyme replacement therapy and SCT are critical for the survival of the patients. Most cases of the DADA2 phenotype, i.e., inflammatory symptoms, can be managed with anti-TNF treatment. Nevertheless, the phenotype of DADA2 has many aspects (vascular, inflammatory, neurological), and treatment available today only targets one of these aspects, i.e., the inflammatory symptoms.

Table 1.4 summarizes the main differences between ADA-SCID and DADA2.

Table 1.4: Comparison of ADA-SCID and DADA2. Credit[18].

	ADA-SCID	**DADA2**
Clinical phenotype	SCID (T-, B-, NK-cells) Early onset (85%) Late onset (15-20%) "Partial deficiency" = normal immune function; ADA1 absent in RBC, 2-70% presence in lymphocytes	Inflammation: Livedo rash, vasculopathy (PAN), recurrent stroke Bone marrow failure: Cytopenia, red cell aplaisa Immune deficiency: Mainly humoral (antibodies), but varies
Metabolic phenotype	Increased Ado and dAdo in plasma Increased dATP and dAXP in RBC Reduced AdoHcyase	No/minimal elevation of Ado and dAdo in plasma * No elevation of RBC dAXP
Genotype/Phenotype	Residual ADA activity correlates with levels of RBD dAXP and phenotype	Residual ADA activity correlates with inflammation and bone marrow failure phenotype

SCID, Severe combined immunodeficiency; RBC, red blood cells; Ado, Adenosine; dAdo, deoxyadenosine; dATP, deoxyadenosine triphosphate; dAXP, all deoxyadenosine phosphate; AdoHcyase, S-adenosylhomocystein hydrolase
*Not systematically researched per 2021

[18]Table adapted and expanded based on Tarrant et al. "Elucidating the pathogenesis of adenosine deaminase 2 deficiency: current status and unmet needs" (accessed May 9th, 2023), licensed under Creative Commons Attribution-NonCommercial-NoDerivatives license (56).

Chapter 2

Methods

2.1 Protein purification

In the following sections, the different techniques used for isolating ADA2 are presented.

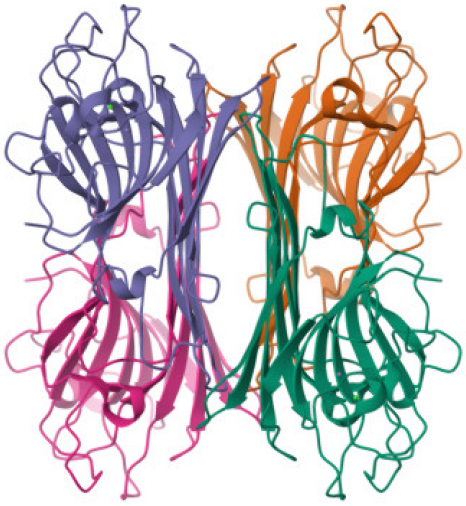
2.1.1 Concanavalin A column chromatography

Concanavalin A column chromatography (Con A) is a general affinity chromatography method (79). Con A is a lectin protein isolated from jack beans, and has a general specificity for carbohydrate compounds and residues (80). More specifically is Con A a mannose-binding protein, hence specific for glycoproteins with N-linked glycans (80; 81). Con A is a tetramer of identical monomers at neutral pH, which is favorable for binding the residues. Ca^{2+} or Mn^{2+} ions are also needed for optimal binding (80). The liquid in which the Con A particles are dissolved is often the same buffer/solvent used for the separation, but this is not required. Generally, the eluting buffer is chosen based on the required pH for the following steps.

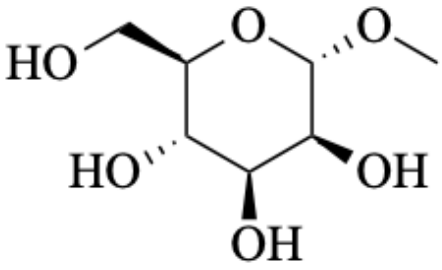
To elute the glycoproteins, a surplus of a competing ligand must be used (80). The competing ligand can have a higher affinity to Con A, though this is not crucial as the competing ligand is in a large surplus. This purification uses a large surplus of competing α-methylmannoside.

[1]Illustration made by Parkin, S., Rupp, B., Hope, H. for "Atomic resolution structure of concanavalin A at 120 K." and reused with permission of International Union of Crystallography (accessed May 11th, 2024) (82).

[2]Illustration made by the candidate in ChemDraw



(a) 3D structure of Concanavalin A. Credit[1].



(b) Structure of α-methylmannoside. Credit[2].

Figure 2.1: Structures of Concanavalin A and α-methylmannoside.

2.1.2 Ion-exchange chromatography and carboxymethyl column chromatography

Ion-exchange chromatography is a separation technique based on the charge of compounds in the mobile phase and the set charge of the stationary column (83). There are two types of ion exchange: Anion exchange and cation exchange. Anion exchange columns have positively charged resins that interact with negatively charged compounds, anions, in the mobile phase. Cation exchange has negatively charged resins interacting with positively charged compounds, cations (83).

Cation-exchange chromatography with carboxymethyl (CM) as the resin is used in this thesis. CM is a weakly acidic resin with a weaker negative charge and is deprotonated at a pH of 4 and higher. The mobile phase cannot have a pH lower than 4, as this will lead to the protonation of the CM molecules and "loss" of its charge (83). CM is coupled to Sepharose, which is a beaded polysaccharide. Sepharose is commonly used for these types of columns.

The pH of the mobile phase must be chosen based on the pI of the target compound. In this case, ADA2 from chicken has a pI of 6, and ADA2 from sheep has a pI of 8.6 (Section 1.3.4). The mobile phase has to have a lower pH than the pI to have a positively charged protein. The mobile phase for chicken is set to 6.5, which results in a weakly positive ADA2. For sheep ADA2, the mobile phase is set to pH 7.2, which results in a stronger positive charge.

A competing NaCl gradient elutes the proteins bound to the column. For a compound with a lower positive charge, a lower concentration of the competing salt is necessary. This means that ADA2 in chicken with a slight positive charge at pH 6.4, will elute with low salt concentrations. ADA2 from sheep will need a stronger competitor due to its higher affinity to the column and, therefore, be in a fraction of a higher concentration. In addition to the degree of positive charge of ADA2, it is important to consider which cation is used as a competitor. CM has a higher affinity to H^+ than Na^+, which means that a surplus of Na^+ is needed (83).

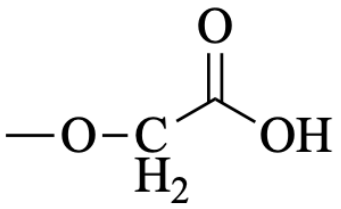


Figure 2.2: Carboxymethyl

This competition is quantified by the *selectivity coefficient*, explained by Equation 2.1 (83):

$$K = \frac{[R^- ADA2^+][Na^+]}{[R^- Na^+][ADA2^+]} \tag{2.1}$$

Where R is CM

2.1.3 SDS-PAGE

SDS-PAGE is the abbreviation for *sodium dodecyl sulfate gel electrophoresis*, and it is a type of gel electrophoresis (84). Electrophoresis is the separation of ions due to the effects of an applied electric field (85). The electric fields create a force on the ions, which results in acceleration and mobilization. A frictional force also occurs, which results in the ions achieving a steady pace along the gel. This results in Equation 2.2, giving the theoretical mobility of an ion (equation and explanations from (85)):

$$u_{\text{ep}} = \frac{q}{f}E = \mu_{\text{ep}}E \tag{2.2}$$

Where

- u_{ep}: Velocity
- q: Charge of ion (coloumbs)
- E: Electric field (V/m)
- qE: Electromotoric force (newtons)
- f: Friction coefficient
- μ_{ep}: Electrophoretic mobility

This equation shows the relation between the size and charge of an ion. f is proportional to the size of the ion, while μ_{ep} is inverse proportional to f and proportional to q. I.e., a larger size results in less mobility, while a higher charge results in higher mobility. For an assumed spherical ion moving through a viscous fluid, the friction coefficient is calculated with *Stokes equation* (85):

$$f = 6\pi\eta r \tag{2.3}$$

Where η [kg m^{-1} s^{-1}] is the viscosity of the fluid and r the radius of the ion.

The friction coefficient and electrophoretic mobility apply regardless of whether gel or capillary electrophoresis is used. Either way, the effect of the applied electric field or voltage on the ions determines the mobility when the same fluid is used. Different methods differ in how the charge is related to the mass.

SDS-PAGE separates proteins based on their molecular weight by denaturing them, thereby giving them a negative charge proportional to their weight (86). The method is commonly used in biomedicine due to its ability to achieve high protein separation. The denaturation is done by an anionic detergent, SDS, which uses a reducing agent *dithiothreitol* (DTT), to reduce the disulfide bonds between thiol groups in cysteines (87).

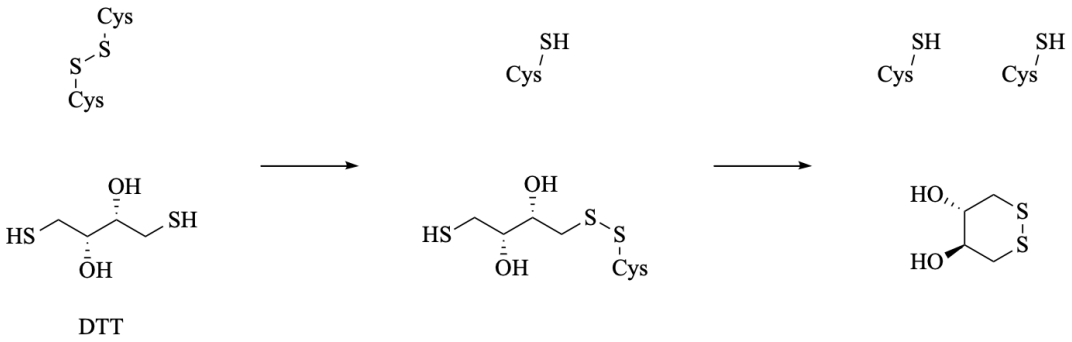


Figure 2.3: Redcution of disulfide bonds between Cysteines with DTT. Credit[3].

The denatured proteins are added to the strongly negative SDS, the running buffer, which binds to the hydrophobic regions in the peptides/proteins and further unfolds them. SDS and the reducing agent, e.g. DTT, ensure that the peptides stay linear. This way, the peptides have the same conditions for migration, namely length and a proportional charge. The amount of SDS the peptides bind to is proportional to the

[3]Illustration made by the candidate in ChemDraw

size of the peptide and the number of hydrophobic areas, which means that peptides of the same size have the same negative charge. When an electric field is applied, the larger peptides have a stronger "pull" towards the positive cathode. However, pores in the polyacrylamide gel do so that the larger peptides experience more friction and are therefore found closer to the starting point (88).

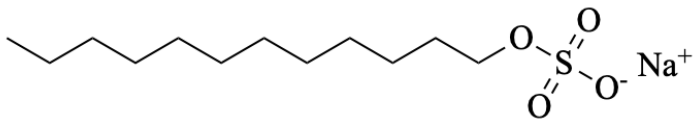


Figure 2.4: Sodium dodecyl sulfate (SDS). Credit[4].

A dye is used to stain the gel and visualize the bonds formed during the electrophoresis. Coomassie blue is often used, but other methods, such as autoradiography when dealing with radiolabeled proteins, are also possible (88).

2.2 Protein characterization and analysis

The following sections present the different analysis techniques used during the purification and analysis of purified ADA2.

2.2.1 Spectrophotometry

Spectrophotometry is a quantification method based on the interaction between a compound and light (89). The method measures the concentration of compound(s) by quantifying the absorbance of light done by the compound(s), i.e. a detector measures the "loss of light" by comparing the initial irradiance, P_0 and detected irradiance, P. In solutions with multiple compounds, spectroscopy cannot be used for specific quantification, but rather the approximate amount of compounds in solution. On the other hand, if there is a solution with only one compound, the method is a convenient tool for quantification. Irradiance is defined as the energy per second per unit area of light, and the ratio is known as *Transmittance*, explained by the following equation (89):

$$T = \frac{P}{P_0} \tag{2.4}$$

[4]Illustration made by the candidate in ChemDraw

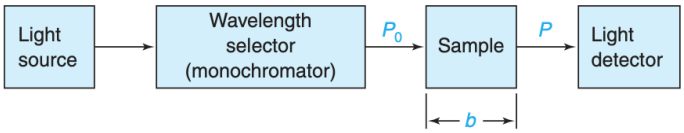


Figure 2.5: Overview of monochromatic spectrophotometry. Credit[5].

It is more common to use *absorbance* (A) or *optical density* (OD) when reviewing spectrophotometric data. In terms of solution with various proteins, one can assume an A equal to 1 is proportional to 1mg proteins per mL of undiluted solution. The absorbance is defined as (89):

$$A = \log \frac{P_0}{P} = -\log T \tag{2.5}$$

Absorbance is more convenient to use when quantifying because it can be used directly to determine the concentration of the target compound using *Beer's law* (89). To use this law, the *molar absorptivity,* ε [M^{-1} cm^{-1}], of the compound must be known. ε specifies a substance's or compound's ability to absorb light at a specific wavelength, λ (89):

$$A = \varepsilon bc \tag{2.6}$$

As seen in Figure 2.5, monochromatic light is used. Monochromatic light means that only one wavelength or a very narrow interval is used in the analysis. Therefore, it is possible to choose a wavelength where the compound has the highest absorbance. When a compound absorbs a photon, it is excited. Depending on the photon's energy, the effect on the compound is different. In the case of spectrophotometry, only the amount of light is interesting, but it is essential to know how the compound is affected by the energy from the absorbed light. Photons with a wavelength of 100-780nm (long-wavelength UV and visible light) are most commonly used because photons of this energy level excite electrons to a higher-energy orbital but do not break any bonds. The excited electron emits a photon of a longer wavelength and lower energy, and the compound returns to its ground state. Short-wavelength UV and radiation of higher energy break bonds and ionize the compound (89; 90).

One requirement for using spectrophotometry is that the compound contains a chromophore. A chromophore is an optically active structure that absorbs photons (90). The structure can be double or triple bonds, conjugated π-systems, aromatic groups, or pep-

[5]Illustration from "Quantitative Chemical Analysis", licensed under W. H. Freeman and Company (accessed October 12th, 2023) (89).

tide bonds. The presence of and/or combination of different chromophores results in a compound's absorptivity, ε, at a specific wavelength. All proteins are optically active, as they contain both peptide bonds and aromatic groups (in some amino acids), and are well suited for spectrophotometry (90).

In this thesis, the Agilent Technologies Cary 60 UV-Vis Spectrophotometer is used. It uses a xenon flash lamp, which emits light of wavelengths 190-1100nm. It is a double-beam photometer with the same layout as Figure 2.6.

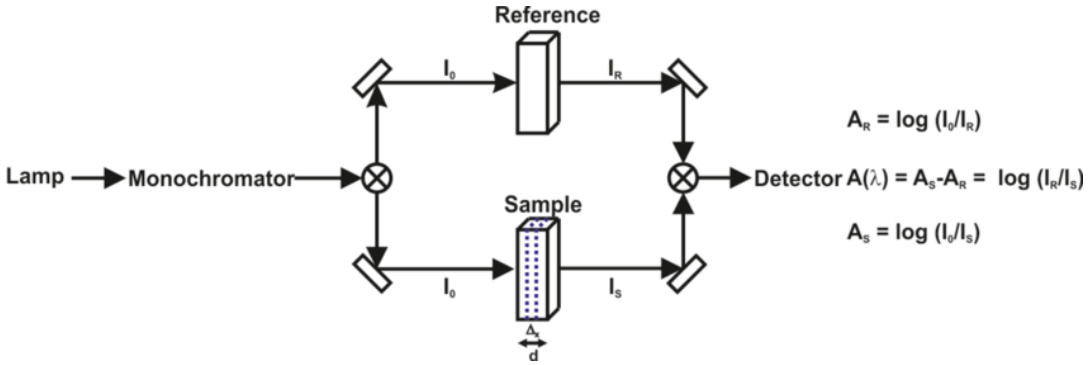


Figure 2.6: Schematic overview of a double beam spectrophotometer. Credit[6].

Figure 2.7: Agilent Technologies Cary 60 UV-Vis. Credit[7].

2.2.2 Enzyme assay for ADA2

Enzyme assays can function as both a quantitative and qualitative technique. It can be used to determine the presence, or lack thereof, of a specific enzyme in a sample by us-

[6]Illustration by Carrara for "Towards new efficient nanostructured hybrid materials for ECL applications" (accessed October 6th, 2023) (91)

[7]Illustration by Agilent Technologies (accessed October 6th, 2023) (92)

ing an indicator. The enzymes' catalytic function can be used as the indicator, e.g., by using a substrate or a reagent that partakes in the reaction and has a color change. For quantitative use, more exact analysis is required, e.g., photometric methods to measure the intensity of a color indicator (93).

Only qualitative enzyme assays are used in this thesis to indicate whether the isolation of ADA2 was successful. The assay used was developed by Iwaki-Egava et al. (2004) (54; 94).

The first step in the assay is the phosphorylation of Ino to hypoxanthine (Hyp) by purine nucleoside phosphorylase (PNP). Hyp is then oxidized to xanthine (Xa) by xanthine oxidase (XO) with nitroblue tetrazolium (NBT) as the oxidizing agent. Xa is further oxidized to uric acid. NBT has a blue color in its reduced state, indicating the presence of Hyp. Because the added substrate is Ado, no Ino or Hyp can be present in the assay without the initial ADA2 hydrolysis of Ado to Ino. This assay can also quantify ADA2 if the color intensity is measured at 540nm (94).

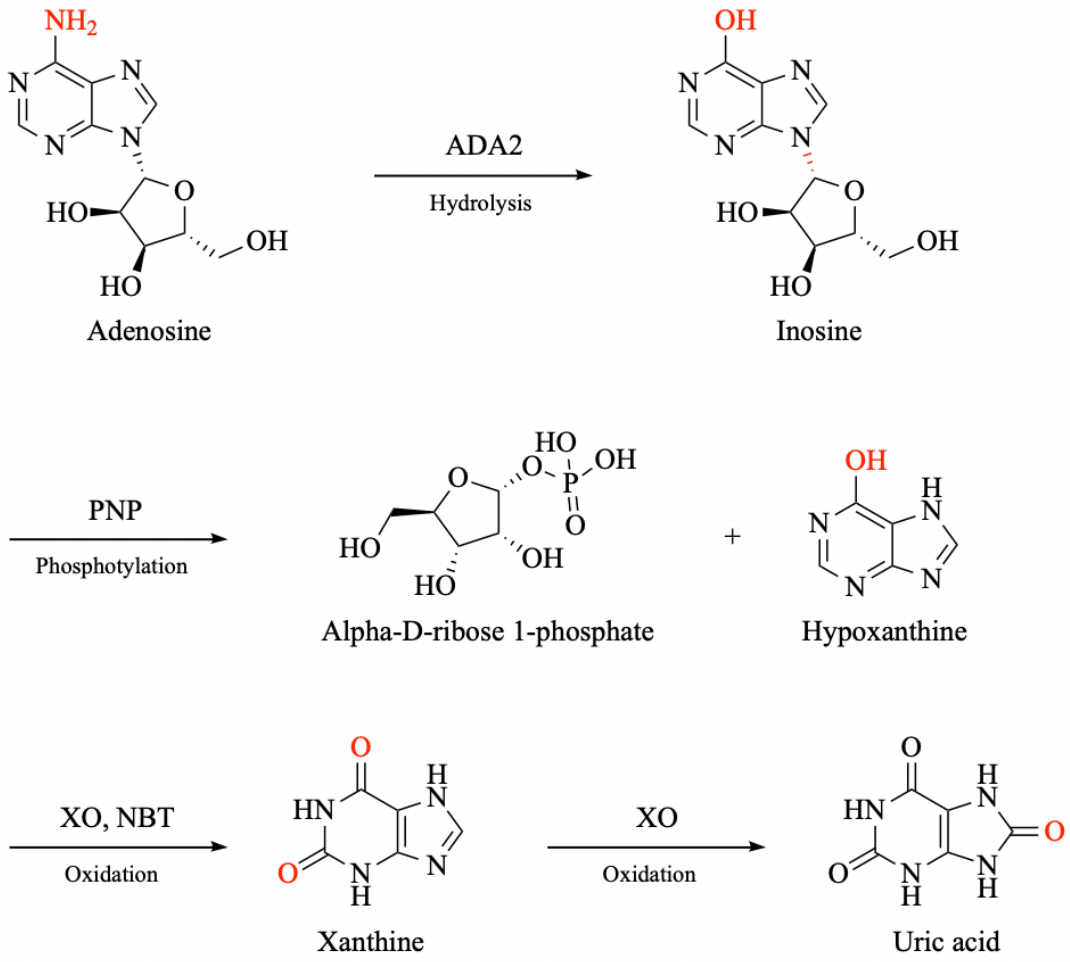


Figure 2.8: Reactions in the enzyme assay testing for ADA2-activity. Credit[8].

2.2.3 Enzyme assay for lysosomal α-mannosidase

In addition to the enzyme assay for ADA2 activity, an assay for α-mannosidse (LAMAN) was first conducted to test for the presence of glycoproteins and activity in general. The assay is simpler than the one for ADA2, as the used assay product functions as an indicator when changing the pH. The substrate, p-nitrophenyl-α-D-mannopyranoside, is added under acidic conditions, as this is the optimal pH for lysosomal α-mannosidase, and incubated at physiological temperature (37°C). The product has a negative charge and yellow color at basic pH, so a base, KOH, is added. It is unclear who first developed the assay, but it has been used in several studies on lysosomal α-mannosidse and α-mannosidosis (34; 95; 96).

[8]Illustration made by the candidate in ChemDraw based on method from "Adenosine deaminase 2 from chicken liver: purification, characterization, and N-terminal amino acid sequence" by Iwaki-Egawa et al., licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Unported license (Accessed October 12th, 2023)(54)

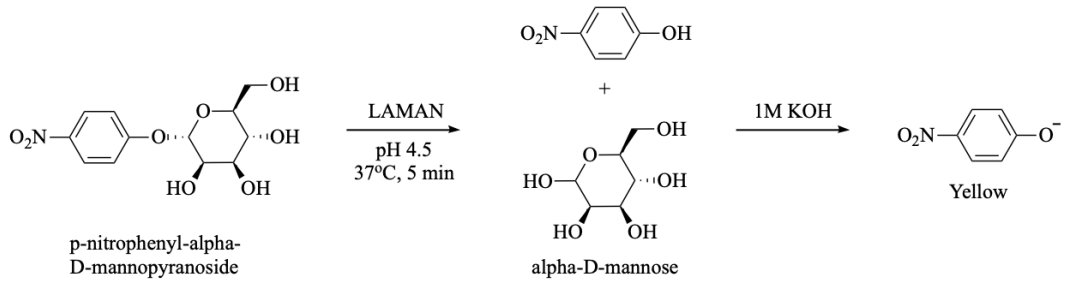


Figure 2.9: Reaction in the enzyme assay testing for α-mannosidase activity. Credit[9].

2.3 Tandem mass spectrometry of proteins

Mass spectrometry (MS) is a technique used to analyze atoms, molecules, and fragments of molecules and is a commonly used method in both organic and biological chemistry. The technique measures and separates atoms and molecules using their mass-to-charge ratio, *m/z*. Since it is *m/z* which is measured, the analytes have to have a positive charge, which can be +1, +2, etc. For analytes with a +1 charge, *m/z* is proportional to the monoisotopic mass of an analyte. Analytes with a +2 charge have a monoisotopic mass double of the *m/z*. The monoisotopic mass is the mass of an analyte when all atoms have a 1:1 ratio of protons and neutrons in the nucleus. It is also possible to use MS without monoisotopic mass, in these cases the abundance of each isotope has to be accounted for (97).

Being based on the *m/z*-ratio, MS uses the charge, the mass, and the ratio to separate the fragments. The fragments are sent into a chamber where they are accelerated by an electric field, E. As the fragments acquire a velocity, v, they are guided through an orthogonal magnetic field of strength, B, which causes the fragments to curve their path with a radius, r. The degree of curving depends on the mass and separates the fragments as they "land" the detector at different sites. In other words, the charge is used to achieve a velocity, and the mass is used to separate, as the charge is consistent for all fragments, i.e. +1, +2, etc. This gives the following equations used to determine the mass (97):

$$\frac{1}{2}mv^2 = zeV \Rightarrow v = \sqrt{\frac{2zeV}{m}} \tag{2.7}$$

$$\frac{mv^2}{r} = \frac{zeBr}{m} \Rightarrow v = \frac{zeBr}{m} \tag{2.8}$$

Combining Equations 2.7 and 2.8

[9]Illustration made by the candidate in ChemDraw

$$\frac{zeBr}{m} = \sqrt{\frac{2zeV}{m}} \Rightarrow \frac{m}{z} = \frac{eB^2r^2}{2V} \tag{2.9}$$

An ionization source fragments the sample and gives the fragments their charge. There are different ionization sources, e.g., *electrospray ionization* (ESI), which uses high voltage and highly charged droplets to ionize and evaporate the sample. ESI is a "soft" ionization source, meaning the molecules in the sample are not fully fragmented or fragmented at all (98). "Hard" ionization, such as *electron impact ionization* (EI), can also be used, but EI is more common for gas chromatography-MS (GC/MS) and smaller molecules. With EI, the sample, which is already in the gaseous phase, is bombarded with electrons that create radical ions and fragment the molecules (99).

In *tandem MS* (MS/MS, MS^2), there are two rounds of ionization and separation. This allows for even higher resolution and separation of ions. The sample is sent through an initial chamber where it is ionized and separated due to the created ions' *m/z*. This is MS1. A selection of these ions, within a specific *m/z*-interval, is further fragmented before a second round of separation and analysis in a second chamber. This is MS2. The ionization method for MS1 and MS2 differ. In MS1, it is more common to use ionization methods such as ESI, i.e., "soft" ionization. In MS2, *collision-induced dissociation* (CID) or photodissociation is more common. In the second round of fragmentation, the sample is already ionized and has a charge. Therefore, it is only necessary to use a method for fragmentation, not fragmentation and ionization. Hence, methods such as CID are useful, as they use a neutral atom or molecule to fragment the ions (98).

The sample preparation depends on the type of sample to be analyzed using MS/MS. When analyzing proteins specifically, the proteins are first digested using proteases, such as trypsin, which is used in this thesis. Trypsin cleaves at the C-terminal of arginine (R) and lysine (K). When using trypsin, it is possible to predict where the proteins will be cleaved and use this when looking for post-translational modifications, such as methylation and glycosylation (100). Specifically for glycosylation, it is important to know the location of potential cleavage sites and known glycosylation sites, as bound glycans can be steric hindrances for trypsin and hinder digestion. Nevertheless, as mentioned in Section 1.2.4, the presence and size of a glycan can vary from protein to protein, and some proteins might be cleaved at a specific site while others might not.

The ionization and fragmentation create specific ions, referred to as a-, b-, and y-ions (Figure 2.10). The monoisotopic masses of the ions of the predicted tryptic pep-

tides can be determined. Knowing the masses of the ions, and which ions can be made is important to know when analyzing spectra and identifying peptides - both in manual and programmed analysis.

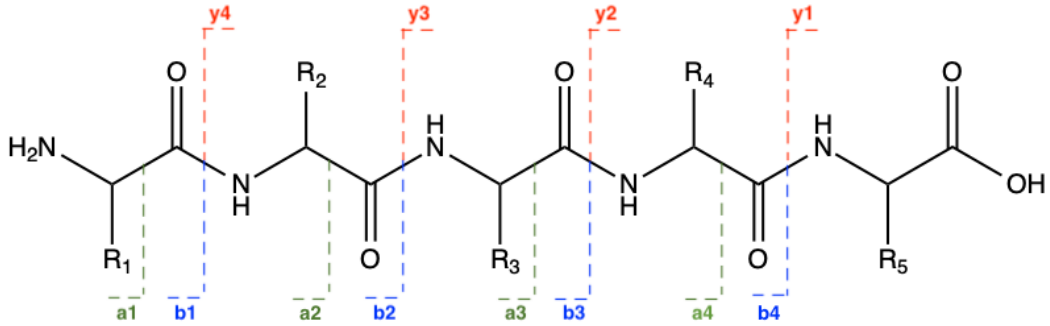


Figure 2.10: A-, b-, and y-ions of polypeptides in MS/MS. Credit[10].

2.3.1 MS/MS of N-linked glycans

MS/MS is the main method for determining the presence of N-linked glycans and their structures. Some methods have already been developed for determining the structures, and these generally include the separation of the glycans from the proteins to which they are bound. Two enzymes used to remove the glycans from their proteins are *peptide N-glycosidase F* (PNGaseF) and *endoglycosidase H* (EndoH) (48; 101). The difference between PNGaseF and EndoH is that EndoH is selective towards high-mannose glycans containing at least four mannoses, which is why PNGaseF is more common as it is more "general" (48).

When analyzing isolated glycans with MS/MS, a standard must be used to have a reference point for determining the amount and structures of glycans in the "unknown" samples. As with any other method, the quality of the standard is important to get reliable results. In addition, the standards must be representative of the samples, and in the case of glycans, there has to be a standard for every possible structure. As Metha (2016) presents in her N-glycan quantification, the availability of large enough amounts of standards with a large range of glycan structures is not yet there (101). Greiner-Tollersrud et al. (2024) also present some troubles with the existing standards and method of isolating the glycans (46):

- PNGaseF is not able to cleave N-glycans lacking α-mannose
- Lack of suitable standards, in terms of range in structures and amount

[10]Illustration made by the candidate in ChemDraw

Especially the analysis of processed glycans, such as *tomi*-glycans (Section 1.2.4), is difficult. In addition to the issues listed above, the possibility of *tomi*-glycans being misconceived as glycans damaged during sample preparation (46).

Due to the reasons listed above, other methods, both new and complementary, should be used when analyzing glycoproteins. One example is conducting MS/MS without isolating the glycans, which is done in this thesis. But, there is very little published research regarding this approach. With this approach, the monoisotopic masses of fragmented glycans, e.g., one GlcNAc or the disaccharide GlcNAc-Man are looked for in the spectra. A whole overview of monoisotopic masses is in Chapter 6, but some examples are presented in the table below:

Table 2.1: Monoisotopic masses of common N-glycans.

N-Glycan	**Monoisotopic mass [Da]**
Fuc	146.0597
Man	162.0528
Man-P	243.0265
GlcNAc – H_2O	185.0688
GlcNAc – 1 Da	202.0794
GlcNAc	203.09852
GlcNAc + 1 Da	204.09852
GlcNAc-Fuc	349.1455
GlcNAc-Fuc + 1 Da	350.1455
GlcNAc2	406.1655
GlcNAc2–Fuc	552.2355
GlcNAc2–Fuc–Man2	876.3355
GlcNAc2–Fuc–Man2 + 1 Da	877.3355
GlcNAc2-Fuc-Man3	1037.4255
GlcNAc2-Fuc-Man3 + 1 Da	1038.4255
GlcNAc-Man	365.1322
GlcNAc2-Man2	730.2755
GlcNAc2-Man3	892.3255
GlcNAc2-Man4	1054.3955
GlcNAc2-Man5	1216.4455

Still, as with peptides, the peak for the whole peptide and the single amino acids is found in a spectrum. Therefore, it is important to know the masses of whole glycans structures and likely fragments, i.e. mono-, di-, trisaccharides, etc. In addition, it is the concept of *tomi*-glycans and glycans more specific for lysosomal glycoproteins. This will be further investigated and discussed in Chapters 3 and 4.

2.4 Experimental

2.4.1 Materials

All materials used were supplied by the Institute of Biomedicine at UiT. Chemicals, enzymes, and substrates were obtained from Sigma-Aldrich, Merck, or GE Healthcare. Chicken liver was produced by Prior and purchased from Meny, Tromsø. The sheep spleen was purchased in bulk from Andslimoen.

2.4.2 Preparation of buffers

A 0.025M potassium phosphate buffer (KPB) with pH 7.2 and 0.02M acetic acid buffer (AcOH) with pH 5 were used as solvents. For practical reasons, 0.25 M KPB was made and diluted to 0.025 M before use. The buffer was made with the $H_2PO_4^-$/HPO_4^{2-}-pair by using potassium phosphate dibasic, K_2HPO_4, and potassium phosphate monobasic, KH_2PO_4. The desired pH of the KPB is equal to the pKa of the pair, therefore it can be made by using equal amounts of the two salts. To make 1L of the buffer, 34.05g KH_2PO_4 and 43.66g K_2HPO_4 were needed. The pH was adjusted with 6M NaOH. The same strategy was used for AcOH: making a 0.2M buffer and diluting it ten times before use. The buffer was made using potassium acetate, CH_3COOK, and concentrated AcOH, CH_3COOH. Since the desired pH of the buffer was 5 and the AcOH has a pKa of 4.7, we cannot use equal amounts of the acid and base. 34.54g potassium acetate, corresponding to a concentration of 0.35M, was dissolved in distilled water. 12mL concentrated AcOH was added, which gave a pH of 5.

ADA2 from sheep has a pI of 8.64 (Section 1.3.4), hence there was only a need for the 0.025M KPB, not AcOH. The buffer was set to a pH of 7.2, as for chicken. The same materials were used and 1L 0.25M was made and diluted to 0.025M before use.

2.4.3 Purification and analysis of ADA2 from chicken liver

An overview of the purification of ADA2 from chicken liver is visualized in Figure 2.11.

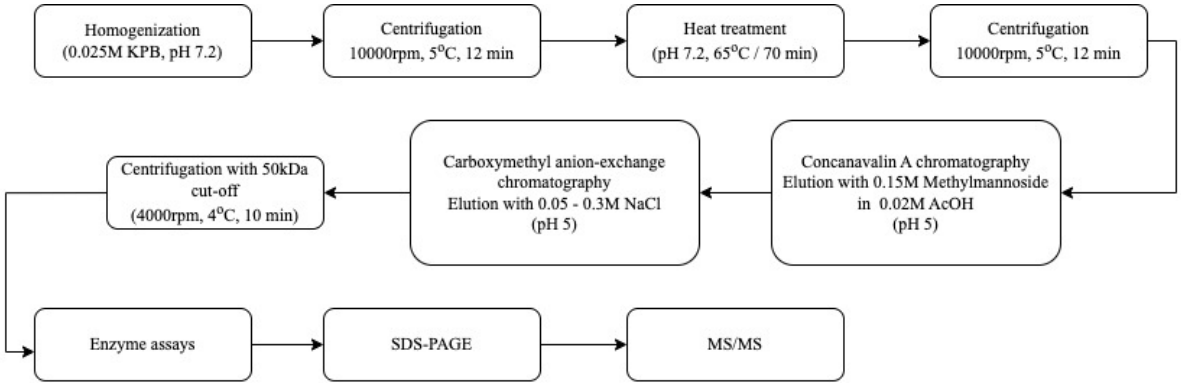


Figure 2.11: Flow chart of the purification process from chicken.

Step 1 Homogenisation

2.5kg Prior chicken liver was homogenized at 4°C with 0.025M KPB (pH 7.2). Approximately 1 cup (2.4dL) of chicken liver and 2 cups (4.8dL) buffer was added to a Waring blender and mixed for 1 minute. 2.5kg of chicken gave 8.2L homogenate, requiring 6.5L KPB. The homogenate was then centrifuged at 5°C and 10 000rpm for 12 minutes in an Avanti J-26XP Centrifuge. The centrifuge holds 6 tubes à 250mL, which were filled 3/4 with homogenate. After centrifugation, 6.3L supernatant was transferred to a new container and stored overnight at 4°C.

Step 2 Heat treatment

The 6.3L supernatant was divided into three Erlenmeyer flasks à 2L before being heated to 65°C. The flasks were heated for one hour before reaching the desired temperature, monitoring the temperature every 5-10 minutes. A spectrophotometric analysis at 280nm was taken of the supernatant before, at 50°C, and after completed heat treatment to monitor the protein concentration. For all three analyses', 10μL supernatant was diluted with 990μL distilled water, and a quartz cuvette was used (Table 3.1).

An Agilent Technologies Cary 60 UV-Vis spectrophotometer set to 280nm was used for all spectrophotometric analyses in the purification of both chicken and sheep.

After heat treatment, the supernatant was cooled at 4°C for 45 minutes before another round of centrifugation was done. The same conditions were used, 10 000 rpm for 12 minutes at 5°C. After centrifugation, 5.2L supernatant remained.

Step 3 Concanavalin A column chromatography

After heat treatment and centrifugation, 25mL of Con A was added to 5.2L supernatant and stirred overnight at 4°C. Before the column was started, the supernatant was without mixing for 30 minutes to allow the Con A to sink to the bottom of the container. 4.2L supernatant was immediately decanted and not added to the column. The remaining 1L supernatant, containing the Con A, was added to a filter-containing column. Additional centrifugation was tested to make the column run faster without success. The column took 2.5 hours. The flasks were washed with 0.025M KPB and added to the column.

Spectrophotometric analyses were done during washing with KPB to ensure all proteins were removed from the Con A. In these analyses, the supernatant sample was not diluted.

After filtration, 16.45mL Con A remained. To remove ADA2 from Con A, 200mL 0.02M AcOH (pH 5) was diluted, and 5.83g α-methylmannoside was added to the buffer. After the 200mL of AcOH, another 16mL of distilled water was added to the column to wash the remaining buffer from the Con A. The final 225mL solution was stored at 4 °C.

Spectrophotometric analyses were taken during the elution with AcOH, every 20-25 mL AcOH added (Table 3.2). This was done to monitor the protein levels in the buffer and inspect if the proteins dissociate from the Con A-molecules.

Step 4 Carboxy-methyl column chromatography

The column was prepared by adding 16mL CM-Sepharose and run through with the 0.02M AcOH (without α-methylmannosid). The column was filled twice to change the solvent and equilibrate the CM. The 225mL of supernatant from the Con A column was then added. This gave the first run-through (RT) fraction. Spectrophotometric analysis was done before and after the RT fraction to find the amount of protein bound to the column. A NaCl gradient was made to elute the remaning proteins from the column. The gradient consisted of 0.05M, 0.075M, 0.1M, 0.15M, and 0.3M NaCl in 50mL AcOH. Spectrophotometric analyses were also done on the five fractions (Table 3.3).

The six fractions were concentrated using 20mL Centrifugal Concentrators with a 50 000Da cut-off (all proteins with an MW < 50kDa are filtrated out). The first centrifu-

gation was done at 3500 rpm at 4 °C for 8 minutes. The remaining rounds were set to 4000 rpm at 4 °C for 8 minutes. The RT fraction had a significantly larger volume than the salt fractions, 225mL against 50mL. Therefore, the RT fraction was distributed over the salt fractions as they emptied. This resulted in a mixing of the fractions, which was not ideal and further discussed in Chapter 4. The next day, the fractions were further concentrated until the remaining volume was less than 1mL, transferred to corresponding Eppendorf tubes à 1.5mL, and stored at 4 °C.

Enzyme assays

The enzyme assay is based on the reactions in Figure 2.8 (Section 2.2.2). Before testing the isolated ADA2 with Ado, two tests with Ino and α-mannosidose (MAN) were done. The two tests were done to check if the assay was working and if the fractions had active enzymes. The solutions needed for the assay are:

- 10mM Ino
- 10mM Ado
- 0.24 units/mL PNP (premade)
- 5mg/mL XO
- 10mM NBT
- 0.02M KPB (pH 6.5)

First, the test with MAN was carried out by adding 250μL substrate, LAMAN, to 20μL of the fractions. The samples were incubated for 5 minutes at 37°C before 250μL 1M KOH was added. The samples obtain a yellow color if there is MAN activity, which they did. 500μL distilled water was added before spectrophotometric analysis at 405nm of the six fractions (Table 3.9).

Second, the test with Ino was carried out as the assay was set up for ADA2. A concentration gradient of 10 was made, with concentrations halved nine times starting with 40μL 10mM Ino. 40μL distilled water was added to all samples. The tenth sample was a zero sample of only 40μL distilled water. A 100μL stock solution with all reagents needed for the assay was then made, consisting of 60μL KPB, 15μL PNP, 15μL NBT, and 10μL XO. 10μL of the stock solution was added to the 2μL of gradient samples

before incubating for 5 minutes at 37°C. The samples turned blue, decreasing with decreasing inosine concentration, which means the assay works.

Before the assay on ADA2, the ADA2 samples were pre-incubated for 5 minutes at 37 °C. A total of 10 samples were made, including the six fractions and four control samples. See Table 2.2 for contents. All samples contained 2μL KPB. The 100μL stock solution was the same as for the Ino-test.

Table 2.2: Sample overview for enzyme assay of chicken liver.

Sample number	Fraction (2µL)	Substrate (2µL)
1	RT	Adenosine
2	0.05M	Adenosine
3	0.075M	Adenosine
4	0.1M	Adenosine
5	0.15M	Adenosine
6	0.3M	Adenosine
7	Blank	Adenosine
8	Blank	Adenosine
9	Blank	Inosine
10	Blank	Blank

After pre-incubation, 10μL stock solution was added. The samples were incubated for 7 minutes at 37 °C.

2.4.4 Purification and analysis of ADA2 from sheep

An overview of the purification of ADA2 from sheep spleen is visualized in Figure 2.12.

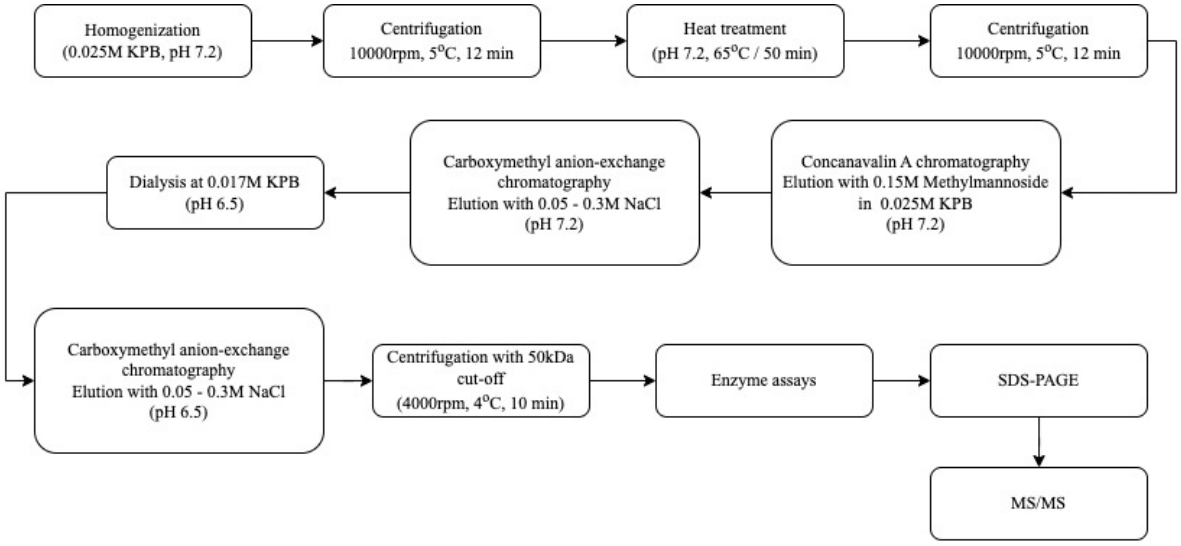


Figure 2.12: Flow chart of the purification process from sheep.

Step 1 Homogenization

1.7kg sheep spleen was homogenized in 5.3L KPB (pH 7.2) at 4°C with a Waring blender. The homogenate was centrifuged at 10 000 rpm and 5°C for 10 minutes. The final volume after centrifugation was 4L. The supernatant was stored at 4°C overnight.

Step 2 Heat treatment

4L supernatant was divided into two Erlenmeyer flasks à 2L and heated in a water bath set to 70°C. As with chicken, the supernatant was heated to 65°C. The heating took approximately 50 minutes, and the temperature was monitored every 5-10 minutes (Figure 3.3). The supernatant was then cooled at 4°C for 40 minutes before another round of centrifugation was done at the same conditions but for 12 minutes. The final volume was 3.4L. Spectrophotometric analyses were done before heat treatment, and before and after centrifugation (Table 3.4). The supernatant was diluted 100 times for all analyses, 10μL supernatant in 990μL distilled water. Distilled water was used as the zero sample.

15ml Con A was added to the supernatant and stirred at 4 °C overnight.

Step 3 Concanavalin A column chromatography

The stirring was stopped one hour before starting the column to let the Con A precipitate. Of the 3.4L supernatant, 3L was decanted off. The remaining 0.4L was centrifuged at 3000 rpm and 4∘C for 5 minutes to decant off as much supernatant as possible, re-

ducing the column volume. After adding all Con A to the column, KPB was used to wash. A spectrophotometric analysis of the eluted KPB was taken to check if all non-bound proteins were removed.

As with chicken liver, 5.83g α-methylmannoside was added in a surplus to dissociate the glycoproteins from Con A. 200mL 0.025M KPB was used instead of AcOH, due to the pI of ADA2 in sheep being 8.6. Spectrophotometric analyses were taken for every 20mL added KPB and of the final fraction (Table 3.5, Figure 3.4).

Step 4 Carboxymethyl column chromatography

20mL CM-Sepharose was used for the column. Before starting with the fraction from Con A, the CM was washed with two columns of KPB (pH 7.2). The same salt gradient as for chicken liver was also made: 50mL of 0.05M, 0.075M, 0.1M, 0.15M, and 0.3M NaCl in KPB. The fraction from Con A is the RT fraction. Spectrophotometric analyses showed that almost all proteins were in this fraction. Therefore, a dialysis of this fraction was done and the pH of the fraction was changed to a KPB of pH 6.5. Still, the salt gradient was eluted, giving five fractions. Spectrophotometric analyses of the fractions are in Table 3.6. These five fractions were concentrated with a cut-off at 50kDa, as with chicken liver. New volumes are in Table 3.10.

A second CM column was done with the dialyzed RT fraction with a pH of 6.5. 18mL CM-Sepharose was added to the column and washed with 0.02M KPB (pH 6.5). The RT-fraction was then eluted, giving a second RT-fraction. A salt gradient was made as previously but with KPB (pH 6.5) as the solvent. Spectrophotometric analyses were taken of these fractions before being concentrated (Table 3.7). New volumes are in Table 3.11.

Enzyme asssy

The same assay and tests were done for sheep spleen as with chicken liver. The MAN test was only done in the concentrated RT fraction from the first CM column. The test with the Ino gradient was not done by itself, only as a positive control during the assay of the fractions from the first CM column.

Two assays were done, one of the fractions from the first CM column and one for the second CM column. The same stock solution was used in both assays: 60μL KPB (pH 6.5), 15μL PNP, 15μL NBT, and 10μL XOD. Before adding the stock solution,

the fractions were pre-incubated for 30 minutes at 37 °C with Ado. Three control samples were included: A negative with Ado, a negative without Ado (only stock and KPB), and a positive with Ino. This gave a total of nine samples for each assay. The nine samples consisted of 2μL 0.02M KPB, 2μL fraction, and 1μL 10mM adenosine. After pre-incubating, 10μL stock solution was added, and the samples were further incubated at 37°C.

A third enzyme assay was performed on the 2. RT fraction. The assay consisted of five samples, all containing 2μL of the 2. RT, 2μL KPB (pH 6.5), and 10μL stock solution. Four different substrates were added to the samples: 10mM adenosine mono-phosphate (AdP), 10mM inosine mono-phosphate (IMP), 10mM Ado, and 10mM Ino. The fifth sample was a blank, negative control. All samples were preincubated for 15 minutes at 37 °C with the substrates before adding the stock solution.

2.4.5 Further enzyme tests

The fractions that showed activity from the previous enzyme assays were combined into one fraction. From chicken liver, fractions 0.05M and 0.075M were combined and concentrated to 75μL. From sheep spleen, fractions 0.1M, 0.15M, and 0.3M from the first CM column were combined and concentrated to 100μL. Both samples were concentrated with a 50 kDa cut-off. Before the two combined samples were concentrated, spectrophotometric analyses at 280nm were taken.

The enzyme assays were performed at different preincubation temperatures to determine the maximum temperature for ADA2 from each animal. They were performed at 60, 65, and 70° C with 10 times diluted samples and 65, 70, and 75° C with undiluted samples. In all assays, three controls - two negatives and one positive - were included as with the previous ones.

2.4.6 SDS-PAGE

X Cell SureLock with Invitrogen Power Ease 500 was used to perform the separation. The gel used was Invitrogen NuPAGE 10% Bis-Tris Gel with 10 wells à 5mm. All 10 wells were used; Wells 1 and 5 were standard solutions with bands of known masses, wells 2-4 were chicken samples, and wells 6-10 were sheep samples. Both chicken and sheep were diluted two, four, and ten times with distilled water, four samples for each

animal. The following solutions were added to the diluted samples:

- 3 µL Invitrogen NuPAGE LDS sample buffer (4x) (REF NP0007)
- 1 µL Invitrogen NuPAGE Sample reducing agent (10x) (REF NP0004)

All ten samples were heated to 83°C for 6 minutes before 13μL were added to the wells. X Cell SureLock was filled with 350mL, 20 times diluted Invitrogen NuPAge MOPS Running buffer (REF NP0001). The analysis was run for 45 minutes at 194V, 120mA, 23.3W at room temperature.

The gel was removed from the cell and washed and stained as described on Invitrogen Simply Blue (REF 46-5034) (Figure 3.11). After staining and washing, the bands at approximately 55kDA and 60kDa were cut out from the undiluted sample lanes and sent to MS/MS.

2.5 MS/MS

Jack-Ansgar Bruun at the University of Tromsø treated the four bands from SDS-PAGE with trypsin and analyzed them with MS/MS. Orbitrap Exploris was used for the analyses.

2.5.1 Analysis of MS/MS raw data using Proteome Discoverer

ThermoFischer's *Proteome Discoverer* (PD) was used to analyze all four samples. Fasta databases with the complete genomes of chicken and sheep and a database with common contaminants were used. All possible glycan structures were initially added to the program as *static modifications*. The masses were summarized by Jan Ole Olsen, found in Chapter 6. By adding the glycan masses as modification, it is possible to have the program search for peptide masses plus the glycan mass and, therefore, determine the presence of a peptide having a specific glycan bound to it.

The *Consensus* and *Processing* workflow in PD was set up as in Figures 2.13 and 2.14 after consulting Jack-Angar Bruun at UiT and The Proteomics Unit at UiB (PROBE). An analysis was set up where one search could look for ten glycans, resulting in 24 analyses per sample, i.e., 96 analyses.

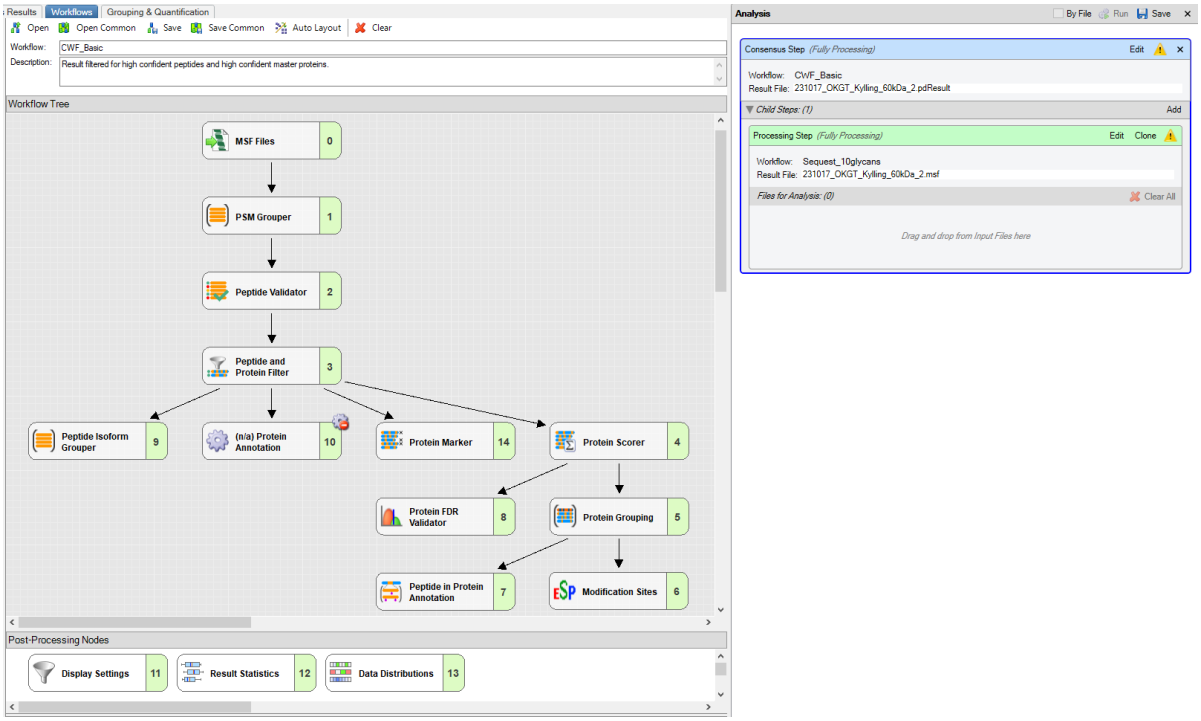


Figure 2.13: Workflow in Consensus step in Proteome Discoverer. (The warnings are due to an expired program license)

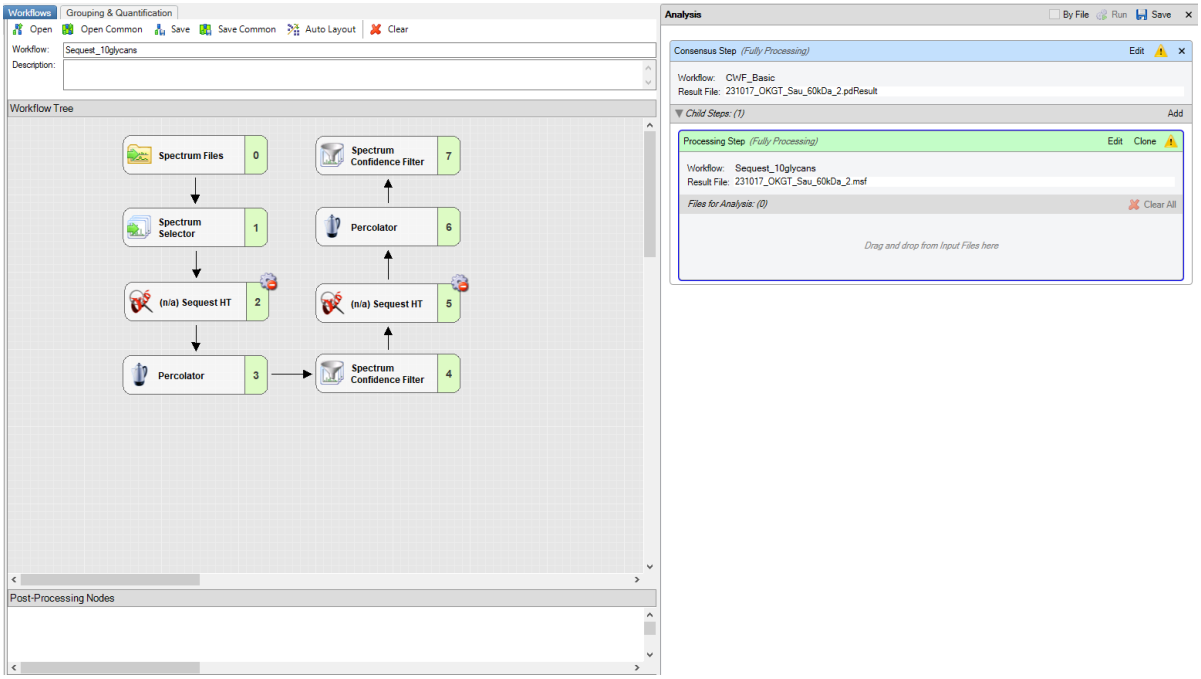


Figure 2.14: Workflow in Processing step in Proteome Discoverer. (The warnings are due to an expired program license)

Each workflow node can be modified by the user in terms of tolerances, minimum and maximum masses, modification, etc. The following table displays the chosen settings:

Table 2.3: Overview of parameter settings in PD workflows.

Node	Parameter	Value	Comments
MSF Files	PSM Filters - Maximum Delta Mass	0 ppm	Peptide-spectrum matches with a larger mass difference between the theoretical and the found peptide m/z are excluded from the final result. When set to zero, no delta mass filter is applied.
PSM Grouper	Peptide Group Modifications - Site Probability Threshold	75	Modifications with a lower site-probability than the specified threshold are not shown for the peptides
Peptide Validator	Validation Mode	Automatic	Control peptide level error rate if possible. Software first ascertains q-values and PEP's for PSMs. If yes, assigns the PSM confidence. If no, the software calculates the q-values and PEP's and assigns confidence.
Peptide and Protein Filter	Target FDR (Strict) for PSMs / Peptides	0.01	
	Target FDR (Relaxed) for PSMs / Peptides	0.05	
	Minimum Peptide Length	6	
	Minimum number of Peptide Sequences	1	
Protein FDR Validator	Target FDR (Strict)	0.01	All proteins with a q-value higher than the specified threshold will receive high confidence.
	Target FDR (Relaxed)	0.05	All proteins with a q-value higher than the specified threshold will receive medium confidence.
Spectrum selector	Min. Precursor Mass	350 Da	
	Max. Precursor Mass	5000 Da	
Sequest HT	Max. Peptide Length	150	
	Min. Peptide Length	6	
	Fragment Mass Tolerance	0.1 Da	
	Precursor Mass Tolerance	15 Da	
Percolator	Target / Decoy Selection	Concetenated	Only the best scoring PSM is written to the input file for percolator.
	Validation Based on	q-value	
	Maximum Delta Cn	0.05	Peptide-spectrum matches with a delta Cn better than or equal to 0.05 are used for percolator
	Target FDR (Strict)	0.01	
	Target FDR (Relaxed)	0.05	
Spectrum Confidence Filter	Spectrum Confidence	Worse Than High	

Due to the complexity of glycan analysis, the diluting effect they have on a protein, and the small amounts of ADA2 in the samples, it was decided to conduct a manual analysis of the raw data. Before manual analysis, the theoretical masses of potential tryptic peptides with glycans had to be determined.

2.5.2 Determination of glycan masses and patterns for MS/MS analysis

The monoisotopic masses of potential N-glycan structures, trimmed and untrimmed were determined using GlycanMass by ExPasy. All masses are in Chapter 6. The monoisotopic masses of untrimmed N-glycans were calculated by Jan Ole Olsen, formerly at the University of Tromsø. The masses of trimmed glycans were determined using *GlycanMass*. Variations of the trimmed glycans, such as +/- 1Da, were also included, as these can be present in the raw data.

2.5.3 Determination of peptide sequences and masses for MS/MS analysis

Using ExPasy PeptideCutter, PeptideMass and GlycanMass, the monoisotopic masses of the most likely peptides to be found with glycans are determined. The monoisotopic masses of the tryptic peptides with glycosylation sites in chicken and sheep are in Tables 3.14 and 3.16, respectively. The a-, b-, and y-ions of each peptide, with charges +1, +2, and +3, is found in Chapter 6 and were found using Peptide Predictor.

2.5.4 Manual analysis of MS/MS raw data for determination of peptide and glycan structures

Five samples were manually analyzed: the 55kDa and 60kDa samples from chicken and sheep purified in the fall of 2023, and a 60kDa sample from sheep purified in the spring of 2023. The earlier 60kDa sheep sample was purified by Greiner-Tollersud using the same method, but with more starting material.

The sum of the monoisotopic masses of tryptic peptides with glycosylation sites and potential glycans were looked for in the raw data. As the resolution of the MS/MS analysis is very good, it is sufficient to look at the theoretical, expected mass, perhaps +/- 1Da, as it is unlikely for a glycosylated peptide to deviate from the theoretical mass.

In the spectra, the molecular ion (M^+) mass must be the same as the mass of the investigated peptide-glycan complex. Only after finding the correct M^+ was it further investigated whether a glycan was present in the spectrum and if it was the correct peptide. Additionally, finding the correct M^+ and the presence of a glycan does not prove it is the correct peptide. Therefore, proof of it being a peptide from ADA2 must also be found.

An algorithm for manual analysis was developed and is summarized as follows:

1. Determine the monoisotopic mass of the peptide with a bound glycan
2. Find the spectra in the raw data file for the determined monoisotopic mass +/- 1Da
3. Investigate the spectra for peaks and neutral losses proving the presence of glycans
4. Investigate the spectra for peaks equal to the monoisotopic mass of the peptide and a-, b-, and y-ions

Step 3, proof of glycans, starts with finding two peaks in particular: *m/z* 204.09 and 366.14. The first peak at approximately 204.09 is the oxonium ion of GlcNAc plus 1Da. This peak generally has a very high intensity and is a clear sign of at least one GlcNAc. The second peak at approximately 366.14 is a GlcNAc and Man, without water, plus 1Da. This peak does not always have such high intensity but is still a clear sign of the presence of a glycan structure. In addition, peaks at 168.08, 185.07, and 202.08 are clear signs of glycans, being the masses of GlcNAc minus two and one water molecule and GlcNAc minus 1Da, respectively.

Step 4, proof of the peptide, consists of looking for the y-ion that is the peptide plus one GlcNAc. I.e., peptide mass plus 203.09Da. This peak is generally more prominent in glycopeptides than the peptide peak itself. In addition, a fragmentation pattern of loss of water, loss of Man, and loss of GlcNAC is often found. The neutral loss of this pattern is 18Da, 162Da, and 203Da, respectively. This fragmentation pattern can also be found as halves, meaning the presence of half the *m/z* of the y-ion and neutral losses of half the mass of water, Man, and GlcNAC, respectively, 9Da, 81Da, and 101.5Da.

Also included in step 4 is determining the amino acid sequence, i.e., specifying which peptide the glycan is bound to. As stated, different peptides can have the same mass. In some spectra, it is possible to determine the sequence as the neutral losses are

single amino acids, e.g., consecutive losses of masses belonging to amino acids. However, this pattern is absent in all spectra, making it difficult to determine which peptide the spectrum belongs to. In these cases, one cannot be certain which peptide it is, and the spectrum cannot be used conclusively. Another way to conclude which peptide is present is to look for a-, b-, and y-ions. Many spectra do not contain the full peptide but rather the smaller, fragmented ions. A third important strategy is determining the difference between the spectra's highest peak and the MH^+.

In addition to determining the glycan structures on peptides from ADA2, other known lysosomal glycoproteins were investigated. In chicken, glycoproteins *Vanin 1* and *N-acetylglucoseaminidase* were investigated. Chicken Vanin-1 is a lysosomal protein, while it is a GPI-anchored protein in humans. Still, the glycan structures found on chicken Vanin-1 were compared to the glycans on chicken ADA2.

Monoisotopic masses of a selection of glycans are found in Section 3.2. All glycan and peptide masses are in Chapter 6. Spectra thought to be of ADA2 peptides with glycans and spectra determined to be of ADA2 with glycans are also found in Chapter 7.

Chapter 3

Results

This chapter includes direct results from analyses conducted during the purification of ADA2 and results from analysis of MS/MS data.

3.1 Isolation and analysis of ADA2

This section presents the results from the purification process and enzyme assays on ADA2 from chicken liver and sheep spleen.

3.1.1 Purification from chicken liver

Temperature development during heat treatment

The expected temperature development of the supernatant was a rapid increase to approximately 50°C (Figure 3.1). At this temperature, most proteins start their irreversible denaturation. ADA2 is a rather heat-stable protein with an expected critical temperature of 75°C.

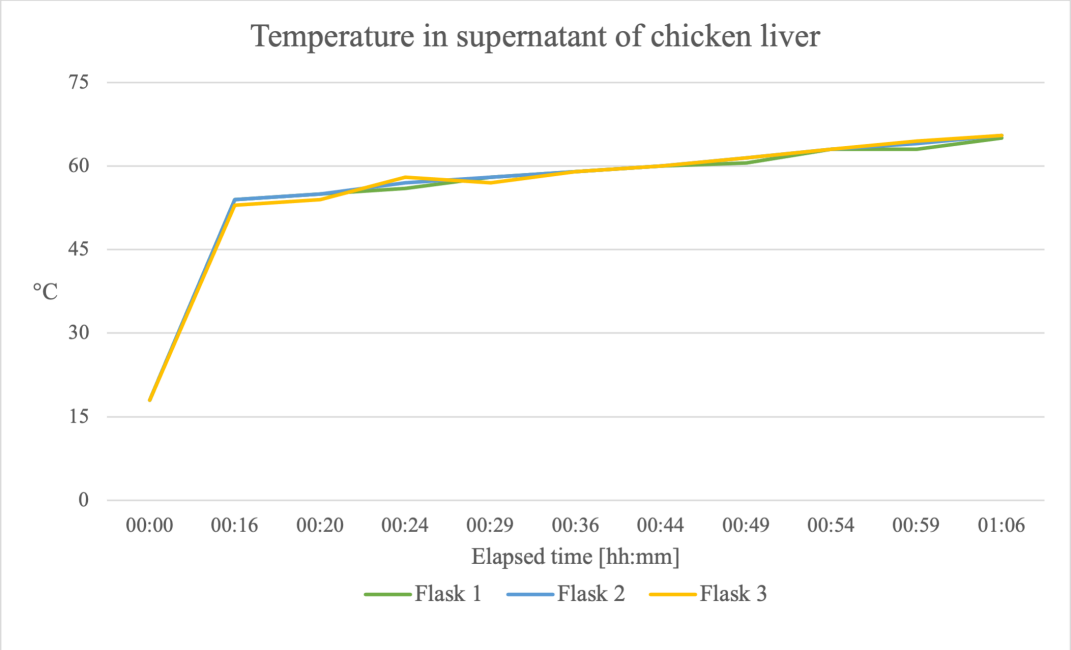


Figure 3.1: Temperature development of chicken liver. Credit[1].

Spectrphotometric analysis during heat treatment

The amount of protein in the supernatant was reduced due to the heat treatment, which was the desired result (Table 3.1). All samples were diluted 100 times (10μL supernatant in 990μL distilled water) and centrifuged before sample reading. As expected, heat treatment reduced absorption by more than 50%, corresponding to a large reduction in the amount of proteins in the supernatant. Assuming an absorbance of 1 is proportional to approximately 1mg/ml, the heat treatment reduced the amount of protein from 120mg/mL to 50mg/mL.

Table 3.1: Spectrophotometric analysis at 280 nm of supernatant from chicken liver. Distilled water was used as zero sample.

Sample type	Absorption at 280 nm
Zero sample	0.0008
Not heat treated	1.206
Heated to 50°C	0.714
Heated to 65°C	0.564

[1]Illustration made by the Candidate in Microsoft Excel

Concanavalin A column and spectrophotometric analyses

25mL of Con A was added before the column. After the column, 16.5mL of Con A remained, which means a loss of 8.5mL of Con A. A loss of both Con A and ADA2 was expected during the purification. In future purifications, keeping these losses even lower will be desirable, as losing 34% of the Con A and the bound proteins is too high.

Elution with α-methylmannoside in AcOH was monitored with spectrophotometric analyses (Figure 3.2 and table 3.2). A reduction in absorption from 0.56 in a 100-times diluted sample to 0.36 in an undiluted sample shows most proteins were dissociated during the elution process.

Table 3.2: Spectrophotometric data of elution with α-methylmannoside in AcOH in Con A column of chicken liver.

Added AcOH with α-methylmannoside [mL]	**Absorption at 280 nm**
0	0.0948
20	0.4899
50	0.6627
70	0.7895
100	0.6216
125	0.4436
130	0.3355
150	0.3187
175	0.2491
200	0.2373
Final fraction (V = 225 mL)	0.3614

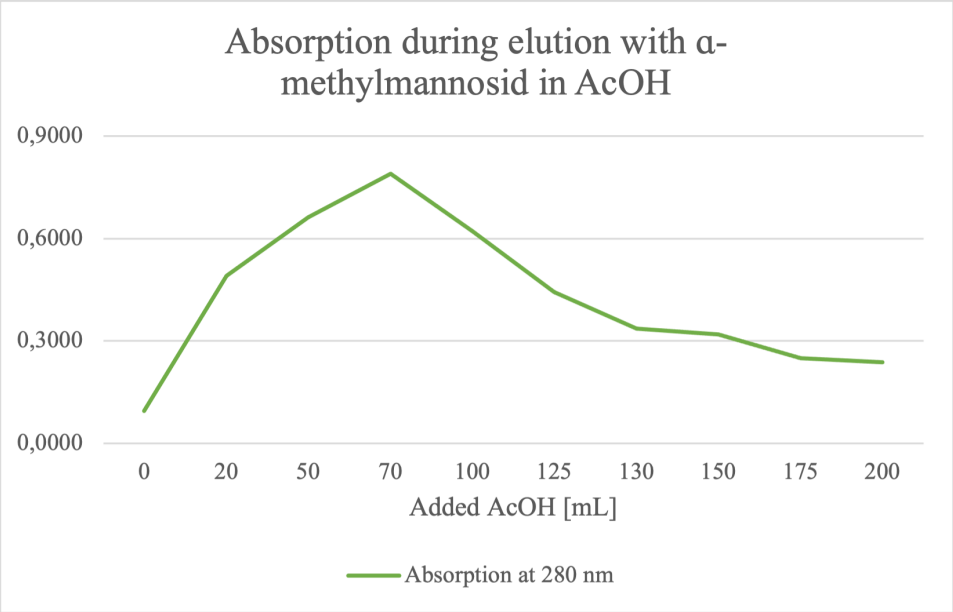


Figure 3.2: Spectrophotometric data of elution with α-methylmannoside in AcOH in Con A column of chicken liver. Credit[2].

CM column and spectrophotometric analyses

Spectrophotometric analyses of the supernatant before adding it to the column showed an absorption of 0.3352. The RT-fraction from the column had an absorption of 0.1177, meaning approximately 1/3 of the proteins from the Con A column were already eluted and did not bind to the CM (Table 3.3). The proteins in the column bound to CM are the ones with a positive charge at pH 5. Based on ADA2's pI, it is expected to elute with a lower salt concentration, i.e., 0.05M or 0.075M NaCl.

Table 3.3: Spectrophotometric analysis of fractions from CM column of chicken liver.

Fraction sample	Absorption at 280 nm
RT	0.1177
0.05 M NaCl	0.066
0.075 M NaCl	0.068
0.1 M NaCl	0.079
0.15 M NaCl	0.158
0.3 M NaCl	0.585

The majority of the proteins were eluted in the 0.3M fraction. As ADA2 is expected to be in the lower concentration fractions, this is favorable as fewer contaminating pro-

[2]Illustration made by the candidate in Microsoft Excel

teins are in the sample.

3.1.2 Purification from sheep spleen

Temperature development during heat treatment

As with chicken liver, it is expected the temperature in the supernatant will plateau at 50°C, and the increase rate will slow down. The supernatant reached this plateau quickly and reached 65°C quicker than the chicken supernatant, 47 and 66 minutes, respectively (Figure 3.3).

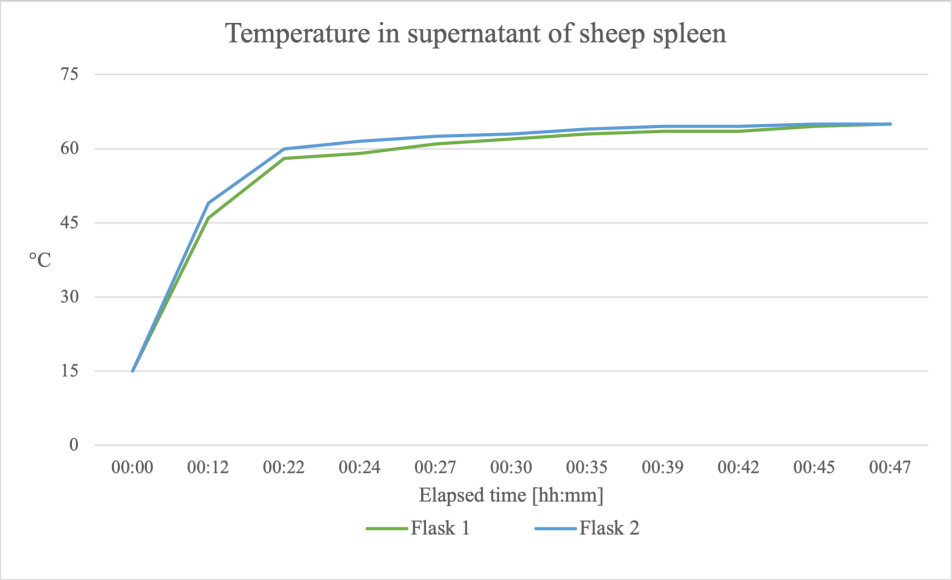


Figure 3.3: Temperature development of sheep spleen.

Spectrphotometric analysis during heat treatment

The heat treatment had the same effect on sheep as on chicken, removing approximately 50% of the proteins (Table 3.4). The samples were diluted 100 times with distilled water, which was also the zero-sample. Where the chicken supernatant was reduced from approximately 120 to 50mg/mL, the sheep supernatant was reduced from approximately 88 to 40mg/mL.

Table 3.4: Spectrophotometric analysis at 280 nm of supernatant from sheep spleen. Distilled water was used as zero sample.

Sample type	**Absorbance at 280 nm**
Zero sample	0.001
Not heat treated	0.880
Heated to 65°C	0.675
After heat treatment and centrifugation	0.402

Concanavalin A column and spectrophotometric analysis

It is clear from the spectrophotometric analyses of the eluted samples that most proteins were eluted from the column during the first 30mL of added KPB (Figure 3.4). At the same time, there is a considerable amount of proteins in the following samples, showing it was necessary with an eluting volume of 200mL.

Table 3.5: Spectrophotometric data of elution with KPB and α-methylmannoside in Con A column of sheep spleen.

Added KPB (pH 7.2) with α-methylmannoside [mL]	**Absorption at 280 nm**
0	0.0582
5	0.7535
30	0.6584
60	0.4892
80	0.4085
100	0.3318
125	0.2196
150	0.2047
160	0.1786
175	0.2833
180	0.2928
190	0.2920
200	0.2149
205	0.0972
Final fraction (V = 205 mL)	0.3729

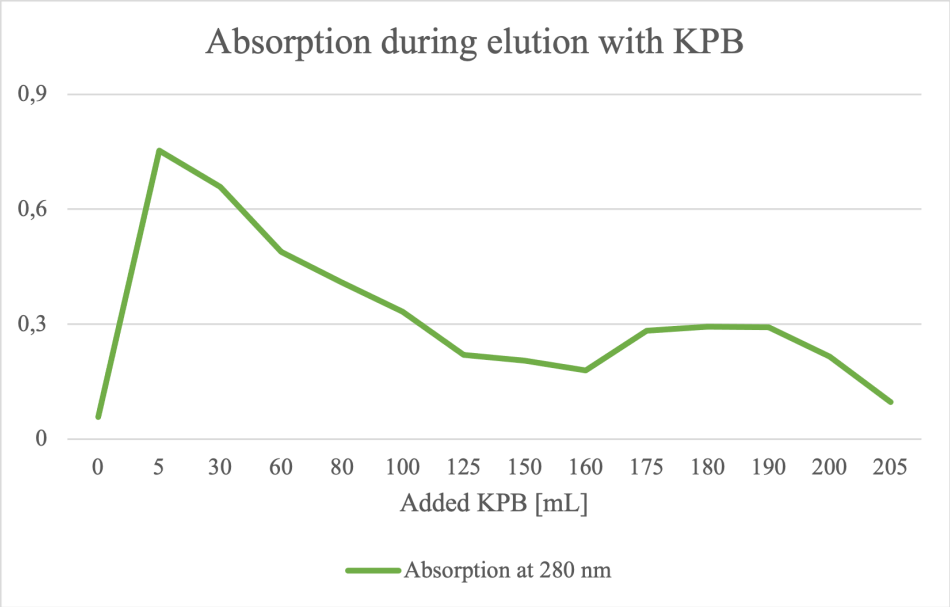


Figure 3.4: Spectrophotometric data of elution with KPB and α-methylmannoside in Con A column of sheep spleen. Credit[3]

The elution was inconsistent, with an increase in the final additions. This could have been caused by an uneven surface of Con A. The absorption did go down to the level of the zero-sample; however, there were some issues in the following steps. Using an additional 30-40mL KPB might have been necessary, despite increasing the volume for the following column, to ensure all proteins were dissociated.

CM column and spectrophotometric analysis

The RT fraction from the first CM column had a higher absorption than the elute from Con A, 0.4160 before the column and 0.4269 in RT fraction. At the same time, the salt gradient had a consistently low absorption, averaging 0.05 (Table 3.6). These low concentrations lead to the decision to do a second CM column on the RT fraction at a different pH to ensure ADA2 bound to the column. Nevertheless, the initial fractions were kept and concentrated.

[3] Illustration made by the candidate in Microsoft Excel

Table 3.6: Spectrophotometric analysis of fractions from the first CM column of sheep spleen.

Fraction sample	Absorption at 280nm
1. RT	0.426
0.05 M	0.0914
0.075 M	0.0276
0.1 M	0.0127
0.15 M	0.0098
0.3 M	0.0094

Dialysis of the RT fraction before a second column, reducing the pH from 7.2 to 6.5, resulted in more proteins binding to the column (Table 3.7). The dialyzed RT fraction from the first CM column had an absorption of 0.7525, a significant increase from the absorption of the first RT fraction, 0.4296. This was expected as dialysis often causes precipitation. The absorption in all fractions showed more proteins bound to the column, and all were concentrated with a cut-off at 50kDa.

Table 3.7: Spectrophotometric analysis of fractions from the second CM column of sheep spleen.

Fraction sample	Absorption at 280nm
2. RT	0.585
0.05 M	0.237
0.075 M	0.104
0.1 M	0.053
0.15 M	0.027
0.3 M	0.010

3.1.3 Enzyme assays

The following three sections present the results from the enzyme assays.

Enzyme assay of ADA2 from chicken liver

The six fractions from the CM column were concentrated to the volumes in Table 3.8. The final volume is important to remember when reviewing the ADA2 activity, as the fraction of higher volumes might have a more diluting effect on ADA2. Though this depends on the total amount of ADA2 in the fraction, it has to be considered. If two fractions have approximately the same activity, based on a visual test of the intensity of

the blue color created by NBT, it is important to take the original volume into account when finding which fractions contain the most ADA2. The fraction with the higher volume will have a higher concentration of ADA2 due to the diluting effect.

Table 3.8: Volumes after concentration of fraction from CM column of chicken liver.

Fraction	Volume after concentration [μL]
RT	220
0.05M NaCl	300
0.075M NaCl	500
0.1M NaCl	950
0.15M NaCl	300
0.3M NaCl	450

A test with α-mannosidase was performed before the assay on ADA2 to ensure that there were active enzymes in the fractions. This test does not need an added indicator because the product, α-mannosidase, is yellow. Therefore, a visual test is enough to see whether there is activity. Still, spectrophotometric measurements were taken where the RT fraction was used as the zero sample because of the spontaneous hydrolyzing of the substrate during storage. Fractions 0.1M and 0.15M had the most activity. Fraction 0.1M had a much higher original volume of 950 μL, compared to 0.15M of 300 μL. This means the 0.1M fraction likely has more α-mannosidase than the 0.15M fraction.

Table 3.9: Spectrophotometric analysis of α-mannosidase test.

Fraction	Absorption at 405nm
RT	0.1568
	0.1562
0.05M NaCl	0.221
0.075M NaCl	0.337
0.1M NaCl	0.368
0.15M NaCl	0.423
0.3M NaCl	0.397

The test with Ino was done to test the assay system and avoid unnecessary loss of ADA2. The Ino-concentration gradient halved the concentration in each of the ten consecutive samples, with the 10th sample being a blank with water. 2μL of the stock solution was added to each sample. After incubation, it was a visible reduction of the

intensity of the blue color from NBT. This corresponds to what was expected and shows that the assay system is functional.

The assay of ADA2 was performed. As seen in Table 2.2, there were four control samples. The seventh and tenth samples were replicas and control whether the assay works on Ado without the presence of ADA2. The ninth sample, with Ino and without a fraction sample, functions as a control of the assay - as the Ino-test previously done.

After 7 minutes of incubation, it was seen that fractions 0.075M, 0.1M, and 0.15M have the strongest color indication. Based on the original fraction volume, fraction 0.1M has the highest concentration of ADA2. Fractions 0.075M and 0.1M were chosen for SDS-PAGE and MS/MS analysis.

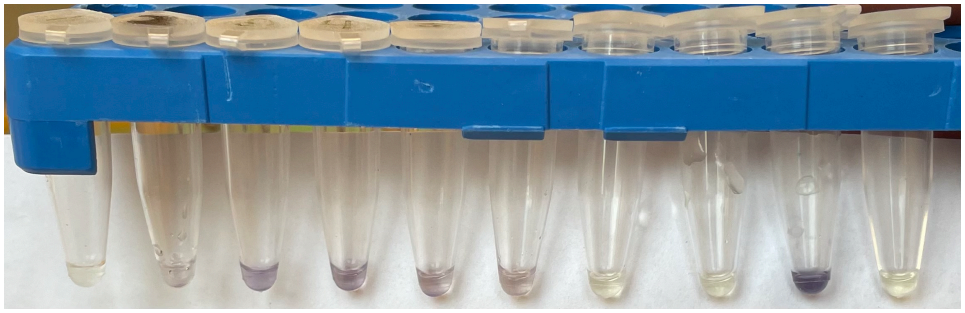


Figure 3.5: Enzyme assay for ADA2 activity in chicken samples. Samples from left to right: 2. RT; 0.05M; 0.075M; 0.1M; 0.15M; 0.3M; Blank w/ Ado; Blank w/ Ado; Blank w/ Ino and wo/ Ado; Blank wo/ Ado and Ino.

The enzyme assay showed that the purification was successful. However, the assay does not give any results on the exact amounts of isolated ADA2; the MS analysis will show this.

Enzyme assay of ADA2 from sheep spleen

The fractions from the two CM columns were concentrated to volumes seen in Tables 3.10 and 3.11. Only 50mL of the second RT fraction was concentrated because of the unlikelihood of ADA2 being in that fraction. This RT fraction includes all proteins with pI < 6.5. ADA2 having a pI of 8.6 makes it likely that ADA2 is bound to the column due to the large difference in buffer pH and pI, therefore being in one of the salt fractions.

Table 3.10: Volumes after concentration of fractions from the first CM column of sheep spleen.

Fraction	Volume after concentration (µL)
1. RT	Not concentrated
0.05 M NaCl	1000
0.075 M NaCl	300
0.1 M NaCl	400
0.15 M NaCl	510
0.3 M NaCl	200

Table 3.11: Volumes after concentration of fractions from the second CM column of sheep spleen.

Fraction	Volume after concentration (µL)
2. RT	210
0.05 M NaCl	585
0.075 M NaCl	500
0.1 M NaCl	450
0.15 M NaCl	350
0.3 M NaCl	300

The enzyme assay of the fractions from the first CM column showed that the most activity was in the unconcentrated RT-fraction (Figure 3.6). Furthermore, there was activity in the 0.05M fraction, likely due to an overlap with the RT fraction. Fractions 0.1M, 0.15M, and 0.3M showed more activity. This was also where activity was expected.

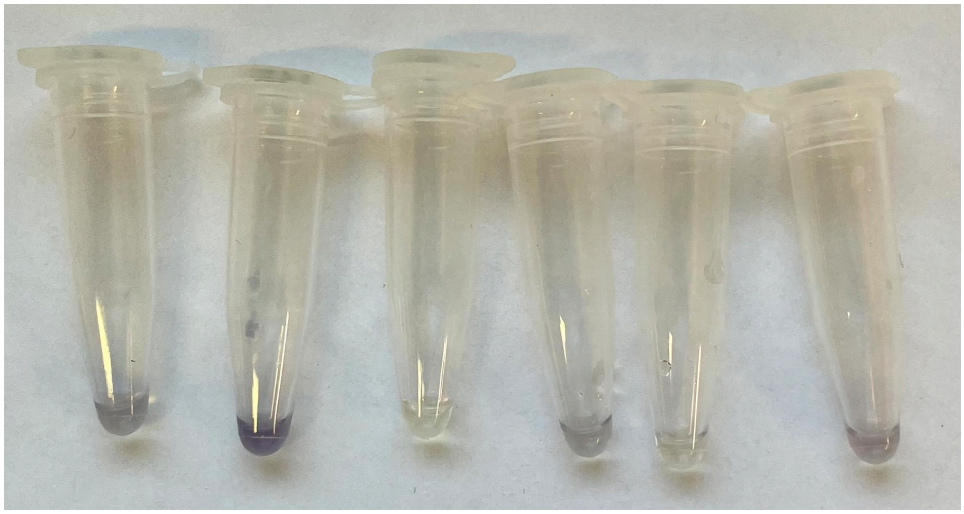


Figure 3.6: Enzyme assay for ADA2 activity in sheep samples from 1. CM column. Samples from left to right: 1. RT; 0.05M; 0.075M; 0.1M; 0.15M; 0.3M.

The assay of the fractions from the second CM column had a strong indication in the concentrated 2. RT fraction, as well as some in the 0.05M fraction, likely due to overlap (Figure 3.7). The other fractions indicated no activity. Lack of activity in these fractions implies ADA2 bound to the first CM column and is present in salt fractions from this column.

Figure 3.7: Enzyme assay for ADA2 activity in sheep samples from 2. CM column. Samples from left to right: 2. RT; 0.05M; 0.075M; 0.1M; 0.15M; 0.3M; Blank w/ Ado; Blank w/ Ino, wo/ Ado; Blank wo/ Ado and Ino.

The assay performed on the 2. RT fraction with different substrates indicated immediate activity in all four samples (Figure 3.8). Sometime after the stock solution was added, the blank sample without any substrate also obtained a slight purple color. Blanks in the previous assays did not change colors, even days after the assay was performed.

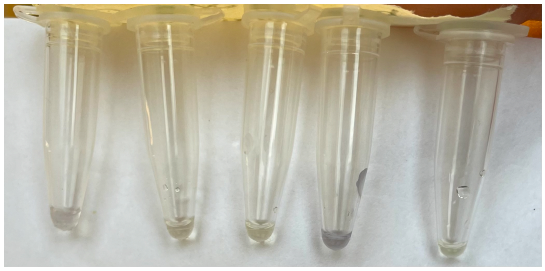
Figure 3.8: Enzyme assay testing for unknown protein activity in 2. RT from 2. CM column on sheep. Sample substrates from left to right: AMP; IMP; Ado; Ino; Blank.

Further enzyme assays to investigate heat stability and specific activtiy

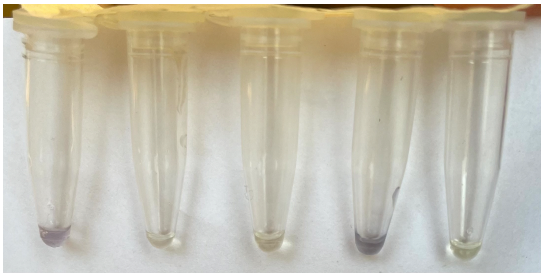
The spectrophotometric analyses taken before concentration showed 2.4387 and 1.7794 for chicken and sheep, respectively.

At 60 and 65°C with ten times diluted samples, sheep ADA2 shows minimal activity while chicken ADA2 shows a rather high amount of activity (Figures 3.9a–3.9c). Based solely on these two assays, it was not certain whether the temperature or the concentration led to the lack of activity.

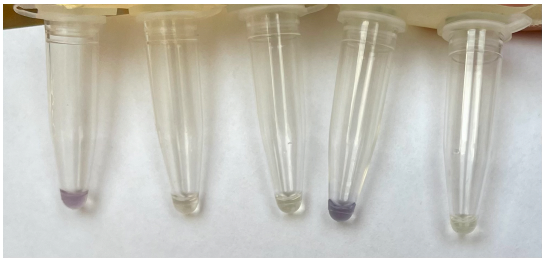
In all assays, the samples sequence from left to right: Chicken, Sheep, Blank w/ Ado, Blank w/ Ino, Blank.



(a) Enzyme assay with 60°C preincubation.



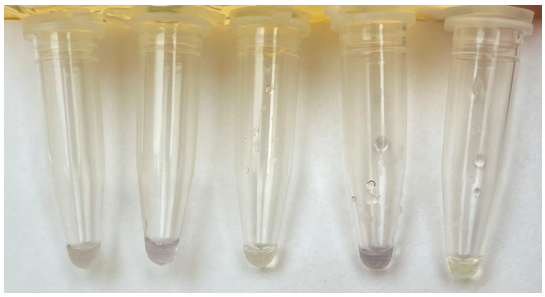
(b) Enzyme assay with 65°C preincubation.



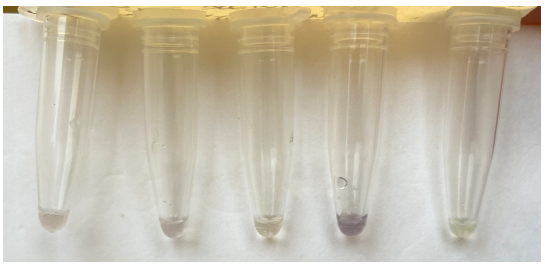
(c) Enzyme assay with 70°C preincubation.

Figure 3.9: Assays of 10 times diluted samples.

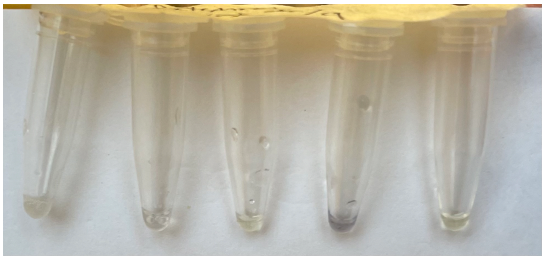
Assays done with undiluted samples at 65, 70, and 75°C indicated it was the concentration that caused the lack of sheep ADA2 activity at 65°C (Figures 3.10a–3.10c). At 65°C, sheep ADA2 showed more activity than chicken ADA2. At 70°C chicken ADA2 had no large change in the activity, while sheep ADA2 had a large decrease. At 75°C there was no activity in either sample, and there was visible denaturation in both samples.



(a) Enzyme assay with 65°C preincubation.



(b) Enzyme assay with 70°C preincubation



(c) Enzyme assay with 75°C preincubation.

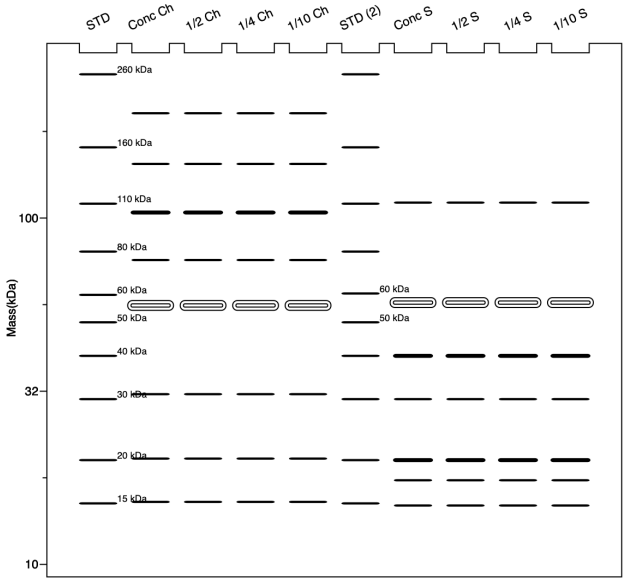
Figure 3.10: Assays of undiluted samples.

Chicken ADA2 has its critical temperature between 70 and 75°C, and sheep ADA2

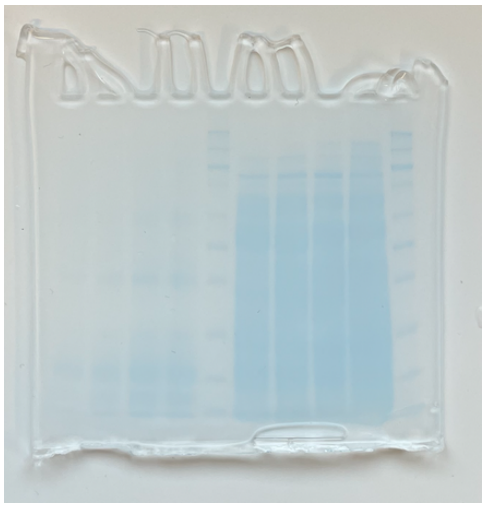
between 65 and 70°C, i.e., chicken ADA2 is more heat stable than sheep ADA2. Sheep ADA2's catalytic activity is also more concentration-dependent than chicken ADA2.

3.1.4 SDS-PAGE

Four different concentrations of each sample were used in case the bands were dragged. Both sheep and chicken presented bands around 50-60kDa, which is the expected band for the ADA2 monomer with a mass of 57kDa. These bands, and a little extra, were cut out and sent to MS/MS.



(a) SDS-PAGE of chicken and sheep. Bold bands are particularly prominent bands. Doubled bands are bands cut out and sent to MS/MS



(b) Photograph of gel from SDS-PAGE. Lanes from left to right are: 1/10 S; 1/4 S; 1/2 S; Conc S; Std; 1/10 Ch; 1/4 Ch; 1/2 Ch; Conc Ch; Std (Own photograph)

Figure 3.11: Illustration of SDS-PAGE

3.2 MS/MS

MS/MS results are presented in the following section. First, results from the analysis of chicken samples are presented, followed by the samples from sheep.

Below are the monoisotopic masses of common monosaccharides and fragmented glycans typically found in spectra of glycopeptides. These are the masses added to the monoisotopic peptide masses during manual analysis. Masses of non-fragmented glycans are found in Chapter 6.

Table 3.12: Monoisotopic masses of common glycans in MS/MS. Masses of monosaccharides are of them being fragmented from a larger glycan

N-Glycan	**Monoisotopic mass [Da]**
Fuc	146.0597
Man	162.0528
Man-P	243.0265
GlcNAc – 2 H_2O	167.0775
GlcNAc – H_2O	185.0688
GlcNAc – 1 Da	202.0794
GlcNAc	203.09852
GlcNAc + 1 Da	204.09852
GlcNAc-Fuc	349.1455
GlcNAc-Fuc + 1 Da	350.1455
GlcNAc2	406.1655
GlcNAc2–Fuc	552.2355
GlcNAc2–Fuc–Man2	876.3355
GlcNAc2–Fuc–Man2 + 1 Da	877.3355
GlcNAc2-Fuc-Man3	1037.4255
GlcNAc2-Fuc-Man3 + 1 Da	1038.4255
GlcNAc-Man	365.1322
GlcNAc2-Man2	730.2755
GlcNAc2-Man3	892.3255
GlcNAc2-Man4	1054.3955
GlcNAc2-Man5	1216.4455

3.2.1 Average masses of glycans

The peaks confirmed to be of glycans were collected, regardless of the peptide they were related to, and the mean mass and standard deviation were found. The data is incomplete because only spectra with precursor masses equal to the calculated masses for ADA2 peptides with glycans were analyzed. The following table presents the *m/z* of all glycan-related peaks in these spectra (Table 3.13).

Table 3.13: Mean observed masses and standard deviation of glycans found in MS spectra.

	Mean observed mass [Da]	Standard deviation
GlcNAc-2H2O (n = 124)	168.07250	0.068
GlcNAc-H2O (n = 118)	186.08401	0.065
GlcNAc (n = 167)	204.08688	0.0147
Man (n = 43)	163.09523	0.25
Fuc (n = 26)	147.10244	0.030
GlcNAcMan (n = 102)	366.134735	0.024
GlcNAcMan2 (n = 2)	528.1987	0.0004
GlcNAc2Man (n = 1)	569.76508	-
GlcNAcFuc (n = 1)	350.1698	-

3.2.2 Chicken ADA2

The monoisotopic masses of the possibly glycosylated peptides are listed in Table 3.14. All discussed spectra are in Chapter 7.

Table 3.14: Masses of tryptic peptides in chicken ADA2 with glycosylation sites.

Position	Sequence	Monoisotopic mass [Da]
175 - 187	**NVT**EFDHSLLR	1329.6750
172 - 187	QLR**NVT**EFDHSLLR	1726.9187
298 - 305	ML**NVS**QIK	931.5233
369 - 381	NILDALML**NTT**R	1373.7409
394 - 403	HPVAK**NLS**LK	1105.6680

Presence and abundance

ADA2 was only detected in a select few of the PD analyses of the raw data, where 10 different glycans were set. Peptides from ADA2 were not found with any glycans bound to them, but still only being "found" in some of the analyses. In addition, ADA2 was only found to be in the 55kDa sample, not in the 60kDa sample.

With the parameter settings displayed in Section 2.5.1, ADA2 was found with high confidence in analyses 2, 4, 5, 6, 11, and 13 in the 55kDa sample. ADA2 was found without glycans, but glycans number 11-13, 30-39, 40-48, 49-58, 99-108, 119-128,

respectively, (numbering corresponds to glycans in Chapter 6), were included as modifications. Analyses 12, 14, 16, 18, and 19 found ADA2 with medium confidence. Glycans set in these analyses were 109-118, 129-138, 149-157, 168-177, and 178-187, respectively. The peptide K.SLIEQSAVFSILK.R, position 92-104, was found in all analyses from the same spectra, #42651.

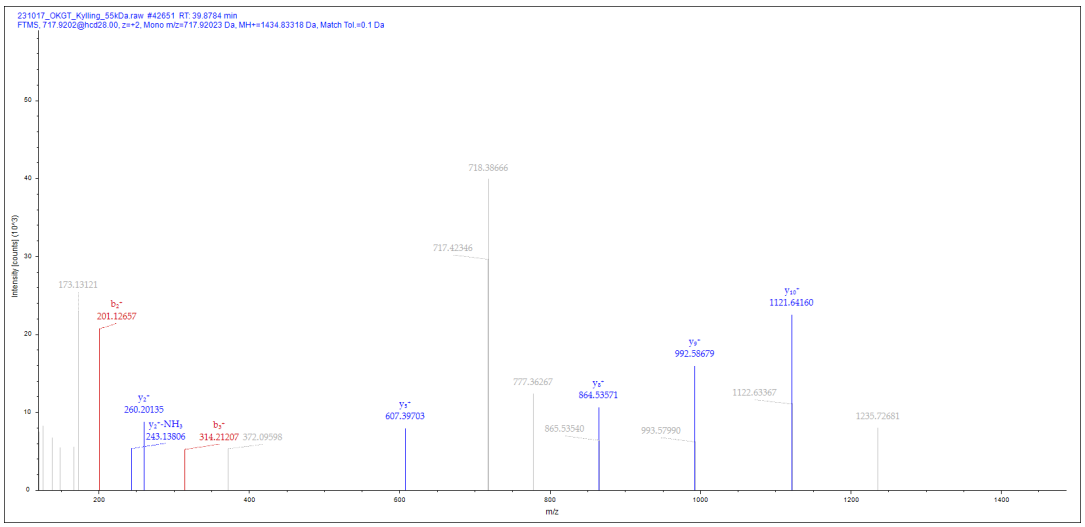


Figure 3.12: Spectra 42651 from 55kDa sample of peptide from ADA2.

Glycan structures

As PD analyses did not find any peptides with glycans, manual analysis was conducted instead, as explained in Section 2.5.4.

Potential spectra from both 55kDa and 60kDa samples were identified to start. The potential spectra were analyzed according to the algorithm explained in Section 2.5.4. Initially, proof of the presence of glycans was the priority. Most spectra did not contain either glycans or a correct peptide, but 7 spectra were promising Table 3.15.

Table 3.15: Potential spectra containing glycosylated ADA2 peptides. All spectra are found in Chapter 7

Sample	Position	Glycan	Spectrum	Mass [Da]	Precursor mass [Da]	Charge
55 kDa	300	GlcNAc2Man	25967	1823.84087	912.42407	+2
	399	GlcNAc	4542	1309.54839	855.27783	+2
	399	GlcNAc2Man2	4617	1836.75322	918.88025	+2
	399	GlcNAcFuc	4909	1455.61431	728.31079	+2
	377	GlcNAc2Man4	5420	2429.10001	810.37152	+3
60 kDa	300	GlcNAc2Man2	21699	1661.77764	831.39246	+2

The following spectra are from the 55kDa sample.

Spectrum #25967 is the sequence MLNVSQIK with GlcNAc2Man (Figure 7.1 and table 7.1). Ions b_2-H_2O and $b_2{}^{+1}$ are present at *m/z* 227.22121 and 245.13451, and the two peaks at *m/z* 931.50995 and 932.51654 correspond to the theoretical mass of the peptide. In addition, peaks at 1499.72827, 1337.66931, 1134.59082/1135.59521, and 931.50995/932.51654 have neutral losses of 162.05896, 203.07849, and 203.08087. *m/z* 1134.59082 is also the base peak, i.e., the most stable ion, and corresponds to the peptide with one GlcNAc, a generally very stable ion. The presence of GlcNAc2-Man was confirmed by the peaks at *m/z* 186.07880, 204.08899, 366.14453, and 163.06203, which are, respectively, GlcNAc-H_2O, GlcNAc, GlcNAcMan, and Man.

Spectrum #4542 is HPVAKNLSLK with one GlcNAc (Figure 7.2 and table 7.2). The $a_6{}^{+2}$-ion of *m/z* 310.18188 and $y_7{}^{+2}$ of *m/z* 387.33304 are present. In addition, peaks at *m/z* 1106.47522 and 1107.47546 are present. There is a 203.88039Da difference between the MH^+ of 1309.54839Da and the peptide mass of 1105.6680Da, corresponding to a GlcNAc bound to the peptide. There is no y-ion for the peptide with one GlcNAc, but the peak at *m/z* 204.09001 that confirms a GlcNAc.

Spectrum #4617 is HPVAKNLSLK with GlcNAc2Man2 (Figure 7.3 and table 7.3). $a_4{}^{+3}$ at 126.05705 and $b_1{}^{+1}$ at 138.05658, in addition to peaks at *m/z* 1106.47156, 1309.56494, 1310.55920, and 1311.56152, might indicate the peptide with one GlcNAc. The mass difference between the MH_+ of 1836.75322Da and peptide mass of 1106.47156Da is 730.28166, corresponding to the mass of GlcNAc2Man2. *m/z* 186.07941, 204.08896, and 366.14514 confirm the presence of both GlcNAc and Man.

Spectrum #4909 is also HPVAKNLSLK with one GlcNAcFuc (Figure 7.4 and table 7.4). The $a_4{}^{+3}$ and $a_6{}^{+2}$-ions are present at, respectivly, *m/z* 126.05709 and 310.18317. Peaks at *m/z* 1106.46960 and 1107.49951 are present, as in #4542 and

#4617, and in addition, *m/z* 1309.55530 and 1310.56616, as in #4617. The loss between these two are 203.08570, the mass of one GlcNAc, and the loss between MH^+ of 1455.61431Da and the peptide mass is 349.94631, the mass of GlcNAcFuc. The peak at *m/z* 204.08983 also confirms the presence of a GlcNAc, though there is no peak for Fuc.

Spectrum #5420 is NILDALMLNTTR with GlcNAc2Man4 (Figure 7.5 and table 7.5). y_3^{+3}-, y_2^{+2}-, and b_7^{+1}-ions are present at *m/z* 126.05698, 138.05664, and 771.39990, respectively. *m/z* 1577.79419 is the y-ion of the peptide with one GlcNAc. The peak at *m/z* 1374.72437 shows the neutral loss of 203.06712, i.e. GlcNAc. The base peak at *m/z* 789.40436 is half the mass of the highest peak, at *m/z* 1578.81421. The difference between MH^+ of 2429.10001Da and peptide mass of 1373.7409Da is 1055.35911Da, corresponding to the mass of GlcNAc2Man4. Glycan-confirming peaks are seen at *m/z* 168.06796, 204.08932, 366.14566, and 163.06284, respectively, GlcNAc-$2H_2O$, GlcNAc, GlcNAcMan, and Man.

The following spectrum is from the 60kDa sample.

Spectrum #21699 is MLNVSQIK with GlcNAc2Man2 (Figure 7.7 and table 7.7). The y_4-H_2O-, y_8-H_2O-ions are peaks *m/z* 457.23911 and 914.48584, respectively, as well as peaks for the peptide it self at *m/z* 931.51135 and 932.52002. The fragmentation pattern of MH^+ of 1661.77764Da and *m/z* 1499.73657, 1337.68213, 1134.59302, and 931.51135 confirms the presence of GlcNAc2Man2, as the losses are 162.04107, 162.05444, 203.08911, and 203.08167. The peaks are also present with +1Da, at *m/z* 1500.71777, 1338.67725, 1135.59302, and 932.52002, and parts of the pattern without H_2O, seen at *m/z* 1482.73535 and 1117.57080. The base peak pair at *m/z* 830.47278 and 831.47894 is half of MH^+. Additional proof of glycans is *m/z* 186.07863, 204.08914, and 366.14401.

Other findings

A full overview of the PD results is in Section 7.1.

Overall, there were more identified proteins in the 55kDa sample, with 142 identified proteins, whereas there were 128 identified in the 60kDa sample. In terms of raw data, the 55kDa has significantly more scans than the 60kDa sample, with 74997 and 54420 scans, meaning there is a higher protein concentration in the 55kDa sample.

In the 55kDa sample *Vanin 2* (Accession A0A8V0XM46) was the most abundant, with 411 PSMs (individual peptide-spectra), 74% coverage, and 35 identified peptides. The second most abundant proteins was *CN hydrolase domain-containing protein* (Accession A0A8VOXW73), followed by *N-acetylglucoseamine-6-sulfatase* (Accession F1NI04). The two have 211 and 127 PSMs, 40 and 66% coverage, and 21 and 30 identified peptides, respectively. This sample is contaminated by *Trypsin - Sus Scrofa (Pig)* (Accession P00761), being the 17th most abundant protein. A greater contaminant was *Keratin, type II cytoskeletal 1* (Accession P04264), a protein found in human skin. Other proteins identified in this sample are: *N-sulfoglucosamine sulfohydrolase* (Accession A0A8V1AHD9), *Alpha-galactosidase* (Accession A0A8V0YVV2), *Alpha-glucosidase* (Accession A0A8V1ALI6), *N-acetylgalactosamine-6-sulfatase* (Accession A0A8V0XWJ8), *Alpha-mannosidase* (Accession A0A8V0Y9V3), and *Alpha-L-fucosidase* (Accession E1BS94).

There were similar results in the 60kDa sample. The most abundant protein is *N-acetylglucoseamine-6-sulfatase*, with 130 PSMs, 64% coverage, and 28 identified peptides. The following two proteins are *Vanin 2* and *Carboxylic ester hydrolase* (Accession A0A8V0YB34) with, respectively, 130 and 87 PSMs, 50 and 56% coverage, and 20 and 28 identified peptides. This sample was also highly contaminated by *Trypsin - Sus Scrofa (Pig)*, which was the ninth most abundant. All remaining proteins mentioned to be in the 55kDa sample were also present in the 60kDa sample.

Generally, the two samples have similar content. Both contain *Carboxylic ester hydrolase* and *CN hydrolase domain-containing protein*, in addition to *N-sulfoglucosamino sulfohydrolase* (Accession A0A8V1AHD9), *Beta-galactosidase* (Accession A0A8V0ZI47), *Alpha-glucosidase* (Accession A0A8V1ALI6), *Alpha-galactosidase* (Accession A0A8V0YVV2), *Alpha-mannosidase* (Accession A0A8V0Y9V3), and *Alpha-L-fucosidase* (Accession E1BS94), to name a few.

3.2.3 Sheep ADA2

The monoisotopic masses of the possibly glycosylated peptides are in Table 3.16.

Out of the peptides presented in Table 3.16, the peptide 303-323 is most likely not found with any glycans due to its size, and a glycan-peptide complex will be in the outer borders of the detection limit. In addition, the peptide starting at 351 is more likely to be found than the one starting at 353. This is due to a possible glycosyaltion

Table 3.16: Masses of tryptic peptides in sheep ADA2 with glycosylation sites.

Position	Sequence	Monoisotopic mass [Da]
227 - 240	GLH**NVS**DFDNSLLR	1586.7921
303 - 323	ASLYPVYEL**NGT**VYSQEWLVR	2487.2554
351 - 358	TK**NES**LIK	932.5411
353 - 358	**NES**LIK	703.3985
426 - 437	NLLDAVIL**NST**R	1328.7532

at Asn-353 blocking a cleavage by trypsin, rather a cleavage in position 351.

Presence and abundance

Unlike the PD analyses of chicken, ADA2 was not found in any of the samples, neither 55kDa nor 60kDa, despite varying the glycans used for each analysis. In addition, there was contamination from *Bos Taurus* and *Sus Scrofa*, respectively, cattle and pig. Strangely enough, there does not seem to be any contamination from chicken, which is odd when the purification of chicken and sheep was performed back-to-back.

The purification performed in the spring of 2023 had several hits on peptides from ADA2. All spectra with their respective peptides are listed in the following table and found in Chapter 7.

Glycan structures

Manual analysis of the 55kDa and 60kDa samples purified in the fall of 2023 contained next to no ADA2 or glycan-containing spectra in general. On the other hand, the 60kDa sample from the spring of 2023 contained a vast amount of glycopeptides and ADA2.

Table 3.17: Spectra with peptides from sheep ADA2. Adapted table exported from PD.

Annotated Sequence	Positions	Theo. MH^+ [Da]	Spectrum	Precursor MH+ [Da]	Modifications
FVAGLAR	560-566	733.43554	15638	733.43364	
TYEEVAHNFAK	324-334	1308.62189	13604	1308.61704	
			13336	1308.61905	
TIFFVLR	265-271	895.54000	42916	895.53679	
QLAINSIR	530-537	914.54179	19066	914.53838	
VGGQLVLTR	100-108	942.57309	19005	942.56792	
SIQTAMGLR	360-368	976.52443	20425	976.51445	
			20526	976.52068	
EEELANQR	109-116	988.46941	4623	988.46520	
SIQTAMGLR	360-368	992.51934	9993	992.51506	1×Oxidation [M6]
TGSILPSMHFFQAK	131-144	1563.79881	36780	1563.79721	
VKFPGIVAGFDLVGR	369-383	1574.90532	47017	1574.90055	
TGSILPSMHFFQAK	131-144	1579.79373	27671	1579.78617	1×Oxidation [M8]
YSTLSETMK	538-546	1059.50269	12894	1059.50090	
SAMETWEER	548-556	1138.48335	16825	1138.48088	
SAMETWEER	548-556	1154.47826	10117	1154.47563	1×Oxidation [M3]
FPGIVAGFDLVGR	371-383	1347.74195	53349	1347.73955	
			53397	1347.74602	
EDTGHTLYDYR	384-394	1369.60188	14128	1369.59917	
DIPIEVCPISNQVLK	458-472	1724.92513	48089	1724.92766	1×Carbamidomethyl [C7]

Table 3.18: Potential spectra containing glycosylated ADA2 peptides. All spectra are found in Chapter 7

Sample	Position	Glycan	Spectrum	Mass [Da]	Precursor mass [Da]	Charge
231017 - 60 kDa	353	GlcNAc-1Da	2928	905.43730	453.22229	+2
230427 - 60 kDa	312	GlcNAc2FucMan3	46086	3524.50779	1175.50745	+3
	353	GlcNAc-H_2O	6040	1117.49089	559.24908	+2
	434	GlcNAc2FucMan-H_2O	17262	2059.89556	1030.45142	+2

All spectra are from 60kDa samples, though from purifications performed in the spring and fall of 2023.

Spectrum #2928, the only one from the purification done in the fall, is NESLIK with one GlcNAc-1Da (Figure 7.8 and table 7.8). Peaks at *m/z* 703.28834 and 905.45038, with a loss of 202.06214Da, indicate the presence of a GlcNAc. In addition, the $a_2{}^{+1}$-ion and b_2-H_2O are present at *m/z* 216.10609 and 226.02315/.12170, respectively. Peaks at *m/z* 168.06796, 186.07877, and 204.08868 confirms the presence of a GlcNAc.

The following three spectra are from the purification done in the spring of 2023.

Spectrum #46086 is ASLYPVYELNGTVYSQEWLVR with GlcNAc2FucMan3 (Figure 7.9 and table 7.9). The difference between the MH+ of 3524.50779 and the peptide mass of 2486.2554 Da is 1038.25239, corresponding to the mass of the glycan structure. There is also the neutral loss of 405.95383, from *m/z* 2892.20923 to 2486.2554, corresponding to the mass of GlcNAc2-1Da. Confirm the peptide are

peaks *m/z* 1657.87756 and 1157.54700, b_{20}^{+2} and b_{15}^{+1}, respectively. *m/z* 163.05974, 168.06531, 186.07658, 204.08644, and 366.13693 confirm the glycan, and are Man, GlcNAc-$2H_2O$, GlcNAc-H_2O, GlcNAc, and GlcNAcMan, respectively.

Spectrum #6040 is TKNESLIK with GlcNAc-H_2O (Figure 7.10 and table 7.10). A difference of 185.94979 Da between the MH+ of 1117.49089 and peptide mass 931.5411 corresponds to GlcNAc-H_2O. Peaks at *m/z* 147.11258, 158.09193, 243.13158, 416.20523, 542.30304, and 831.39636, respectively, ions y_1^{+1}, a_3^{+2}, y_2-NH_3, b_5^{+1}, and y_7^{+1}. *m/z* 186.08704 and 204.13474 confirms GlcNAc-presence.

Spectrum #17262 is with NLLDAVILNSTR GlcNAc2FucMan-H_2O (Figure 7.11 and table 7.11). The difference between MH+ of 2059.89556 and the peptide mass of 1327.7532 corresponds to this glycan structure. In addition, the presence of y_2^{+2}, y_{11}^{+2}, and a_{11}^{+1} at 138.05501, 607.27600/.77484, and 1214.55432, respectively, indicates the correct peptide. *m/z* 168.06525, 186.07423, and 204.08669, confirms GlcNAc. There are no peaks corresponding to Fuc or Man. The base peak at 1030.57068 is half the mass of the MH+.

Other findings

A full overview of the PD results is in Section 7.2.

Compared to chicken, very little protein is isolated from sheep spleen in the fall purification. Overall, the 60kDa sample had 48809 scans and 30 identified proteins, 14 of which were contaminants. The 55kDa sample had 41308 scans, though only 22 identified proteins and no contaminants.

The three most abundant proteins in the 55kDa sample were different Keratin-proteins, *Keratin 5, 3* and *type I cytoskeletal* (Accessions W5Q687, W5Q5V9, and W5Q6L8). The three have 6, 5, and 4 PSMs, 6, 3, and 13% coverage, and 2, 1, and 4 identified peptides. While the majority of identified proteins were other Keratin-type proteins, *N-acetylglucoseamine-6-sulfatase* (Accession W5NZ62) and *N-sulfoglucosamine sulfohydrolase* (Accession W5NUI6) were identified. However, only one scan was found for each, i.e., 1 PSM and 1 peptide, resulting in 1 and 2% coverage, respectively.

In the 60kDa sample, the six most abundant proteins were contaminants from pig

and human, with *Trypsin* from pig being the most abundant. The following protein, the most abundant protein from sheep, was *Alpha-2-macroglobulin* (Accession W5NSA6), with 6 PSMs, 6 identified peptides, and 6% coverage. *N-sulfoglucosamine sulfohydrolase* and *N-acetylglucoseamine-6-sulfatase* are also identified in this sample, more abundant than in the 55kDa sample. 4 PSMs of *N-sulfoglucosamine sulfohydrolase* was determined, corresponding to 2 peptides and 9% coverage. *N-acetylglucoseamine-6-sulfatase* had 3 PSMs, 1 peptide and 2% coverage. In addition, *N-acetylgalactoseamine-6-sulfatase* (Accession W5PVT3) was identified in 1 scan, corresponding to 1 PSM, 1 peptide, and 5% coverage.

The results above are from the purification done in the fall of 2023. In the 60kDa sample from the spring, the most abundant proteins are *Complement C4* (Accession W5NUX8) and *Alpha-2-macroglobulin domain-containing protein* (Accession W5NU00, W5NSA6), with 1090 and 1086 PSMs, 65 and 52 identified peptides, and 35 and 38% coverage. The sample is much larger than the previous two, with 402 identified proteins and 90504 scans, contributing to the higher number of PSMs per identified protein. As with the other samples, contaminants are highly abundant, though being from *Bos Taurus* (Cattle) in this sample. Other abundant proteins are *Alpha-2-macroglobulin*, *Alpha-mannosidase* (Accession W5PS45), *Alpha-glycosidase* (Accession W5NW80), *N-acetylglucoseamine-6-sulfatase*, and *Beta-galactosidase* (Accession W5QFF0). *Adenosine deaminse*, ADA2, is the 87th most abundant protein, with 20 PSMs, 14 identified peptides, and 23% coverage.

Chapter 4

Discussion

The following chapter will discuss the purification, enzyme assays, MS results, and the existing knowledge of ADA2, as well as other relevant topics. Possible flaws and successes of the purification and how they might have affected the MS results will be discussed. In addition, the enzyme assays will be discussed and related to the final section of the chapter, i.e. the function of ADA2. Furthermore, the results from the MS analysis will be discussed, with the purification process in mind, and how the results can be used to find the function and localization of ADA2. Lastly, the findings will be integrated with the existing knowledge and ideas regarding ADA2, as well as other related topics.

4.1 Purification and enzyme assays

Overall, there seems to be significantly less ADA2 in chicken liver and sheep spleen compared to pig brain and spleen. Previous purifications of the pig brain and spleen resulted in large amounts of ADA2 (65). The large difference may suggest some evolutionary drift of ADA2's localization and function in the different species, though this is further discussed in Section 4.3.

4.1.1 Chicken ADA2

Despite having few MS results, the purification of chicken liver was successful. The heating effect was as expected, with more than 50% of the proteins denaturating. The temperature increase was also as expected, reaching a plateau at approximately 50°C. It is important to remember that spectrophotometric analyses only give a quantitative estimate, not a qualitative one. In addition, denatured proteins can affect instrument detection, resulting in a higher absorbance due to molecules blocking and diverging the

transmitted light. Nevertheless, the analysis is used to get an idea of the amounts in the supernatant, not precise measurements.

The Con A column effectively isolated ADA2 and other glycoproteins, as all non-glycosylated proteins were removed. As mentioned in Section 3.1.1, there was a loss of Con A, and it is reasonable to assume there was a loss of ADA2, as well. During the elution with α-methylmannosid there was residual absorption when the elution was ended, it was stopped at 0.2373. Ideally, the elution would continue until analysis showed an absorption of 0 or closer to the initial zero sample of 0.0948. As seen in the MS results, there is little ADA2 in the final samples, and choosing to end the elution early might be a contributing factor. Nonetheless, the difference in ADA2's and α-methylmannosid's affinity to Con A is large enough to assume ADA2 quickly dissociated from the Con A molecules.

The CM column was also as expected, with most proteins eluting with the solutions of higher salt concentrations. ADA2 was expected to be present in the lower concentration fractions, and the enzyme assays confirmed this. Based on the color indication and the fraction volume, 0.075M and 0.1M were chosen for SDS-PAGE.

The 0.15M fraction also indicated activity. However, it was not chosen for SDS-PAGE as the fraction had a low volume, 300μL, and double the absorption before concentration compared to the chosen fractions (Table 3.3). The 0.075M and 0.1M factions had volumes 500 and 950μL, respectively. Therefore, it was concluded that the fraction was likely highly polluted with other proteins, which is undesirable for SDS-PAGE and MS analysis. In addition, a similar color indication and a lower volume indicate that this fraction has a smaller amount of ADA2. In retrospect, it should have been included in the SDS-PAGE despite having a smaller volume and lower ADA2 concentration. Including this fraction would have increased the amount of other proteins and possibly contaminants. But, seeing as there was a large loss of Con A and an early stop for the CM-elution, the overall loss of ADA2 became larger than necessary.

4.1.2 Sheep ADA2

The sheep samples generally contained less protein and ADA2 than the chicken samples. 1.7kg sheep spleen was homogenized, and 2.5kg chicken liver was used. As this is not a large difference in starting material, and the spleen is a more macrophage-rich organ, one would expect more ADA2 from sheep spleen than chicken liver - if sheep

ADA2 expression is controlled by a macrophage-specific promoter. If not, one would expect approximately the same amounts.

The columns performed on sheep were more problematic and likely affected the final MS results. Nevertheless, the purification was successful as ADA2 was isolated and in the fractions initially predicted. Yet, one of the problems was the uneven dissociation of glycoproteins from Con A. This might be due to an uneven Con A surface, and it is possible that a significant amount of proteins were not included in the final fraction. Another problematic aspect was that the same batch of Con A was used for both purifications. It is possible to reuse the Con A if it is thoroughly washed and used on the same species. However, using it on different species is not optimal or recommended. In addition, it is very difficult to clean the Con A completely, as there will always be some glycoproteins that are not disassociating.

The CM columns, first performed at pH 7.2 and then at pH 6.5, were also somewhat concerning. The fraction from Con A had a higher absorption on the day the column was performed, meaning some proteins denatured during storage. In addition, the first RT fraction with pH 7.2 showed that no proteins had bound to the column, as the absorption in all salt fractions was below 0.1. This could be below or at the instrument's detection limit (not specified in the instrument brochure), meaning it is impossible to differentiate between detector noise and signals. Therefore, it was assumed the fractions were blanks and no proteins bound to the column.

Dialysis to reduce the pH to 6.5 was performed to increase the charge of ADA2 and create a stronger bond between the protein and the CM molecules. This also meant that ADA2 would be in salt fractions of higher concentrations, i.e., 0.15M or 0.3M, not 0.075M and 0.1M as with chicken. The dialysis caused, as expected, visible protein denaturation, affecting the spectrophotometric analyses, which increased from 0.42 to 0.75. Yet, the absorption in the salt fractions was still low.

However, the previous purification of sheep spleen in the spring had similar tendencies regarding the low amounts. As the enzyme assays showed, the initial expectations of where ADA2 would be were also correct. In other words, the dialysis and second CM column were unnecessary.

Initially, the enzyme assays gave contradicting results. Fractions from the first CM column indicated ADA2 activity in all fractions except the 0.75M fraction. At the same

time, the control samples (not shown in Figure 3.6) did not suggest any fault in the assay system. Based on the buffer pH of 7.2 and ADA2's pI of 8.6, there should not have been any indication in the RT fraction, but clearly, there was. A possible reason was that the buffer affected ADA2's affinity to the column, but with a low buffer concentration of 0.025M, this is not likely. A second possibility was the presence of another protein with pI < 7.2 with similar activity to ADA2 or a protein using NBT as an oxidation agent in another reaction. The activity indicated in 0.05M is likely due to an overlap with the RT fraction, not protein activity. The fractions were collected manually, and some overlap was expected.

Fractions 0.1M, 0.15M, and 0.3M from the first column were expected to contain activity. 0.15M had a much weaker indication, which can result from a higher sample volume compared to the other two fractions (Table 3.10). Based on the indications observed in 0.1M and 0.3M, the higher sample volume is most likely the cause. A higher volume dilutes the protein concentration, resulting in less protein per aliquot used in the assay.

Only the RT and 0.05M fractions from the second CM column indicated activity. As previously stated, if ADA2 were present in this column, it would be eluted with 0.15M or 0.3M fractions. The lack of indication in these fractions shows that ADA2 was isolated in the first CM column. As with the first CM column, the indication in 0.05M is likely due to an overlap with the RT fraction. However, the cause of the activity was still unknown.

An additional assay, including AMP and IMP, was done to investigate the cause of the worrisome activity in the 2. RT fraction. The assay gave surprising results as all the samples gave indications of activity. In addition, the blank sample also changed color after some time. The blanks in previous assays did not change colors, even days after they were performed. This means NBT was stable, and the negative controls and blanks could be trusted.

AMP and IMP are known not to be substrates for ADA2, meaning there should not be any indication in these samples. If ADA2 were the only protein that could cause an indication, only samples with Ino and Ado would be positive samples. In other words, there was a protein in the 2. RT fraction, which either has AMP as a substrate or directly affects the assay indicator, NBT. Alternatively, there is the possibility of PNP converting Ado to Adenin, in addition to the conversion of Ino to Hypoxhanthine. There is no

evidence that PNP has Ado as a substrate, which means there must be a protein with Adenin as a substrate and an enzymatic reaction where NBT is used as the oxidizing agent. However, this does not explain the samples with AMP and IMP, as PNP does not convert IMP. Rather, it suggests that the most likely explanation is a protein or another compound in the 2. RT fraction, such as a cofactor that uses NBT directly. This explanation would also apply to the alleged activity in the 1. RT fraction and both 0.05M fractions.

4.1.3 Enzyme assay for heat stability

Additional enzyme assays testing the heat stability showed that chicken ADA2 could withstand at least 70°C. This means the homogenate can be heated to 70°C, denaturing even more proteins without risking a loss of ADA2. It is desirable to remove as much protein in each step of the purification, and having as high a temperature as possible is an efficient way to do this. The activity disappeared when the sample was incubated at 75 °C, meaning the critical temperature is between 70 and 75°C. Further assays should be performed to find the exact temperature. It is crucial to know whether it is, e.g., 71 or 74°C, especially if the next purification is to be performed with heat treatment at 70°C. For example, heating to 70°C and ADA2's critical temperature being 71°C leaves little room for an accidental temperature increase.

Sheep ADA2 is slightly less heat stable than chicken ADA2. At 70°C, almost all activity had disappeared, while at 65°C, sheep ADA2 had a stronger indication than chicken ADA2. Based on the dramatically decreased indication when increasing the temperature to 70°C, its critical temperature might be closer to 65°C. As with chicken, it is critical to find the exact temperature because this could have affected this purification. If the critical temperature of sheep ADA2 is closer to 65°C, some ADA2 might have been denatured during the heat treatment step because it was stopped at 65°C. This might also have contributed to the low concentrations in the columns and MS analysis.

High heat stability indicates strong intramolecular forces and high stability in a protein. Strong intramolecular forces also mean a protein has a better ability to not be degraded or affected by other enzymes or harsh conditions. ADA2 being heat stable, therefore, indicates it is well suited for, e.g., lysosomal conditions, both in terms of not being affected by proteases and acidic conditions. This will be further discussed in Section 4.3.

4.1.4 Enzyme assay for specific activity

Assays were also performed to investigate the specific activity for Ado. The diluted samples from chicken and sheep were tested at different temperatures, and it was clear that chicken ADA2 had a much higher activity for Ado across different temperatures and sample concentrations. The spectrophotometric analysis taken before concentrating with a 50kDa cut-off indicated that the chicken sample contained almost twice the amount of protein. Although the samples are not solely ADA2, it is reasonable to assume the sample from chicken contains more ADA2 than the sheep sample. In addition, it was also evident that there was generally more ADA2 and proteins obtained from chicken than sheep. Nevertheless, diluting the sample from chicken did not affect its activity, but it highly affected the sheep sample. The assays of each species indicate that chicken ADA2 has a higher specific activity for Ado than sheep ADA2.

As presented in Section 1.3.4 and Figure 1.13, sheep ADA2 has a glycosylation site at the opening of the binding pocket. Assuming most of the ADA2 molecules are glycosylated at this site, with glycans of varying sizes, a large part of the ADA2 dimers might not bind Ado effectively and convert it to Ino. This means a higher concentration of ADA2 is needed to make up for the amount with steric hindrances. On the other hand, Ado is a rather small molecule, and seeing that the K_m-value is quite similar to that of chicken (0.1-0.3mM), the glycan might not hinder the metabolism of free Ado to a large degree. Rather, considering the theory of ADA2 having a DNA-related function, the glycan can affect DNA binding, as it is a large molecule. In other words, the theory might not be correct for sheep ADA2. However, the assay used Ado, not DNA, as the substrate, so it is impossible to conclude the glycan's effect on DNA binding.

Lastly, sheep ADA2's assumed lower specific activity for free Ado might be due to a lower ADA2 concentration in the samples. But, the glycosylation site's placement ought to have some effect on the protein's function. To better understand this, it will be necessary to have assays with the same concentrations of ADA2 in all samples and test with both free Ado and DNA.

4.2 MS/MS

The following section will discuss the results of the MS analysis. The samples sent to MS analysis were cut-out bands from the SDS-PAGE gel and will be referred to based on the mass of the band, e.g., 55kDa-sample.

The average masses found of the different peaks belonging to glycans, both mono- and oligosaccharides, will not be directly used in the discussion of the spectra. However, the masses will are valuable in that they can be used for a future development of an automated program for determining glycan-containing spectra. This is not the topic of this thesis, but it would have been an invaluable tool to have in the analysis, as the manual analysis of raw data is time-consuming. Not only is it time-consuming to find the correct theoretical masses, but there is also the aspect of finding spectra with both evidence of glycans and the correct peptide.

4.2.1 Chicken ADA2

The findings from the PD analysis were rather odd, as ADA2 was only detected in some of the analyses. All analyses had the same overall settings, only varying which glycans were set as static modifications. In addition, it seems as though ADA2 is not present in the 60kDa sample, only the 55kDa. The theoretical mass of ADA2 is 57kDa, meaning it is not unlikely for the protein to be found in both samples as they are cut out from the SDS-PAGE gel by hand with standard sample bands as reference.

Glycosylated peptides

Although all spectra listed in Section 3.2.2 were promising in being glycosylated ADA2 peptides, there was still some conflicting information in most of the spectra. For example, spectra #25967, thought to be MLNVSQIK with GlcNAc2Man, had conflicting information on the difference between the highest peak and the MH^+, 1499.72827 and 1823.94097Da, respectively. The difference of 324Da does not correspond to any glycan structures or fragments of the peptide. Despite the spectra containing two smaller b-ions, fragment patterns fitting to the glycosylated peptide, and the peak for the peptide mass, the unidentified loss of 324Da disapproves the spectra.

Spectrum #4542, on the other hand, had more compelling evidence of it being HPVAKNLSLK with one GlcNAc. The base peak of the spectra is the peptide mass of 1067.53552Da, with a 203.88039Da difference between this peak and the MH^+. One concerning peak is at *m/z* 534.73535. The peak has high intensity, but is not a known a-, b- or y-ion, half the mass of the base peak or half the mass of any peak. The inability to identify this peak discredits the spectrum. In comparison, spectra #4617, thought to be the same peptide with GlcNAc2Man2, contains a peak at *m/z* 918.48114, which

is half of MH^+ and fragment patterns corresponding to the glycosylated peptide. Similar to #4542, it contains one unidentified peak, *m/z* 739.91150. This peak has a much lower intensity, meaning it is a more unstable ion, but still, it discredits spectrum #4617.

Spectrum #4909 is also HPVAKNLSLK with GlcNAcFuc, which is a *tomi*-glycan. As with the other two spectra of this peptide, the peaks around *m/z* 600-700 are concerning. The peak at *m/z* 728.32050 is half the mass of MH^+, though the peak at *m/z* 638.28796 is unidentified. Other peaks, such as *m/z* 363.08612 and 723.92181, also remain unidentified, though these peaks are likely to be products of smaller fragmentations. Nevertheless, this spectrum is also discredited due to the unidentified peaks, as the previous two.

Spectrum #5420 is the most promising spectrum, being of NILDALMLNTTR with GlcNAc2Man4. The spectrum contains many peaks of lower intensities which were difficult to identify, though the peaks of higher intensity can be identified. Especially the base peaks at *m/z* 789.40436 and 789.90265 being identified as half the mass of the highest peak is important. The neutral loss of a GlcNAc is also important for confirming glycosylation. Some remaining peaks, such as *m/z* 1053.50952 and 972.47614, are the already seen fragmentation pattern, but halved. The loss between the two peaks is 81.02978Da, which is half the mass of a Man, 162.05956Da. The loss of halved masses can be seen with *m/z* 1053.50952, 972.47614, 891.44507, and 810.36786, the losses being 81.02978, 81.03107, and 81.07721, respectively, adding up to 162.05956, 162.06214, and 162.15442, i.e., Man3. This pattern is not seen in the previous spectra and is why those are discredited while #5420 is not.

The only potential spectrum from the 60kDa sample, #21699, is MLNVSQIK with GlcNAc2Man2. The spectrum contains many low-intensity peaks, comparable to noise, making it difficult to find and identify peaks. The clear fragmentation pattern and the difference between MH^+ and peptide mass correspond to GlcNAc2 and are clear proof of it being a glycosylated peptide. However, it is not proof that it is the correct peptide. The presence of y-ions and the peptide mass gives more confidence, but due to the large number of peaks, one cannot be certain about the peptide.

In sum, only one spectrum is determined to be of a glycosylated peptide from ADA2, namely #5420. GlcNAc2Man4 is a trimmed glycan, which can say something about its location in the cell or if it is extracellular. One glycosylated peptide is not sufficient to have a certain conclusion regarding its location, but it provides an

indication of where it has been transported. As it is trimmed, we cannot know the initial glycan structure, whether it was high-mannose, hybrid, or branched. In addition, MS/MS does not give information regarding the bonds between the Man, whether it is a Man(α1-6, α1-3)Man2(α1-6)Man, Man(α1-6, α1-3)Man2(α1-3)Man, Man(α1-6, α1-3)Man2(α1-2)Man, or Man(α1-6)Man(α1-6, α1-3)Man2.

Other findings

The five most abundant proteins in the 55kDa sample are all N-glycosylated proteins. The literature does not state the localization of the most abundant protein, *Vanin 2* (102). However, an analysis of the glycan structures performed by Greiner-Tollersurd indicates that it is a lysosomal protein. This is based on the protein having trimmed glycan structures, which are only achievable by processing by lysosomal glycosidases, i.e. *tomi*-glycans presented in Section 1.2.4. The second most abundant protein, *CN hydrolase domain-containing protein*, does not have its localization stated in the literature either (103). But, having 86% sequence overlap with *Vanin 2*, it might also be lysosomal. The glycan structures of this protein have not yet been investigated, and this is a mere assumption. The third most abundant protein, *N-acetylglucoseamine-6-sulfatase* is a known lysosomal, glycoprotein (104). The remaining proteins mentioned are also all lysosomal glycoproteins that participate in glycan metabolism.

The 60kDa sample had similar content to the 55kDa sample, though there was some variation in the abundance of the proteins. The most abundant protein was *N-acetylglucoseamine-6-sulfatase*, followed by *Vanin 2* and *Carboxylic ester hydrolase*. The latter was also highly present in the 55kDa sample.

Vanin 2 is an interesting protein, as there is limited literature about it. Greiner-Tollersrud's glycan analysis of the protein is also interesting as it shows that it is a lysosomal protein in chicken. In humans, the protein is a GPI-anchored cell membrane protein (105). This demonstrates that a protein can have different localizations in different species. Concerning ADA2, the possibility of it having different localizations in different species might explain the large variations in K_m, specific activity, and overall function. In other words, the evolutionary aspect has to be considered when investigating ADA2, and its function in chickens might be very different from the one in humans. Nonetheless, the findings in other species are important in understanding the function of ADA2 and how it might have evolved.

The other proteins found in the samples are important findings, particularly in two ways, as they indicate: 1) the purification favored glycoproteins, and 2) the purification favored lysosomal proteins. Not all proteins in the samples are glycosylated or lysosomal proteins. Nevertheless, the overweight of these proteins indicates ADA2 falls into one of these groups. It is certainly a glycoprotein; the question is whether it is lysosomal or extracellular. This cannot be answered solely based on these results.

4.2.2 Sheep ADA2

Unlike the PD analyses of chicken, ADA2 was not found in the samples from the fall of 2023, neither 55kDa nor 60kDa, despite varying the glycans for each analysis. This initial result was rather concerning; however, the enzyme assay indicated some ADA2 activity in the samples. Therefore, the reason PD did not "find" any ADA2 is likely one or more of the following: Glycans' diluting effect on proteins, ADA2-peptides without glycans having concentrations below the detection limit, or PD not having statistical confidence in the results. Most likely, it is a combination of all these factors, i.e., a diluting effect of the peptides and an overall low concentration of ADA2.

The spring purification indicates that ADA2 is indeed present in sheep spleen. Hence, the presence of ADA2 is confirmed and a rather high presence as well. It is concerning that ADA2 does not show in the fall purification, as the same methods were used. However, it is likely due to the issues of the purification process and using less starting material. The exact amount of sheep spleen used in the spring purification was not disclosed to the candidate, but based on the high abundance in that purification it was concluded that it would be sufficient with less material.

The initial observation during manual analysis was that the sheep samples had significantly fewer proteins. In addition, there is significantly more contamination compared to the chicken samples, including contamination from *Bos Taurus* and *Sus Scrofa*, respectively, cattle and pig. These contaminants are likely due to the equipment being used to purify ADA2 from these two species. There does not seem to be any contamination from chicken, but this is reasonable considering chicken ADA2's pI and the pH used in the CM column, respectively, 5.9 and 7.2, causing chicken ADA2 to be in the RT fraction. The pI of other chicken glycoproteins are not known, but it is reasonable to assume most proteins were eluted in the RT fraction as having a pI > 7.2 is not so common. Additionally, as PD only uses a FASTA-database with likely contaminants and a database for the species being analyzed, there is the possibility of wrong identifi-

cation.

Glycosylated peptides

The only possible glycosylated peptide found in the sample from the fall purification is spectrum #2928, possibly of NESLIK. The spectrum contains many peaks, indicating considerable fragmentation, making it difficult to identify the single peaks and neutral losses. The findings presented in Section 3.2 are insufficient to be confident that the spectrum is of NESLIK with GlcNAc-1Da.

Spectrum #46086 from the spring sample is likely to be of the peptide ASLYPVYEL-NGTVYSQEWLVR with GlcNAc2FucMan3. The peak at *m/z* 1175.59204 does discredit the spectrum, as no fragment ions have this mass, and it is not part of a halved fragmentation pattern. In addition, a loss of GlcNAc-1Da is not a common fragmentation pattern. It is more common to find peaks of the peptide with GlcNAc-1Da. Despite some compelling evidence, the spectrum is likely not of this peptide.

The second spectrum presented, #6040, thought to be of TKNESLIK with GlcNAc-H_2O, is more compelling. The spectrum has many peaks corresponding to ions of the peptide, and the base peak at *m/z* 560.31390 is likely the MH^+ halved. Though 560.31390Da is higher than the MH^+ when doubled, there is a peak at *m/z* 559.28790, which gives MH^{2+} when doubled. This is feasible as the precursor charge is +2. Therefore, this spectrum is likely of TKNESLIK with GlcNAc-H_2O.

The final spectrum found in this sample was #17262, containing LLDAVILNSTR with GlcNAc2FucMan-H_2O. The base peak at *m/z* 1030.57068 is half of MH_+ plus 1Da. What discredits this spectrum is that the peaks at *m/z* 799.40125 and 915.44403 are unidentified. The peaks do not correlate with a fragmentation pattern of glycans or the peptide, halved or not. As there is not a large number of peaks in the spectrum, the peaks cannot be deemed as noise or results of smaller losses from other peaks. Hence, the spectrum is likely not correct.

Other findings

As initially mentioned in this subsection, the samples have a high degree of contamination. Compared to chicken, the high degree of contamination and low amounts of identified proteins indicated that the purification was at fault. Nevertheless, the pres-

ence of *N-acetylglucoseamine-6-sulfatase*, *N- sulfoglucosamine sulfohydrolase*, and *N-acetylgalactoseamine-6-sulfatase* is an indication that glycoproteins were isolated. All three proteins are glycosylated, lysosomal proteins with a main function of hydrolyzing their respective substrates, i.e., metabolic enzymes (106–108). Despite not having identified ADA2 in the fall samples, identifying these lysosomal glycoproteins and the positive results from the enzyme assays indicate that ADA2 is glycosylated and perhaps lysosomal. As with the results from chicken, this analysis alone is insufficient to conclude the localization of ADA2.

The most abundant proteins in the sample from the spring, *Complement C4* and *Alpha-2-macroglobulin*, are secretory glycoproteins (109; 110). The other proteins mentioned, such as α-mannosidase and N-acetylglucosamine-6-sulfatase, also identified in the fall samples, indicated the high concentration of these proteins in sheep spleen. Both proteins are lysosomal glycoproteins contributing to the metabolism of glycans. Especially α-mannosidase is an important protein, as discussed in Section 1.2.3, in that it is crucial to regulate the Man-concentration in the lysosomes.

The results from these samples confirm that ADA2 is a glycoprotein, but as there is such a small amount of ADA2, there is not enough to determine a "trend" in its glycan profile, i.e., *tomi* or not. As the samples contain both secreted and lysosomal glycoproteins, the only conclusion that can be drawn with certainty is that ADA2 is glycosylated - which is already known. The one spectrum found by manual analysis, likely to be KNESLIK with GlcNAc-H_2O, does not give any indications either, as a single GlcNAc does not offer any clear indication regarding its localization.

Analysis of both chicken and sheep have similar results, with simply having low concentrations of ADA2. Typically, enzymes present at lower concentrations have an enzymatic function, while enzymes of higher concentrations have a DNA-binding effect. This is reasonable as one enzyme can perform its function efficiently on many substrates, whereas DNA-binding requires the enzyme to be "occupied" until a DNase, e.g., DNaseI or II, depending on the localization, metabolizes it. Hence, low concentrations of ADA2, considering the diluting effect of glycans, indicate that ADA2 has an enzymatic function in these two species.

4.3 The function and localization of ADA2

As explained in Section 1.3.5, the existing theories on ADA2 are based on it being secreted from macrophages/monocytes and has an enzymatic function converting extracellular Ado to Ino. As mentioned initially, this thesis will explore the possibility of ADA2 being a lysosomal protein with a DNase-like function, either combined with Ado metabolism or as the sole function. Based on the large difference between ADA2's K_m-value and the physiological Ado concentration, this hypothesis came to life. This is also one of the factors discrediting most theories, as Hershfield, Tarrant, and Kelly (2021) explained in their review article (56). The K_m-value, being 2-3mM, is significantly higher than Ado's physiological, extracellular concentration of >1μM (2; 51; 54). This indicates that ADA2 has a very low conversion rate under physiological conditions, and a full or partial protein dysfunction should not have as critical consequences as seen in DADA2 if this is ADA2's only purpose. Therefore, it is proposed that the metabolism of Ado is not ADA2's main function.

Comparing the protein's K_m-values in different species suggests a shift in the function through evolution. Multiple studies have shown ADA2 has a rather high Ado-conversion in chickens and pigs; chickens' high conversion was also shown in this thesis (54; 65). Sheep ADA2 is not discussed to the same degree in the literature, but as seen in the assays performed here, it does have a rather high conversion rate, comparable to chicken, as long as the protein concentration is high enough. In other words, the enzyme assays showed the two species have a similar conversion rate, though sheep ADA2 is more concentration-dependent and has a lower specific activity. In addition, as mentioned in Section 4.2, a low protein concentration generally implies an enzymatic function, not a binding function. The low concentration of ADA2 found in the chicken samples might imply an enzymatic function for free Ado. On the other hand, sheep require more investigation, as two purifications were performed with the same method but resulted in very different ADA2 concentrations not correlating to the difference in starting material. Additionally, the glycosylation site near sheep ADA2's active site should be investigated further, which will be discussed later in this section.

Another significant aspect is that ADA2 is completely missing from rodents. There has been an evolutionary elimination of ADA2 from rodents, and ADA1 remains as the single enzyme responsible for Ado metabolism. Factors such as the "inconvenient" glycosylation site in sheep ADA2 and the high K_m of human ADA2 can also indicate this is occurring in sheep and humans, i.e., an evolutionary elimination. Or, rather than

a full elimination, a change from Ado conversion to another function.

A second and third property that varies between species is the theoretical pI and pH optimum. This characteristic was presented in Section 1.3 where the variations between chicken, sheep, and humans were displayed. The pH properties of a protein greatly affect its function and can indicate where it is localized. Chicken ADA2's pH optimum was determined to be 5.99 by Iwaki-Egawa et al. (2004), and its theoretical pI is 6.5 (54). This means that if the protein is in an environment with pH 5.99, it has a slight negative charge. In the extracellular space, with an approximate pH of 7.4, it will have a stronger negative charge (60). The pH optimum is not determined for sheep ADA2, but its theoretical pI is 8.6. Assuming an extracellular function, sheep ADA2 will have a positive charge. Human ADA2, with an optimum pH of 6.5 and pI of 7.75, will have a negative charge at its optimum and essentially be neutral at physiological, extracellular pH. In other words, all three species have ADA2 with different charges at physiological, extracellular pH. Regarding the enzymatic conversion of Ado, seen in Figure 1.9, the conversion itself is not dependent on the protein's charge, and the varying charge of the three species should not matter to a large degree. It is the Zn^{2+}-ion in the active site that is the electrophile in the reaction, and its charge is independent of the protein. This might indicate that the protein's overall charge is not significant for the function, and considering the species are different in other aspects as well, the charge difference might not be critical. Nonetheless, it is noteworthy as charge and pH properties are generally important for a protein's function.

Local, extracellular acidosis is induced at a site of inflammation, reducing the pH to 6.0-7.0 (111). As the Ado concentration increases at inflammation sites, it is reasonable to assume the ADA2 concentration also increases to regulate the Ado levels (51). This has also been found to occur in conditions such as HIV and Tuberculosis, and ADA2 has also been used as a biomarker in some cases (56). However, it has not been investigated whether the elevated ADA2 levels are due to elevated Ado levels or if there is another cause. With elevated Ado levels and inflammation-induced acidosis, both chicken and human ADA2 will have a more effective Ado metabolism since the pH is closer to their pH optimum of 5.99 and 6.5, respectively. It appears that ADA2's effectiveness increases in parallel with the increased Ado levels. In terms of charge, chicken ADA2 is essentially neutral, while human and sheep ADA2 has a positive charge. The conversion of Ado to Ino, as mentioned, seems to be unrelated to the charge of the protein, as it is solely based on the affinity of Ado to the binding site and the Zn^{2+}-ion. Still, it is odd that human and sheep ADA2 have such a strong positive charge under

these conditions that it ought to have some purpose. For example, a strong positive charge is optimal for DNA binding.

It is seen that DNA is released from cells to the extracellular space during inflammation as *cell free DNA* (cfDNA) (112). This is primarily associated with systemic inflammation, as seen in DADA2, but one can also assume it also occurs under local inflammation. Free DNA and cfDNA are highly pro-inflammatory molecules and are likely to cause further inflammatory responses, both systemically and locally. At these local sites, DNases are responsible for metabolizing the DNA and reducing the activation of inflammatory pathways. One proposition is that ADA2 contributes here by binding and metabolizing free and cfDNA via Ado, i.e., functioning as an extracellular endonuclease. This would explain the purpose of the strong positive charge, as this is optimal for DNA binding. This also aligns with the increased ADA2 levels seen in inflammatory conditions like HIV and Tuberculosis. To investigate this proposition further, it is necessary to know if and how ADA2 binds to DNA and if there is any selectivity for different DNA structures.

The main proposition and hypothesis for human ADA2 are that it is lysosomal with a DNA function, i.e., an endonuclease, not extracellular with an Ado function. The pH properties of ADA2, especially how they translate to ADA2's charge, are a strong argument for a lysosomal nature. The environment of an inflammation-induced acidosis is comparable to lysosomal conditions, which is highly acidic with an approximate pH of 5 (26). ADA2 would have the same DNA binding and degrading effect in the lysosomes as it would extracellularly during localized inflammation and acidosis. As a lysosomal endonuclease, a deficiency will lead to much higher stress on the cell types that are particularly responsible for DNA degradation, i.e., locations with a high degree of "DNA waste." Macrophages, particularly the ones in the bone marrow, are related to the maturation of erythrocytes and have a higher load of DNA waste than, e.g., macrophages in the liver. As erythrocytes are nucleus-less cells, the macrophages are responsible for degrading the nucleases and all their contents, resulting in vast amounts of DNA (113). Hence, macrophages in the bone marrow will be under enormous stress with an ADA2 deficiency, regardless of whether ADA2 expression is promoted specifically in monocytes/macrophage lysosomes or generally in all cell types. ADA2 is still present in macrophages, even without a monocyte/macrophage-specific promoter. Nevertheless, the stress on the bone marrow macrophages is likely to cause symptoms of the CNS, which has been seen in DADA2 patients. In the four screenings of patients done by Nanthapisal et al., Navon et al., Zhou et al., and Batu et al., 42.6% presented symptoms

of CNS (72–75).

Staying on the topic of erythrocytes, anemia is also a symptom found in DADA2 patients (114–116). Anemia can be caused by multiple factors in the DADA2 phenotype. However, if ADA2 is an endonuclease, a feasible cause is the increased stress of bone marrow macrophages due to a lack of ADA2 activity. The increased stress would then lead to a compensating down-regulation of the production of erythrocytes. None of the three screenings has a conclusion regarding the cause. Yet, seeing that anemia is a common symptom in patients with DNaseII deficiency, a known lysosomal endonuclease, there is potentially a similarity between the nature of DNaseII and ADA2 (33). Parallels to DNaseII and deficiency of this protein will be further discussed.

ADA2 being a lysosomal protein and not specific to and secreted by macrophages, one would expect the concentration of ADA2 to be rather consistent in all types of tissues and organs. Considering today's theories of it being secreted from macrophages, the amount of purified ADA2 would vary significantly depending on the purified organ. For example, the spleen is a more macrophage-rich organ than the kidneys, and one would expect a higher yield from the spleen. A "general" lysosomal protein would be expected to have a more consistent concentration in all organs and tissues. There is still some variation as the lysosomes have different enzyme compositions depending on their cell type and where the cell is localized, but not to the extent seen for cell type-specific proteins (117). For example, an endonuclease is more critical in the bone marrow than the skin, and there is a higher ratio of these enzymes in bone marrow lysosomes. However, despite the variation in enzyme composition, the lysosome is present in all cell types, and one should find a decent amount of ADA2 regardless of the purified organ. Unfortunately, this concept is also lacking in research, as there have only been purifications from one organ from different species: Porcine brain, chicken liver, and sheep spleen. In the future, purifying ADA2 from different organs in these species will give valuable insights, e.g., chicken spleen and sheep liver. This will allow for a more comprehensive comparison of the different species. Comparing macrophage-rich and macrophage-poor organs, e.g., liver and spleen, from the same species will also present whether ADA2 expression is promoted in macrophages and, hence, if its function is related to macrophages or is generalized.

The leading argument up until now has been that ADA2 is a lysosomal protein. However, looking at the pI and optimum pH of human ADA2, which are 7.75 and 6.5, respectively, these values do not necessarily "fit" into the general characteristics of a

lysosomal protein. According to a study comparing the median pI's of proteins in the various organelles, the median for lysosomal proteins is 6.6 (118). There are variations, with pIs ranging from pH 4 to 10, but the most abundant proteins have pIs closer to 6. Interestingly, DNaseII, the best comparison when assuming ADA2 is a lysosomal endonuclease, has a pI of 8.3 (Expasy Compute pI/Mw). In other words, ADA2 has a high pI compared to the median but is perhaps too low to have optimal DNase function. The lysosomal pH is around 4.7, meaning the higher the pI, the stronger the positive charge and, consequently, the stronger the binding to negatively charged DNA. Chicken ADA2 has a very weak positive charge, further solidifying that chicken ADA2 has an Ado-related function, not DNA-related. However, sheep ADA2 has a pI higher than human ADA2, resulting in a strong positive charge in the lysosomes that are well-suited for DNA binding. This solidifies sheep ADA2 having a function related to DNA, perhaps even more so than in humans. Yet, sheep ADA2 has the glycosylation site in the opening of the active site, hindering possible DNA binding and discrediting the idea of a DNA-related function.

Returning to the topic of evolutionary changes. It is a complex topic as it has occurred over thousands of years, and it can be difficult to pinpoint exactly when, where, and for what purpose different species' protein functions diverge. Regardless, it is clear that ADA2's role has not yet been "eliminated" from humans, as it has with rodents, since a deficiency leads to severe autoinflammatory disease. While ADA1's function seems consistent across species, the large variations in ADA2 imply evolutionary changes. Investigating different mutations will give interesting insights, especially human and sheep ADA2 mutations. Human ADA2, where the glycosylation seat near the active site is added, and sheep ADA2 without the glycosylation site can indicate the exact effect of the glycosylation site. Investigating mutated versions of human ADA2 will also give information regarding which amino acid properties enhance specific protein properties, e.g., charge, pI, and affinity. Other possible mutations can be varying the hydrophobic properties of the binding site, the sizes of the more exposed amino acids, or adding amino acids with different charges. In addition, an even more in-depth analysis of the amino acid sequences from different species will be interesting in determining more trends regarding charge and hydrophobic/hydrophilic properties in the different domains, especially the catalytic domain, including the active seat and binding site. The trend of lower species like flesh flies, having more hydrophobic active sites and very high Ado-activity is already determined (49). Uncovering other trends will help find the function of ADA2 in more species than the ones mentioned here.

Already discussed in this section and Section 4.1, is the evolutionary addition, and possibly removal, of an ADA2 promoter specific for monocytes/macrophages. Also mentioned was how this promoter seems absent in sheep ADA2 due to the low amounts in this purification, comparable to that in chicken liver. This means sheep ADA2 is not specific for macrophages and instead has a more general function. Along with the glycan possibly hindering DNA binding, sheep ADA2 most likely stands for a general "housekeeping" of free Ado. However, the glycan might also affect the binding of free Ado, despite it being a small molecule and ADA2 having a rather low K_m-value. If the glycan affects Ado binding, it raises the question of whether sheep ADA2 has a completely different function. This argument cannot be made for chicken as there have not been any purifications of macrophage-rich organs, but this should be done in the future. It is valuable to know if multiple species have lost the macrophage promoter. The promoter remains in pigs, indicated by large amounts of ADA2 being purified from macrophage-rich organs. This also seems to be the case for human ADA2, but this should be further investigated for a clear conclusion.

The Ado function seems completely missing in humans due to the very high K_m-value (2.5mM). Also, human ADA2 does not have the "inconvenient" glycosylation site. Along with the positive charge at acidic conditions, this strongly indicates a DNA-related function. As mentioned, chicken and sheep ADA2 have indications of still having a function for free Ado despite the aspects already discussed. And, the indications of a DNA-related function in humans might be an indication itself of chicken and sheep ADA2 still metabolizing free Ado.

Unfortunately, the MS analysis did not give as many results as desired regarding the glycan structures. Yet, the lack of results is still a result, as the concentration of proteins can indicate its function, as previously explained. However, concerning glycan profiling, a high abundance of data is necessary to discover the structure trends. Finding a few structures is insufficient to confidently say something about a protein's glycan profile. Profiling is a valuable tool when investigating the localization of a protein, i.e., secreted vs. lysosomal. *Tomi*-glycans are one way of determining if a protein is lysosomal, but phosphorylation also indicates a lysosomal nature. The phosphorylation process is part of the transportation pathway utilized by lysosomal glycoproteins, which depends on the mannose-6-phosphate (M6P) receptor pathway. The pathway was briefly mentioned in Section 1.2.4.

Only the proteins meant to be translocated to the lysosomes are phosphorylated

during the N-glycans biosynthesis. This means secreted glycoproteins do not have phosphorylated glycans. Hence, determining the presence of phosphorylated glycans on ADA2 will signify it is a lysosomal protein. An abundant lysosomal glycoprotein, N-acetylglucosamine-6-sulphatase, has been determined to have phosphorylated N-glycans and *tomi*-glycans, both in this analysis and previous ones. By comparing the glycan profile of ADA2, which will be determined in future analyses, to those of known lysosomal proteins, one can see if ADA2 is lysosomal or not. However, it is important to note that lysosomal phosphatases remove the phosphate groups on the glycans, and lysosomal proteins tend to have a combination of phosphorylated and unphosphorylated glycans. Therefore, the glycan profile of a lysosomal glycoprotein will consist of phosphorylated and unphosphorylated glycans, *tomi*-glycans, and also "normally" trimmed glycans. The presence of some structures that can only exist after processing by lysosomal glycosidases indicates that a protein is lysosomal, which means not every glycan has to be a lysosome-specific structure.

In Section 1.2.3, it was explained how LSDs mainly have autoinflammatory phenotypes, which DADA2 aligns with. In cases of DNaseII deficiency, the accumulation of DNA in the lysosomes causes the activation of the cytosolic STING pathway, leading to type I IFN production and inflammation. Assuming ADA2 has a similar function of binding and degrading DNA and retaining it in the lysosomes, a deficiency can lead to an accumulation of unaccounted-for DNA. To investigate this effect, one would traditionally make a knock-out model of mice or rats, like with DNaseII (32; 33). However, as ADA2 is absent in rodents, a knock-out model will not be representative, and other species must be used. A species where ADA2 has a seemingly similar function should be chosen. Hence, chickens and sheep are unsuitable as their ADA2 function is likely different from that in humans. ADA2 appears to have a rather clear function for Ado in these two species, which it does not have in humans. Sheep ADA2 seems to be more comparable to human ADA2. The K_m-values differ, but the pI's, hence the charge, are similar, and the glycosylation site near the active site, reducing its availability, can be thought to have a similar effect on the function for Ado as the high K_m in humans. On the other hand, sheep ADA2 will most likely not be able to bind DNA due to the glycan. When the purpose is to investigate a DNase-like function, a different species with clearer indications of this should be chosen. Still, it would be interesting to have knock-out models of sheep to investigate its Ado function.

Assuming ADA2 has a lysosomal DNase function and DNA is the accumulating macromolecule, there is still the question of how the DNA causes inflammation. One

explanation is that DNA leaks out to the cytosol and activates the STING pathway unless the DNA is degraded by a cytosolic DNase, like TREX1. This theory is based on the spontaneous pore formation in membranes, presented in Section 1.2.1 (14). The idea that membranes are "open" without a specific purpose is radical, especially the lysosomal membrane, as the lysosome is the "garbage bin" of the cell. Spontaneous pores and leakage means cell waste, such as microbes, phagocytosed material, etc., will accumulate in the cytosol. Another hypothesis is that DNA accumulation induces pore formation or membrane damage, causing leakage. Both theories are possible; perhaps it is a combination where accumulated DNA increases the frequency of spontaneous pore formation and does not necessarily damage the membrane. Either way, TREX1 is the endonuclease regulating cytosolic DNA, and it is the last line of defense for avoiding activation of STING and increasing type I IFN production.

Inflammation due to activation of STING can only be explained by cytosolic DNA binding to cGAS. As mentioned, it is known that DNA in the cytosol, with various origins, is degraded by TREX1 and other endo- and exonucleases, avoiding STING-activation. However, if STING is the basis for the autoinflammation in DADA2, the only explanation is DNA originating from the lysosomes. There has to be some form of translocation from the lysosomes due to, for example, openings in the membrane. Regardless of how the translocation, or leakage, occurs, DNA accumulation will result in the leaked material containing more DNA than leaked material without accumulation. This means there will be more DNA for cytosolic DNases, i.e., TREX1, to degrade. The accumulated amount might oversaturate TREX1, resulting in over-activation of the STING pathway and systemic inflammation.

The theory regarding cholesterol crystals leading to lysosomal membrane damage and breakage brings up the topic of specific macromolecules increasing the frequency of pore formation. Compared to the first-mentioned theory of spontaneous pores, the second theory relates the pores to the accumulated macromolecule. This topic needs substantial research, especially regarding DNA having a similar effect on pore formation and membrane damage when accumulated, i.e., accumulated DNA inducing pores and, hence, leakage. Still, these concepts are new concerning DNA, with many variables and unknowns. Nevertheless, if it is the case that STING-activation is the cause of autoinflammation, either of these theories or a combination, are reasonable explanations.

The accumulating macromolecule greatly impacts the pathomechanism and severity

of a disease. Whether DNA, Ado, or both accumulate affects what mechanism initiates inflammation and how it unfolds. A good example is the severity of α-mannosidsis in cattle versus humans. In humans, a trisaccharide accumulates with a GlcNAc-core, while in cattle, it is a tetrasaccharide, as the oligosaccharide has a GlcNAc2-core. The disease is much more severe in cattle than humans, solely due to the additional GlcNAc in the accumulating oligosaccharide (119). Although this does not correlate exactly with DADA2, it does visualize the importance of knowing the accumulating substrate caused by a deficient enzyme.

On the other hand, diseases that DADA2 can be compared to are deficiencies of DNaseII and ADA1. The phenotype of a DNaseII deficiency is autoinflammation, including recurring fevers, rheumatism, and vasculitis. As mentioned, this results from accumulating DNA, which causes inflammation reactions driven by STING activation. ADA-SCID, the deficiency of ADA1, results in the intracellular accumulation of Ado. In addition, dATP accumulates and inhibits DNA synthesis (77). However, in ADA-SCID, the acquired immune system is downregulated by preventing DNA synthesis in B- and T-cells and is therefore not autoinflammatory; this differentiation is important. Regardless, there would be some similarities between the deficiencies of ADA1 and ADA2 if they have the same accumulating substrate, i.e., Ado. At first glance, DNaseII deficiency and DADA2 have more common ground than DADA2 and ADA-SCID, as they are both autoinflammatory diseases causing increased expression of type I IFN. In addition, elevated Ado levels have not been found in DADA2 patients, though this has not been researched in-depth yet. Despite both deficiencies of ADA2 and DNaseII being autoinflammatory diseases, it is not enough to state that ADA2 and DNaseII have the same substrate and function, as many other autoinflammatory diseases exist. But, it is one aspect that has to be considered, along with the previously discussed topics.

It is clear in this discussion that there are many layers to understanding ADA2's function, and many more are not mentioned here. The most fundamental aspects, such as the K_m-value, charge, and evolutionary differences, are discussed and show the complexity of understanding the purpose of even the most rudimentary characteristics of ADA2. In a way, most findings contradict, such as a deficiency not resulting in elevated Ado. In addition, there is compelling evidence that human ADA2 does not hydrolyze Ado *in vivo*. Also, the significant differences from pig and chicken, which likely have a function of hydrolyzing Ado, strengthen the initial hypothesis of the thesis. At least, it discredits the existing theories. In addition, many similarities to another lysosomal endonuclease, DNaseII, strengthen the aspect of ADA2 having a lysosomal function,

especially regarding the disease phenotype and pH properties. However, to confidently determine this, more analyses of the protein and patients must be conducted. Additionally, having a full glycan profile of ADA2 is valuable as it can be used to find the exact location of the protein. Understanding ADA2 means discovering both the localization and the function, not one or the other. But finding one will make it easier to discover the other. Most importantly, the protein's function must be known to understand the disease, but having an in-depth understanding of the symptoms is valuable in finding the function. In the future, extensive glycan analyses of ADA2 from different species and tissues, mutagenesis of human ADA2, and better screenings of patients will be necessary.

Lastly, all research on ADA2 aims to find the optimal treatment for DADA2 patients. Enzyme replacement treatment (ERT) is a logical choice, as it, in simple terms, gives the patient the lacking enzyme. ERT is also favorable as the glycans on the ER will guide the protein to the macrophages' lysosomes via the M6P pathway. The downside to ERT is that it has to be given intravenously every 2-3 weeks, as it does not allow the patient to produce their own proteins. Additionally, the effects of an ADA2 "overdose" are also not known. Any possible negative effects of too high ADA2 concentration have to be determined, in addition to ADA2's normal function. Therefore, bone marrow transplantation (BMT) or stem cell therapy (SCT) might be better. With these types of treatments, the patient is given the ability to produce their own functional ADA2. However, the development of BMT and SCT is complex, and the protein's function must be well understood. It is also important to know the consequences of protein levels that are too high and how to avoid an overproduction of the protein. This applies to all three treatment strategies.

Today, DADA2 is treated with anti-TNFα, which tackles the resulting autoinflammation. The treatment is effective for most patients, but it does not treat the foundation of the disease, only the symptoms. ADA-SCID is generally treated with ERT or hematopoietic SCT (HSCT), where this is possible. The latter approach, SCT, is likely the most suitable for DADA2, as mentioned previously. There are already a few cases of successful HSCT in DADA2 patients having hematological symptoms (22). The ultimate goal is to find a treatment for every DADA2 patient, regardless of their symptoms. Whether this ends up being ERT, SCT, or BMT is not what is important; giving the patients the opportunity for a healthy life is.

Chapter 5

Conclusions and Future Work

In this thesis, ADA2 was purified from chicken liver and sheep spleen, followed by enzyme assays, an extensive MS/MS analysis, and an in-depth discussion of the protein. The purification was done according to previous purifications, but MS analysis showed that these procedures need further development to increase the yield and better isolate ADA2. Nevertheless, the results from the assays and MS analysis provided some insights when discussing ADA2's function. Furthermore, the results were combined with the existing research and theories on the protein. It was argued that the very foundation of the protein, i.e., its function and localization, presented in the literature might be inaccurate. Instead of human ADA2 being secreted from monocytes/macrophages and hydrolyzing free Ado, it was suggested that ADA2 is lysosomal with a DNA-related function and specific for monocytes/macrophages. To ground this statement, parallels were drawn to other lysosomal proteins and endonucleases in terms of pH properties, pathomechanism, and glycan structures. Particularly, similarities to the lysosomal DNaseII give strong indications of both a lysosomal nature and a DNA-related function.

A comparison of the properties of sheep, chicken, and human ADA2 uncovered some significant differences, especially in terms of a direct Ado function. While it seems as though sheep and chicken ADA2 still have some function related to Ado, there is compelling evidence that this is not the case for human ADA2. Differences in localization were also found, where sheep ADA2 does not seem to be specifically promoted in monocyte/macrophages, while human ADA2 likely is.

In addition to the discussion on ADA2, MS/MS data directly related to masses of glycan structures and fragments was obtained. Hopefully, this data will improve glycan analysis and profiling in the future. This thesis has clarified the importance of glycobiology when investigating proteins, and it is a field that should have sufficient analytical

tools.

Unfortunately, the findings do not allow for a certain conclusion on the objectives of the thesis, which was to find the physiological function of ADA2. However, the topics mentioned in the discussion will be further investigated, and Greiner-Tollersrud will carry on the research. This novel theory on ADA2 is not widely recognized in the literature. Yet, as the existing theories all fall short, this approach should be more seriously considered as it provides some possible explanations for previously unanswered questions.

Finally, the primary objective for all researchers investigating ADA2 is to develop a treatment for DADA2 patients and improve their quality of life. Although the exact function of ADA2 has not been determined, the research conducted over the past year has provided valuable insights that have contributed to this goal.

Chapter 6

Appendix Glycan and peptide masses

6.1 Glycan masses

Table 6.1: Monoisotopic masses of all N-glycans

Glycan	Monoisotopic mass [Da]
Man1GlcNAc2 (+568)	568.2116
Man1Fuc1GlcNAc2 (+714)	714.2695
Man2GlcNAc2 (+730)	730.2644
Man2Fuc1GlcNAc2 (+876)	876.3223
Man3GlcNAc2 (+892)	892.3172
GlcNac1Man1Fuc1GlcNAc2 (+917)	917.3489
Man3Fuc1GlcNAc2 (+1038)	1038.3751
Man4GlcNAc2 (+1054)	1054.37
GlcNAc1Man2Fuc1GlcNAc2 (+1079)	1079.4017
GlcNAc1Man3GlcNAc2 (+1095)	1095.3966
P1Man4GlcNAc2 (+1134)	1134.3487
Man4Fuc1GlcNAc2 (+1200)	1200.4279
Man5GlcNAc2 (+1216)	1216.4228
GlcNAc1Man3Fuc1GlcNAc2 (+1241)	1241.4545
Gal1GlcNAc1Man3GlcNAc2 (+1257)	1257.4494
GlcNac2Man2Fuc1GlcNAc2 (+1282)	1282.4811
P1Man5GlcNAc2 (+1296)	1296.4015
GlcNAc2Man3GlcNAc2 (+1298)	1298.476
Man5Fuc1GlcNAc2 (+1362)	1362.4807
Man6GlcNAc2 (+1378)	1378.4756

Table 6.1 continued from previous page

Gal1GlcNAc1Man3Fuc1GlcNAc2 (+1403)	1403.5073
Gal1Fuc1GlcNAc2Man3GlcNAc2 (+1403)	1403.5073
GlcNAc1Man5GlcNAc2 (+1419)	1419.5022
GlcNAc2Man3Fuc1GlcNAc2 (+1444)	1444.5339
P1Man6GlcNAc2 (+1458)	1458.4543
Gal1GlcNAc2Man3GlcNAc2 (+1460)	1460.5288
GlcNac3Man3GlcNAc2 (+1501)	1501.5554
Man6Fuc1GlcNAc2 (+1524)	1524.5335
AcSia1Gal1GlcNAc1Man2Fuc1GlcNAc2 (+1532)	1532.5499
Man7GlcNAc2 (+1540)	1540.5284
GcSia1Gal1GlcNAc1Man2Fuc1GlcNAc2 (+1548)	1548.5448
AcSia1Gal1GlcNAc1Man3GlcNAc2 (+1548)	1548.5448
Gal1Fuc1GlcNAc1Man3Fuc1GlcNAc2 (+1549)	1549.5652
GlcNAc1Man5Fuc1GlcNAc2 (+1565)	1565.5601
Gal1GlcNAc1Man5GlcNAc2 (+1581)	1581.555
GlcNAc1Man6GlcNAc2 (+1581)	1581.555
Gal1GlcNAc2Man3Fuc1GlcNAc2 (+1606)	1606.5867
Gal1Fuc1GlcNAc2Man3GlcNAc2 (+1606)	1606.5867
P1Man7GlcNAc2 (+1620)	1620.5071
GlcNAc2Man5GlcNAc2 (+1622)	1622.5816
Gal2GlcNAc2Man3GlcNAc2 (+1622)	1622.5816
GlcNAc3Man3Fuc1GlcNAc2 (+1647)	1647.6133
GlcNAc1P1Man6GlcNAc2 (+1661)	1661.5337
Gal1GlcNac3Man3GlcNAc2 (+1663)	1663.6082
AcSia1Gal1GlcNAc1Man3Fuc1GlcNAc2 (+1694)	1694.6027
AcSia1Gal1GlcNAc1Man3Fuc1GlcNAc2 (+1694)	1694.6027
P2Man7GlcNAc2 (+1700)	1700.4858
Man8GlcNAc2 (+1702)	1702.5812
GlcNAc4Man3GlcNAc2 (+1704)	1704.6348
GcSia1Gal1GlcNAc1Man3Fuc1GlcNAc2 (+1710)	1710.5976
Gal1Fuc1GlcNAc1Man5GlcNAc2 (+1727)	1727.6129
Gal1GlcNAc1Man5Fuc1GlcNAc2 (+1727)	1727.6129
GlcNAc1Man6Fuc1GlcNAc2 (+1727)	1727.6129
Gal1GlcNAc1Man6GlcNAc2 (+1743)	1743.6078
AcSia1Gal1GlcNAc2Man3GlcNAc2 (+1751)	1751.6242
Gal1Fuc1GlcNAc2Man3Fuc1GlcNAc2 (+1752)	1752.6446

Table 6.1 continued from previous page

Gal2GlcNAc2Man3Fuc1GlcNAc2 (+1768)	1768.6395
Gal2Fuc1GlcNAc2Man3GlcNAc2 (+1768)	1768.6395
GlcNAc2Man5Fuc1GlcNAc2 (+1768)	1768.6395
P1Man8GlcNAc2 (+1782)	1782.5599
Gal1GlcNAc2Man5GlcNAc2 (+1784)	1784.6344
GlcNAc2Man6GlcNAc2 (+1784)	1784.6344
GlcNAc1P1Man6Fuc1GlcNAc2 (+1807)	1807.5916
Gal1GlcNAc3Man3Fuc1GlcNAc2 (+1809)	1809.6661
Gal1Fuc1GlcNAc3Man3GlcNAc2 (+1809)	1809.6661
Gal1GlcNAc1P1Man6GlcNAc2 (+1823)	1823.5865
Gal2GlcNAc3Man3GlcNAc2 (+1825)	1825.661
GlcNAc4Man3Fuc1GlcNAc2 (+1850)	1850.6927
P2Man8GlcNAc2 (+1862)	1862.5386
GlcNAc2P1Man6GlcNAc2 (+1864)	1864.6131
Man9GlcNAc2 (+1864)	1864.634
Gal1GlcNAc4Man3GlcNAc2 (+1866)	1866.6876
AcSiaGal1GlcNAc1Man5GlcNAc2 (+1872)	1872.6504
Gal1Fuc1GlcNAc1Man5Fuc1GlcNAc2 (+1873)	1873.6708
Gal1Fuc1GlcNAc1Man6GlcNAc2 (+1889)	1889.6657
Gal1GlcNAc1Man6Fuc1GlcNAc2 (+1889)	1889.6657
AcSia1Gal1GlcNAc2Man3Fuc1GlcNAc2 (+1897)	1897.6821
AcSia1Gal21GlcNAc2Man3GlcNAc2 (+1913)	1913.677
Gal2Fuc2GlcNAc2Man3GlcNAc2 (+1914)	1914.6974
Gal2Fuc1GlcNAc2Man3Fuc1GlcNAc2 (+1914)	1914.6974
Gal1Fuc1GlcNAc2Man5GlcNAc2 (+1930)	1930.6923
Gal1GlcNAc2Man5Fuc1GlcNAc2 (+1930)	1930.6923
GlcNAc2Man6Fuc1GlcNAc2 (+1930)	1930.6923
Gal1GlcNAc2Man6GlcNAc2 (+1946)	1946.6872
AcSia1Gal1GlcNAc3Man3GlcNAc2 (+1954)	1954.7036
Gal1Fuc1GlcNAc3Man3Fuc1GlcNAc2 (+1955)	1955.724
Gal1Fuc1GlcNAc1P1Man6GlcNAc2 (+1969)	1969.6444
Gal1GlcNAc1P1Man6Fuc1GlcNAc2 (+1969)	1969.6444
Gal2GlcNac3Man3Fuc1GlcNAc2 (+1971)	1971.7189
Gal2Fuc1GlcNAc3Man3GlcNAc2 (+1971)	1971.7189
Gal3GlcNAc3Man3GlcNAc2 (+1987)	1987.7138
GlcNAc2P1Man6Fuc1GlcNAc2 (+2010)	2010.671

Table 6.1 continued from previous page

Gal1GlcNAc4Man3Fuc1GlcNAc2 (+2012)	2012.7455
Gal1Fuc1GlcNAc4Man3GlcNAc2 (+2012)	2012.7455
AcSia1Gal1GlcNAc1Man5Fuc1GlcNAc2 (+2018)	2018.7083
Gal1GlcNAc2P1Man6GlcNAc2 (+2026)	2026.6659
Glc1Man9GlcNAc2 (+2026)	2026.6868
Gal2GlcNAc4Man3GlcNAc2 (+2028)	2028.7404
AcSia1Gal1GlcNAc1Man6GlcNAc2 (+2034)	2034.7032
Gal1Fuc1GlcNAc1Man6Fuc1GlcNAc2 (+2035)	2035.7236
GlcNAc5Man3Fuc1GlcNAc2 (+2053)	2053.7721
AcSia1Gal21GlcNAc2Man3Fuc1GlcNAc2 (+2059)	2059.7349
Gal2Fuc2GlcNAc2Man3Fuc1GlcNAc2 (+2060)	2060.7553
Gal1GlcNAc5Man3GlcNAc2 (+2069)	2069.767
AcSia1Gal1GlcNAc2Man5GlcNAc2 (+2075)	2075.7298
Gal1Fuc1GlcNAc2Man5Fuc1GlcNAc2 (+2076)	2076.7502
Gal1Fuc1GlcNAc2Man6GlcNAc2 (+2092)	2092.7451
Gal1GlcNAc2Man6Fuc1GlcNAc2 (+2092)	2092.7451
AcSia1Gal1GlcNAc3Man3Fuc1GlcNAc2 (+2100)	2100.7615
AcSia1Gal1GlcNAc1P1Man6GlcNAc2 (+2114)	2114.6819
Gal1Fuc1GlcNAc1P1Man6Fuc1GlcNAc2 (+2115)	2115.7023
AcSia1Gal2GlcNAc3Man3GlcNAc2 (+2116)	2116.7564
Gal2Fuc2GlcNAc3Man3GlcNAc2 (+2117)	2117.7768
Gal2Fuc1GlcNAc3Man3Fuc1GlcNAc2 (+2117)	2117.7768
Gal3GlcNAc3Man3Fuc1GlcNAc2 (+2133)	2133.7717
Gal3Fuc1GlcNAc3Man3GlcNAc2 (+2133)	2133.7717
AcSia1Gal1GlcNAc4Man3GlcNAc2 (+2157)	2157.783
Gal1Fuc1GlcNAc4Man3Fuc1GlcNAc2 (+2158)	2158.8034
Gal1Fuc1GlcNAc2P1Man6GlcNAc2 (+2172)	2172.7238
Gal1GlcNAc2P1Man6Fuc1GlcNAc2 (+2172)	2172.7238
Gal2GlcNAc4Man3Fuc1GlcNAc2 (+2174)	2174.7983
Gal2Fuc1GlcNAc4Man3GlcNAc2 (+2174)	2174.7983
AcSia1Gal1GlcNAc1Man6Fuc1GlcNAc2 (+2180)	2180.7611
Gal3GlcNAc4Man3GlcNAc2 (+2190)	2190.7932
AcSia2Gal2GlcNAc2Man3GlcNAc2 (+2204)	2204.7724
AcSia1Gal2Fuc1GlcNAc2Man3Fuc1GlcNAc2 (+2205)	2205.7928
Gal1GlcNAc5Man3Fuc1GlcNAc2 (+2215)	2215.8249
Gal1Fuc1GlcNAc5Man3GlcNAc2 (+2215)	2215.8249

Table 6.1 continued from previous page

AcSia1Gal1GlcNAc2Man5Fuc1GlcNAc2 (+2221)	2221.7877
Gal2GlcNAc5Man3GlcNAc2 (+2231)	2231.8198
AcSia1Gal1GlcNAc2Man6GlcNAc2 (+2237)	2237.7826
Gal1Fuc1GlcNAc2Man6Fuc1GlcNAc2 (+2238)	2238.803
AcSia1Gal1GlcNAc1P1Man6Fuc1GlcNAc2 (+2260)	2260.7398
AcSia1Gal2GlcNAc3Man3Fuc1GlcNAc2 (+2262)	2262.8143
Gal2Fuc2GlcNAc3Man3Fuc1GlcNAc2 (+2263)	2263.8347
AcSia1Gal3GlcNAc3Man3GlcNAc2 (+2278)	2278.8092
Gal3Fuc2GlcNAc3Man3GlcNAc2 (+2279)	2279.8296
Gal3Fuc1GlcNAc3Man3Fuc1GlcNAc2 (+2279)	2279.8296
AcSia1Gal1GlcNAc4Man3Fuc1GlcNAc2 (+2303)	2303.8409
AcSia1Gal1GlcNAc2P1Man6GlcNAc2 (+2317)	2317.7613
Gal1Fuc1GlcNAc2P1Man6Fuc1GlcNAc2 (+2318)	2318.7817
AcSia1Gal2GlcNAc4Man3GlcNAc2 (+2319)	2319.8358
Gal2Fuc2GlcNAc4Man3GlcNAc2 (+2320)	2320.8562
Gal2Fuc1GlcNAc4Man3Fuc1GlcNAc2 (+2320)	2320.8562
Gal3GlcNAc4Man3Fuc1GlcNAc2 (+2336)	2336.8511
Gal3Fuc1GlcNAc4Man3GlcNAc2 (+2336)	2336.8511
AcSia2Gal2GlcNAc2Man3Fuc1GlcNAc2 (+2350)	2350.8303
Gal4GlcNAc4Man3GlcNAc2 (+2352)	2352.846
AcSia1Gal1GlcNAc5Man3GlcNAc2 (+2360)	2360.8624
Gal1Fuc1GlcNAc5Man3Fuc1GlcNAc2 (+2361)	2361.8828
Gal2GlcNAc5Man3Fuc1GlcNAc2 (+2377)	2377.8777
Gal2Fuc1GlcNAc5Man3GlcNAc2 (+2377)	2377.8777
AcSia1Gal1GlcNAc2Man6Fuc1GlcNAc2 (+2383)	2383.8405
Gal3GlcNAc5Man3GlcNAc2 (+2393)	2393.8726
AcSia2Gal2GlcNAc3Man3GlcNAc2 (+2407)	2407.8518
AcSia1Gal3GlcNAc3Man3Fuc1GlcNAc2 (+2424)	2424.8671
Gal3Fuc3GlcNAc4Man3GlcNAc2 (+2425)	2425.8875
Gal3Fuc2GlcNAc3Man3Fuc1GlcNAc2 (+2425)	2425.8875
AcSia1Gal1GlcNAc2P1Man6Fuc1GlcNAc2 (+2463)	2463.8192
AcSia1Gal2GlcNAc4Man3Fuc1GlcNAc2 (+2465)	2465.8937
Gal2Fuc2GlcNAc4Man3Fuc1GlcNAc2 (+2466)	2466.9141
AcSia1Gal3GlcNAc4Man3GlcNAc2 (+2481)	2481.8886
Gal3Fuc2GlcNAc4Man3GlcNAc2 (+2482)	2482.909
Gal3Fuc1GlcNAc4Man3Fuc1GlcNAc2 (+2482)	2482.909

Table 6.1 continued from previous page

Gal4GlcNAc4Man3Fuc1GlcNAc2 (+2498)	2498.9039
Gal4Fuc1GlcNAc4Man3GlcNAc2 (+2498)	2498.9039
AcSia1Gal1GlcNAc5Man3Fuc1GlcNAc2 (+2506)	2506.9203
AcSia1Gal2GlcNAc5Man3GlcNAc2 (+2522)	2522.9152
Gal2Fuc2GlcNAc5Man3GlcNAc2 (+2523)	2523.9356
Gal2Fuc1GlcNAc5Man3Fuc1GlcNAc2 (+2523)	2523.9356
Gal3GlcNAc5Man3Fuc1GlcNAc2 (+2539)	2539.9305
Gal3Fuc1GlcNAc5Man3GlcNAc2 (+2539)	2539.9305
AcSia2Gal2GlcNAc3Man3Fuc1GlcNAc2 (+2553)	2553.9097
Gal4GlcNAc5Man3GlcNAc2 (+2555)	2555.9254
AcSia2Gal3GlcNAc3Man3GlcNAc2 (+2569)	2569.9046
Gal3Fuc3GlcNAc3Man3Fuc1GlcNAc2 (+2571)	2571.9454
AcSia2Gal2GlcNAc4Man3GlcNAc2 (+2610)	2610.9312
AcSia1Gal3GlcNAc4Man3Fuc1GlcNAc2 (+2627)	2627.9465
Gal3Fuc3GlcNAc4Man3GlcNAc2 (+2628)	2628.9669
Gal3Fuc2GlcNAc4Man3Fuc1GlcNAc2 (+2628)	2628.9669
AcSia1Gal4GlcNAc4Man3GlcNAc2 (+2643)	2643.9414
Gal4Fuc2GlcNAc4Man3GlcNAc2 (+2644)	2644.9618
Gal4Fuc1GlcNAc4Man3Fuc1GlcNAc2 (+2644)	2644.9618
AcSia1Gal2GlcNAc5Man3Fuc1GlcNAc2 (+2668)	2668.9731
Gal2Fuc2GlcNAc5Man3Fuc1GlcNAc2 (+2669)	2669.9935
AcSia1Gal3GlcNAc5Man3GlcNAc2 (+2684)	2684.968
Gal3Fuc2GlcNAc5Man3GlcNAc2 (+2685)	2685.9884
Gal3Fuc1GlcNAc5Man3Fuc1GlcNAc2 (+2685)	2685.9884
Gal4GlcNAc5Man3Fuc1GlcNAc2 (+2701)	2701.9833
Gal4Fuc1GlcNAc5Man3GlcNAc2 (+2701)	2701.9833
AcSia2Gal3GlcNAc3Man3Fuc1GlcNAc2 (+2715)	2715.9625
AcSia2Gal2GlcNAc4Man3Fuc1GlcNAc2 (+2756)	2756.9891
AcSia2Gal3GlcNAc4Man3GlcNAc2 (+2772)	2772.984
Gal3Fuc3GlcNAc4Man3Fuc1GlcNAc2 (+2775)	2775.0248
AcSia1Gal4GlcNAc4Man3Fuc1GlcNAc2 (+2789)	2789.9993
Gal4Fuc3GlcNAc4Man3GlcNAc2 (+2791)	2791.0197
Gal4Fuc2GlcNAc4Man3Fuc1GlcNAc2 (+2791)	2791.0197
AcSia2Gal2GlcNAc5Man3GlcNAc2 (+2814)	2814.0106
AcSia1Gal3GlcNAc5Man3Fuc1GlcNAc2 (+2831)	2831.0259
Gal3Fuc3GlcNAc5Man3GlcNAc2 (+2832)	2832.0463

Table 6.1 continued from previous page

Gal3Fuc2GlcNAc5Man3Fuc1GlcNAc2 (+2832)	2832.0463
AcSia1Gal4GlcNAc5Man3GlcNAc2 (+2847)	2847.0208
Gal4Fuc2GlcNAc5Man3GlcNAc2 (+2848)	2848.0412
Gal4Fuc1GlcNAc5Man3Fuc1GlcNAc2 (+2848)	2848.0412
AcSia3Gal3GlcNAc3Man3GlcNAc2 (+2861)	2861.0000
AcSia2Gal3GlcNAc4Man3Fuc1GlcNAc2 (+2919)	2919.0419
AcSia2Gal4GlcNAc4Man3GlcNAc2 (+2935)	2935.0368
Gal4Fuc4GlcNAc4Man3GlcNAc2 (+2936)	2937.0776
Gal4Fuc3GlcNAc4Man3Fuc1GlcNAc2 (+2937)	2937.0776
AcSia2Gal2GlcNAc5Man3Fuc1GlcNAc2 (+2960)	2960.0685
AcSia2Gal3GlcNAc5Man3GlcNAc2 (+2976)	2976.0634
Gal3Fuc3GlcNAc5Man3Fuc1GlcNAc2 (+2978)	2978.1042
AcSia1Gal4GlcNAc5Man3Fuc1GlcNAc2 (+2993)	2993.0787
Gal4Fuc3GlcNAc5Man3GlcNAc2 (+2994)	2994.0991
Gal4Fuc2GlcNAc5Man3Fuc1GlcNAc2 (+2994)	2994.0991
AcSia3Gal3GlcNAc3Man3Fuc1GlcNAc2 (+3007)	3007.0579
AcSia3Gal3GlcNAc4Man3GlcNAc2 (+3064)	3064.0794
AcSia2Gal4GlcNAc4Man3Fuc1GlcNAc2 (+3081)	3081.0947
Gal4Fuc4GlcNAc4Man3Fuc1GlcNAc2 (+3083)	3083.1355
AcSia2Gal3GlcNAc5Man3Fuc1GlcNAc2 (+3122)	3122.1213
AcSia2Gal4GlcNAc5Man3GlcNAc2 (+3138)	3138.1162
Gal4Fuc4GlcNAc5Man3GlcNAc2 (+3140)	3140.157
Gal4Fuc3GlcNAc5Man3Fuc1GlcNAc2 (+3140)	3140.157
AcSia3Gal3GlcNAc4Man3Fuc1GlcNAc2 (+3210)	3210.1373
AcSia3Gal4GlcNAc4Man3GlcNAc2 (+3226)	3226.1322
AcSia3Gal3GlcNAc5Man3GlcNAc2 (+3267)	3267.1588
AcSia2Gal4GlcNAc5Man3Fuc1GlcNAc2 (+3284)	3284.1741
Gal4Fuc4GlcNAc5Man3Fuc1GlcNAc2 (+3286)	3286.2149
AcSia3Gal4GlcNAc4Man3Fuc1GlcNAc2 (+3372)	3372.1901
AcSia3Gal3GlcNAc5Man3Fuc1GlcNAc2 (+3413)	3413.2167
AcSia3Gal4GlcNAc5Man3GlcNAc2 (+3429)	3429.2116
AcSia4Gal4GlcNAc4Man3GlcNAc2 (+3517)	3517.2276
AcSia3Gal4GlcNAc5Man3Fuc1GlcNAc2 (+3575)	3575.2695
AcSia4Gal4GlcNAc4Man3Fuc1GlcNAc2 (+3663)	3663.2855
AcSia4Gal4GlcNAc5Man3GlcNAc2 (+3720)	3720.307
AcSia4Gal4GlcNAc5Man3Fuc1GlcNAc2 (+3866)	3866.3649

6.2 Chicken - Peptide masses and a-, b-, y-ions for MS/MS

6.2.1 MLNVSQIK

Table 6.2: Ions of MLNVSQIK

H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Peptide	**M**	**L**	**N**	**V**	**S**	**Q**	**I**	**K**
OH	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
b(1+)	132.05	245.13	359.18	458.24	545.28	673.33	786.42	
y(1+)		801.48	688.4	574.36	475.29	388.26	260.2	147.11
a(1+)	104.05	217.14	331.18	430.25	517.28	645.34	758.42	
b(2+)	66.53	123.07	180.09	229.63	273.14	337.17	393.71	
y(2+)		401.25	344.7	287.68	238.15	194.63	130.6	74.06
a(2+)	52.53	109.07	166.09	215.63	259.14	323.17	379.72	
b(3+)		82.38	120.4	153.42	182.43	225.12	262.81	
y(3+)		267.83	230.14	192.12	159.1	130.09	87.4	
a(3+)		73.05	111.07	144.09	173.1	215.79	253.48	
b-H20	114.04	227.12	341.16	440.23	527.27	655.32	768.41	896.5
b-NH3	115.02	228.11	342.15	441.22	528.25	656.31	769.39	897.49
y-H20	914.51	783.47	670.39	556.35	457.28	370.25	242.19	129.1
y-NH3	915.5	784.46	671.37	557.33	458.26	371.23	243.17	130.09
a-NH3	87.03	200.11	314.15	413.22	500.25	628.31	741.4	869.49

6.2.2 NILDALMLNTTR

Table 6.3: Ions of NILDALMLNTTR

H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Peptide	**N**	**I**	**L**	**D**	**A**	**L**	**M**	**L**
OH	**12**	**11**	**10**	**9**	**8**	**7**	**6**	**5**
b(1+)	115.05	228.13	341.22	456.25	527.28	640.37	771.41	884.49
y(1+)		1260.7	1147.61	1034.53	919.5	848.47	735.38	604.34
a(1+)	87.06	200.14	313.22	428.25	499.29	612.37	743.41	856.5
b(2+)	58.03	114.57	171.11	228.63	264.15	320.69	386.21	442.75
y(2+)		630.85	574.31	517.77	460.26	424.74	368.2	302.67
a(2+)	44.03	100.57	157.12	214.63	250.15	306.69	372.21	428.75
b(3+)		76.72	114.41	152.75	176.43	214.13	257.81	295.5
y(3+)		420.9	383.21	345.52	307.17	283.49	245.8	202.12
a(3+)		67.39	105.08	143.42	167.1	204.8	248.48	286.17
b-H20	97.04	210.12	323.21	438.24	509.27	622.36	753.4	866.48
b-NH3	98.02	211.11	324.19	439.22	510.26	623.34	754.38	867.46
y-H20	1356.73	1242.69	1129.6	1016.52	901.49	830.46	717.37	586.33
y-NH3	1357.71	1243.67	1130.59	1017.5	902.48	831.44	718.36	587.32
a-NH3	70.03	183.11	296.2	411.22	482.26	595.35	726.39	839.47

H	**9**	**10**	**11**	**12**
Peptide	**N**	**T**	**T**	**R**
OH	**4**	**3**	**2**	**1**
b(1+)	998.53	1099.58	1200.63	
y(1+)	491.26	377.21	276.17	175.12
a(1+)	970.54	1071.59	1172.63	
b(2+)	499.77	550.29	600.82	
y(2+)	246.13	189.11	138.59	88.06
a(2+)	485.77	536.3	586.82	
b(3+)	333.52	367.2	400.88	
y(3+)	164.42	126.41	92.73	
a(3+)	324.19	357.87	391.55	
b-H20	980.52	1081.57	1182.62	1338.72
b-NH3	981.51	1082.56	1183.6	1339.7
y-H20	473.25	359.2	258.16	157.11
y-NH3	474.23	360.19	259.14	158.09
a-NH3	953.51	1054.56	1155.61	1311.71

6.2.3 HPVAKNLSLK

Table 6.4: Ions of HPVAKNLSLK

H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Peptide	**H**	**P**	**V**	**A**	**K**	**N**	**L**	**S**
OH	**10**	**9**	**8**	**7**	**6**	**5**	**4**	**3**
b(1+)	138.07	235.12	334.19	405.23	533.32	647.36	760.45	847.48
y(1+)		969.61	872.56	773.49	702.45	574.36	460.31	347.23
a(1+)	110.07	207.12	306.19	377.23	505.33	619.37	732.45	819.48
b(2+)	69.54	118.06	167.6	203.12	267.16	324.19	380.73	424.24
y(2+)		485.31	436.78	387.25	351.73	287.68	230.66	174.12
a(2+)	55.54	104.07	153.6	189.12	253.17	310.19	366.73	410.25
b(3+)		79.05	112.07	135.75	178.45	216.46	254.15	283.16
y(3+)		323.88	291.52	258.5	234.82	192.12	154.11	116.42
a(3+)		69.71	102.74	126.42	169.11	207.13	244.82	273.83
b-H20	120.06	217.11	316.18	387.21	515.31	629.35	742.44	829.47
b-NH3	121.04	218.09	317.16	388.2	516.29	630.34	743.42	830.45
y-H20	1088.66	951.6	854.55	755.48	684.44	556.35	442.3	329.22
y-NH3	1089.64	952.58	855.53	756.46	685.42	557.33	443.29	330.2
a-NH3	93.05	190.1	289.17	360.2	488.3	602.34	715.43	802.46

H	**9**	**10**
Peptide	**L**	**K**
OH	**2**	**1**
b(1+)	960.56	
y(1+)	260.2	147.11
a(1+)	932.57	
b(2+)	480.79	
y(2+)	130.6	74.06
a(2+)	466.79	
b(3+)	320.86	
y(3+)	87.4	
a(3+)	311.53	
b-H20	942.55	1070.65
b-NH3	943.54	1071.63
y-H20	242.19	129.1
y-NH3	243.17	130.09
a-NH3	915.54	1043.64

6.3 Sheep - Peptide masses and a-, b-, y-ions for MS/MS

6.3.1 GLHNVSDFDNSLLR

Table 6.5: Ions of GLHNVSDFDNSLLR

H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Peptide	**G**	**L**	**H**	**N**	**V**	**S**	**D**	**F**
OH	**14**	**13**	**12**	**11**	**10**	**9**	**8**	**7**
b(1+)	58.03	171.11	308.17	422.22	521.28	608.32	723.34	870.41
y(1+)		1529.77	1416.69	1279.63	1165.59	1066.52	979.48	864.46
a(1+)	30.03	143.12	280.18	394.22	493.29	580.32	695.35	842.42
b(2+)	29.52	86.06	154.59	211.61	261.15	304.66	362.18	435.71
y(2+)		765.39	708.85	640.32	583.3	533.76	490.25	432.73
a(2+)	15.52	72.06	140.59	197.61	247.15	290.66	348.18	421.71
b(3+)		57.71	103.4	141.41	174.43	203.44	241.79	290.81
y(3+)		510.6	472.9	427.21	389.2	356.18	327.17	288.82
a(3+)		48.38	94.06	132.08	165.1	194.11	232.45	281.48
b-H20	40.02	153.1	290.16	404.2	503.27	590.31	705.33	852.4
b-NH3	41	154.09	291.15	405.19	504.26	591.29	706.32	853.38
y-H20	1568.78	1511.76	1398.68	1261.62	1147.57	1048.51	961.47	846.45
y-NH3	1569.77	1512.74	1399.66	1262.6	1148.56	1049.49	962.46	847.43
a-NH3	13.01	126.09	263.15	377.19	476.26	563.29	678.32	825.39

H	**9**	**10**	**11**	**12**	**13**	**14**
Peptide	**D**	**N**	**S**	**L**	**L**	**R**
OH	**6**	**5**	**4**	**3**	**2**	**1**
b(1+)	985.44	1099.48	1186.51	1299.6	1412.68	1568.78
y(1+)	717.39	602.36	488.32	401.29	288.2	175.12
a(1+)	957.44	1071.49	1158.52	1271.6	1384.69	1540.79
b(2+)	493.22	550.24	593.76	650.3	706.84	784.89
y(2+)	359.2	301.69	244.66	201.15	144.61	88.06
a(2+)	479.23	536.25	579.76	636.3	692.85	770.9
b(3+)	329.15	367.17	396.18	433.87	471.57	523.6
y(3+)	239.8	201.46	163.45	134.43	96.74	59.05
a(3+)	319.82	357.83	386.84	424.54	462.23	514.27
b-H20	967.43	1081.47	1168.5	1281.59	1394.67	1550.77
b-NH3	968.41	1082.45	1169.49	1282.57	1395.65	1551.76
y-H20	699.38	584.35	470.31	383.28	270.19	157.11
y-NH3	700.36	585.34	471.29	384.26	271.18	158.09
a-NH3	940.42	1054.46	1141.49	1254.58	1367.66	1523.76

6.3.2 ASLYPVYELNGTVYSQEWLVR

Table 6.6: Ions of ASLYPVYELNGTVYSQEWLVR

H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Peptide	**A**	**S**	**L**	**Y**	**P**	**V**	**Y**	**E**
OH	**21**	**20**	**19**	**18**	**17**	**16**	**15**	**14**
b(1+)	72.04	159.08	272.16	435.22	532.28	631.35	794.41	923.45
y(1+)		2416.22	2329.19	2216.1	2053.04	1955.99	1856.92	1693.85
a(1+)	44.05	131.08	244.17	407.23	504.28	603.35	766.41	895.46
b(2+)	36.53	80.04	136.58	218.12	266.64	316.18	397.71	462.23
y(2+)		1208.61	1165.1	1108.56	1027.02	978.5	928.96	847.43
a(2+)	22.53	66.04	122.59	204.12	252.65	302.18	383.71	448.23
b(3+)		53.7	91.39	145.75	178.1	211.12	265.47	308.49
y(3+)		806.08	777.07	739.37	685.02	652.67	619.64	565.29
a(3+)		44.37	82.06	136.42	168.77	201.79	256.14	299.16
b-H20	54.03	141.07	254.15	417.21	514.27	613.33	776.4	905.44
b-NH3	55.02	142.05	255.13	418.2	515.25	614.32	777.38	906.42
y-H20	2469.25	2398.21	2311.18	2198.09	2035.03	1937.98	1838.91	1675.84
y-NH3	2470.23	2399.19	2312.16	2199.08	2036.01	1938.96	1839.89	1676.83
a-NH3	27.02	114.06	227.14	390.2	487.26	586.32	749.39	878.43
H	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Peptide	**L**	**N**	**G**	**T**	**V**	**Y**	**S**	**Q**
OH	**13**	**12**	**11**	**10**	**9**	**8**	**7**	**6**
b(1+)	1036.54	1150.58	1207.6	1308.65	1407.72	1570.78	1657.81	1785.87
y(1+)	1564.81	1451.73	1337.69	1280.66	1179.62	1080.55	917.48	830.45
a(1+)	1008.54	1122.58	1179.6	1280.65	1379.72	1542.78	1629.82	1757.88
b(2+)	518.77	575.79	604.3	654.83	704.36	785.89	829.41	893.44
y(2+)	782.91	726.37	669.35	640.84	590.31	540.78	459.25	415.73
a(2+)	504.77	561.8	590.31	640.83	690.36	771.9	815.41	879.44
b(3+)	346.18	384.2	403.21	436.89	469.91	524.26	553.28	595.96
y(3+)	522.28	484.58	446.57	427.56	393.88	360.85	306.5	277.49
a(3+)	336.85	374.87	393.87	427.56	460.58	514.93	543.94	586.63
b-H20	1018.52	1132.57	1189.59	1290.64	1389.71	1552.77	1639.8	1767.86
b-NH3	1019.51	1133.55	1190.57	1291.62	1390.69	1553.75	1640.78	1768.84
y-H20	1546.8	1433.72	1319.67	1262.65	1161.61	1062.54	899.47	812.44

Table 6.6 continued from previous page

y-NH3	1547.79	1434.7	1320.66	1263.64	1162.59	1063.52	900.46	813.43
a-NH3	991.51	1105.56	1162.58	1263.63	1362.69	1525.76	1612.79	1740.85
H	**17**	**18**	**19**	**20**	**21**			
Peptide	**E**	**W**	**L**	**V**	**R**			
OH	**5**	**4**	**3**	**2**	**1**			
b(1+)	1914.91	2100.99	2214.08	2313.14				
y(1+)	702.39	573.35	387.27	274.19	175.12			
a(1+)	1886.92	2073	2186.08	2285.15				
b(2+)	957.96	1051	1107.54	1157.08				
y(2+)	351.7	287.18	194.14	137.6	88.06			
a(2+)	943.96	1037	1093.54	1143.08				
b(3+)	638.98	701	738.7	771.72				
y(3+)	234.8	191.79	129.76	92.07				
a(3+)	629.64	691.67	729.37	762.39				
b-H20	1896.9	2082.98	2196.07	2295.13	2451.23			
b-NH3	1897.89	2083.97	2197.05	2296.12	2452.22			
y-H20	684.38	555.34	369.26	256.18	157.11			
y-NH3	685.37	556.32	370.25	257.16	158.09			
a-NH3	1869.89	2055.97	2169.05	2268.12	2424.22			

6.3.3 TKNESLIK

Table 6.7: Ions of TKNESLIK

H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Peptide	**T**	**K**	**N**	**E**	**S**	**L**	**I**	**K**
OH	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
b(1+)	102.06	230.15	344.19	473.24	560.27	673.35	786.44	
y(1+)		831.49	703.4	589.36	460.31	373.28	260.2	147.11
a(1+)	74.06	202.16	316.2	445.24	532.27	645.36	758.44	
b(2+)	51.53	115.58	172.6	237.12	280.64	337.18	393.72	
y(2+)		416.25	352.2	295.18	230.66	187.14	130.6	74.06
a(2+)	37.53	101.58	158.6	223.12	266.64	323.18	379.72	
b(3+)		77.39	115.4	158.42	187.43	225.12	262.82	
y(3+)		277.84	235.14	197.12	154.11	125.1	87.4	
a(3+)		68.06	106.07	149.09	178.1	215.79	253.49	
b-H20	84.04	212.14	326.18	455.23	542.26	655.34	768.43	896.52
b-NH3	85.03	213.12	327.17	456.21	543.24	656.33	769.41	897.5
y-H20	914.53	813.48	685.39	571.35	442.3	355.27	242.19	129.1
y-NH3	915.52	814.47	686.37	572.33	443.29	356.25	243.17	130.09
a-NH3	57.03	185.13	299.17	428.21	515.25	628.33	741.41	869.51

6.3.4 NESLIK

Table 6.8: Ions of NESLIK

H	**1**	**2**	**3**	**4**	**5**	**6**
Peptide	**N**	**E**	**S**	**L**	**I**	**K**
OH	**6**	**5**	**4**	**3**	**2**	**1**
b(1+)	115.05	244.09	331.13	444.21	557.29	
y(1+)		589.36	460.31	373.28	260.2	147.11
a(1+)	87.06	216.1	303.13	416.21	529.3	
b(2+)	58.03	122.55	166.07	222.61	279.15	
y(2+)		295.18	230.66	187.14	130.6	74.06
a(2+)	44.03	108.55	152.07	208.61	265.15	
b(3+)		82.04	111.05	148.74	186.44	
y(3+)		197.12	154.11	125.1	87.4	
a(3+)		72.7	101.72	139.41	177.1	
b-H20	97.04	226.08	313.11	426.2	539.28	667.38
b-NH3	98.02	227.07	314.1	427.18	540.27	668.36
y-H20	685.39	571.35	442.3	355.27	242.19	129.1
y-NH3	686.37	572.33	443.29	356.25	243.17	130.09
a-NH3	70.03	199.07	286.1	399.19	512.27	640.37

6.3.5 NLLDAVILNSTIR

Table 6.9: Ions of NLLDAVILNSTIR

H	**1**	**2**	**3**	**4**	**5**	**6**	7	8
Peptide	**N**	**L**	**L**	**D**	**A**	**V**	I	L
OH	**12**	**11**	**10**	**9**	**8**	**7**	6	5
b(1+)	115.05	228.13	341.22	456.25	527.28	626.35	739.44	852.52
y(1+)		1214.71	1101.63	988.54	873.52	802.48	703.41	590.33
a(1+)	87.06	200.14	313.22	428.25	499.29	598.36	711.44	824.52
b(2+)	58.03	114.57	171.11	228.63	264.15	313.68	370.22	426.76
y(2+)		607.86	551.32	494.78	437.26	401.74	352.21	295.67
a(2+)	44.03	100.57	157.12	214.63	250.15	299.68	356.22	412.77
b(3+)		76.72	114.41	152.75	176.43	209.46	247.15	284.85
y(3+)		405.58	367.88	330.19	291.84	268.16	235.14	197.45
a(3+)		67.39	105.08	143.42	167.1	200.12	237.82	275.51
b-H20	97.04	210.12	323.21	438.24	509.27	608.34	721.42	834.51
b-NH3	98.02	211.11	324.19	439.22	510.26	609.32	722.41	835.49
y-H20	1310.74	1196.7	1083.62	970.53	855.51	784.47	685.4	572.32
y-NH3	1311.73	1197.68	1084.6	971.52	856.49	785.45	686.38	573.3
a-NH3	70.03	183.11	296.2	411.22	482.26	581.33	694.41	807.5

H	**9**	**10**	**11**	**12**
Peptide	**N**	**S**	**T**	**R**
OH	**4**	**3**	**2**	**1**
b(1+)	966.56	1053.59	1154.64	
y(1+)	477.24	363.2	276.17	175.12
a(1+)	938.57	1025.6	1126.65	
b(2+)	483.79	527.3	577.82	
y(2+)	239.12	182.1	138.59	88.06
a(2+)	469.79	513.3	563.83	
b(3+)	322.86	351.87	385.55	
y(3+)	159.75	121.74	92.73	
a(3+)	313.53	342.54	376.22	
b-H20	948.55	1035.58	1136.63	1292.73
b-NH3	949.54	1036.57	1137.62	1293.72
y-H20	459.23	345.19	258.16	157.11
y-NH3	460.22	346.17	259.14	158.09
a-NH3	921.54	1008.57	1109.62	1265.72

Chapter 7

Appendix MS/MS results and spectra

7.1 Results - PD analysis of chicken

231017_OKGT_Kylling_55kDa.xlsx

231017_OKGT_Kylling_60kDa.xlsx

7.2 Results - PD analysis of sheep

2310177_OKGT_Sau_55kDa.xlsx

231017_OKGT_Sau_60kDa.xlsx

230427_OKGTollersrud_OKGT_sau_S1.xlsx

7.3 Spectra

231017_OKGT_Kylling_55kDa.raw #25967 RT: 27.9544 min: MLNVSQIK-GlcNAc2Man

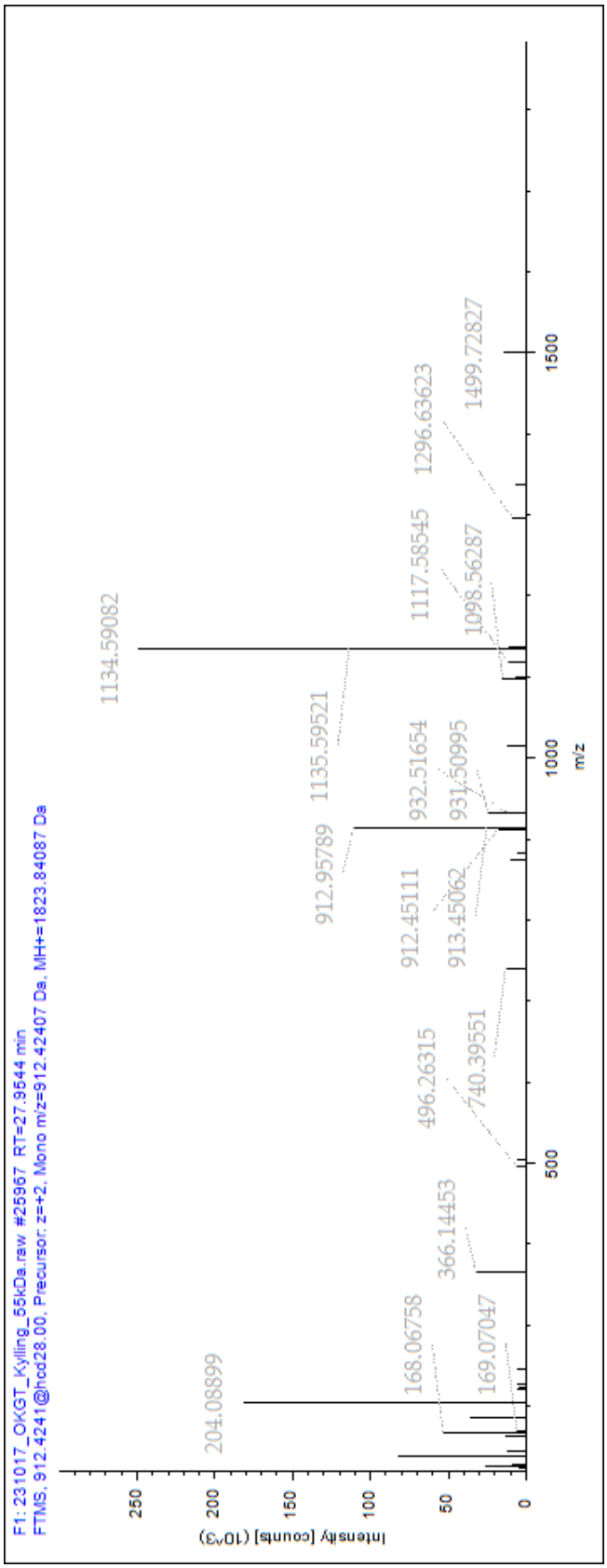


Figure 7.1: Chicken 55kDa #25967

Table 7.1: Chicken 55kDa #25967

m/z	Intensity [counts]	*m/z*	Intensity [counts]
123.05495	5025	740.39551	13098
124.58801	5217	873.49939	10371
126.05643	26133	883.51465	5967
127.04060	9041	911.94617	7263
138.05661	82584	912.45111	17637
145.05185	12624	912.95789	110883
163.06203	12484	913.45062	25419
168.06758	53408	931.50995	24285
169.07047	5900	932.51654	12185
186.07880	35627	1014.54779	12053
204.08899	181785	1098.56287	15186
205.09343	7446	1099.58496	6286
221.12950	5195	1117.58545	11244
222.09897	5105	1134.59082	249004
227.22121	5334	1135.59521	113725
245.13451	5856	1136.60852	11042
366.14453	32161	1296.63623	8852
496.26315	6192	1337.66931	6941
505.26288	5960	1499.72827	14278

231017_OKGT_Kylling_55kDa.raw #4542 RT: 11.1545 min: HPVAKNLSLK-GlcNAc

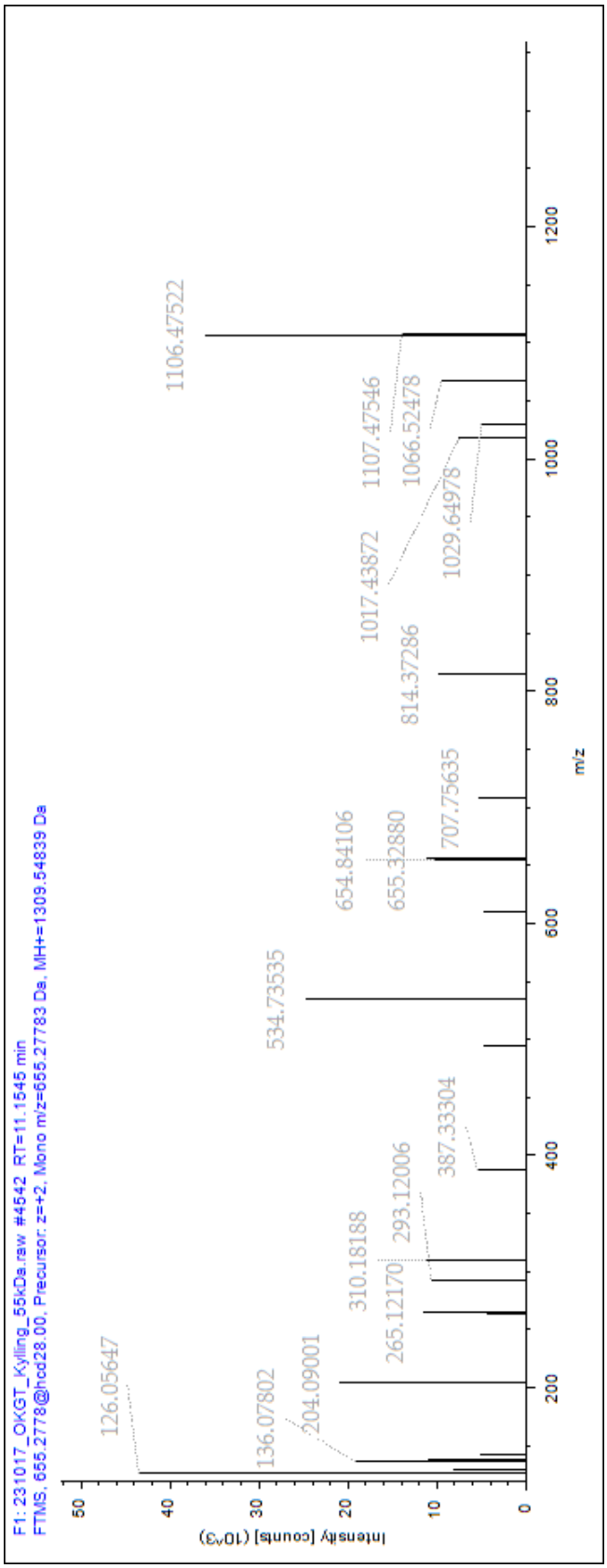


Figure 7.2: Chicken 55kda #4542

Table 7.2: Chicken 55kda #4542

m/z	Intensity [counts]
126.05647	43529
129.10460	8205
136.07802	19188
138.05711	11008
142.46548	5023
204.09001	20936
205.20036	4725
263.71558	4290
265.12170	11435
293.12006	10619
310.18188	11266
387.33304	5496
388.05457	4485
494.66531	4698
534.73535	24734
609.65063	4682
654.84106	10355
655.32880	11275
656.27289	6543
707.75635	5274
814.37286	9801
1017.43872	7546
1029.64978	4995
1066.52478	9539
1067.53552	7307
1106.47522	36047
1107.47546	13979

231017_OKGT_Kylling_55kDa.raw #4617 RT: 11.2163 min: HPVAKNLSLK-GlcNAc2Man2

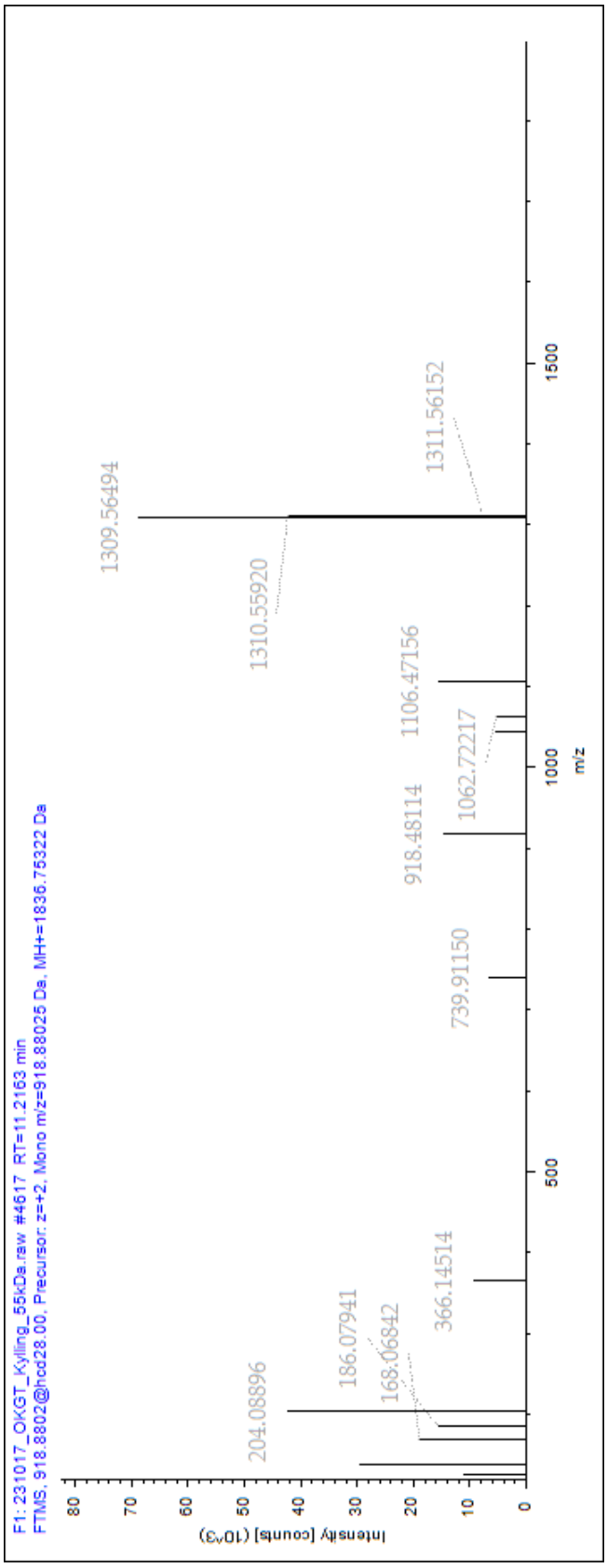


Figure 7.3: Chicken 55kda #4617

Table 7.3: Chicken 55kda #4617

m/z	Intensity [counts]
126.05705	11016
138.05658	29684
168.06842	18931
186.07941	15743
204.08896	42427
366.14514	9471
739.91150	6527
918.48114	14742
1044.57068	5602
1062.72217	5298
1106.47156	15461
1309.56494	68819
1310.55920	42339
1311.56152	7648

231017_OKGT_Kylling_55kDa.raw #4909 RT: 11.4757 min: HPVAKNLSLK-GlcNAcFuc

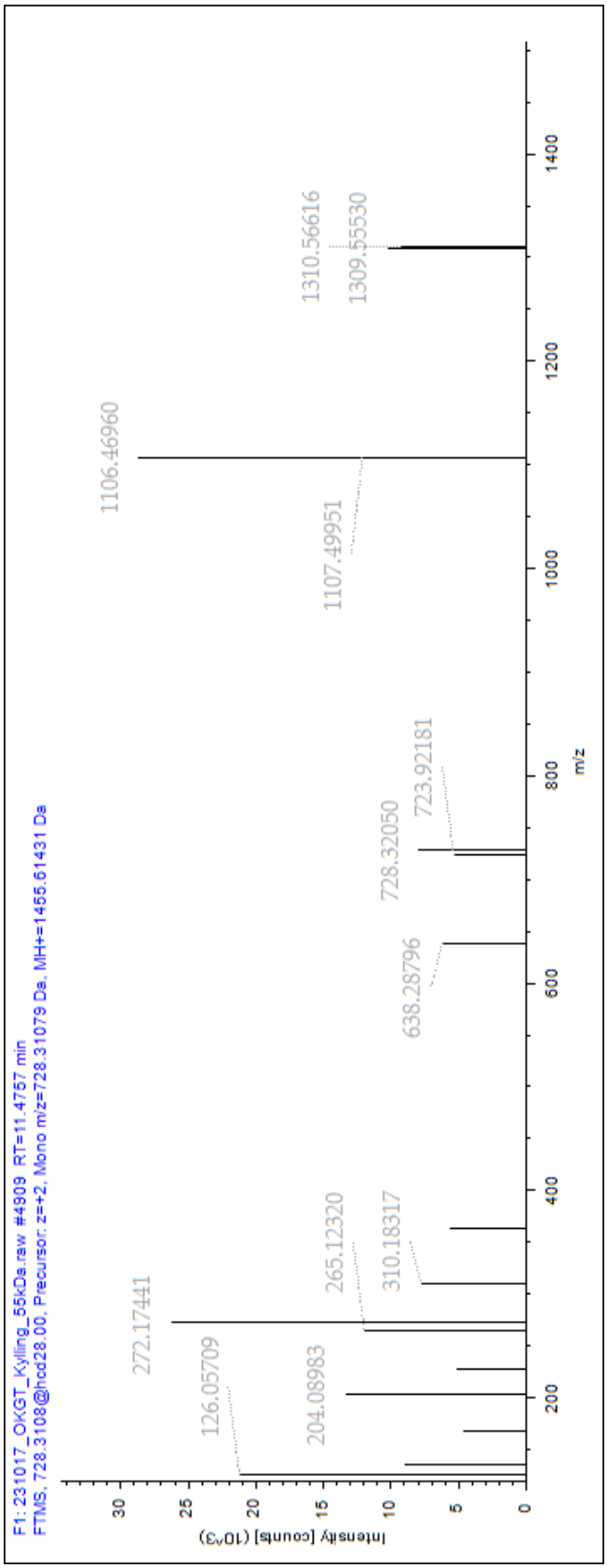


Figure 7.4: Chicken 55kda #4909

Table 7.4: Chicken 55kda #4909

m/z	Intensity [counts]
126.05709	21220
136.07698	9023
168.26805	4689
204.08983	13366
228.39624	5158
265.12320	11969
272.17441	26223
310.18317	7761
363.08612	5636
638.28796	6245
723.92181	5360
728.32050	7924
1106.46960	28689
1107.49951	12107
1309.55530	10259
1310.56616	9344

231017_OKGT_Kylling_55kDa.raw #5420 RT: 11.9474 min: NILDALMLNTTR-GlcNAc2Man4

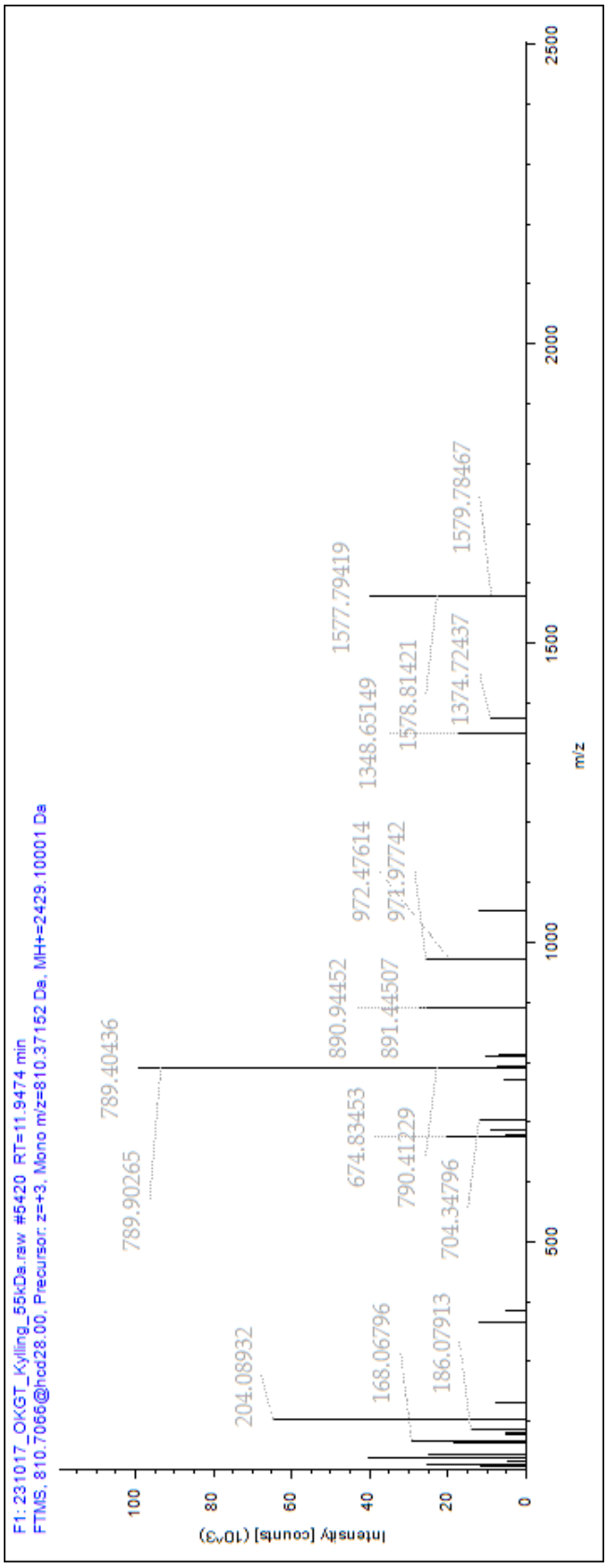


Figure 7.5: Chicken 55kda #5420

Table 7.5: Chicken 55kda #5420

m/z	Intensity [counts]	*m/z*	Intensity [counts]
126.05698	11579	771.39990	5898
127.04060	25361	789.40436	99153
129.10426	6530	789.90265	93458
134.60713	5075	790.41229	22820
138.05664	40685	790.90637	10595
139.06110	5740	792.37946	7339
145.05078	24920	810.36786	10495
163.06284	18758	811.36066	6988
168.06796	29264	890.94452	25580
179.28024	5190	891.44507	27487
179.96338	5443	891.93915	7749
186.07913	14212	971.97742	25623
204.08932	64826	972.47614	19594
230.15469	7722	972.97919	11561
366.14566	12072	1053.50952	12127
384.70184	5183	1348.65149	17484
674.83453	20606	1349.67053	12597
675.33917	8120	1374.72437	9227
679.36493	5218	1577.79419	39948
687.86932	9013	1578.81421	22761
704.34796	12015	1579.78467	8929

231017_OKGT_Kylling_55kDa.raw #42651 RT: 39.8784 min: SLIEQSAVFSILK

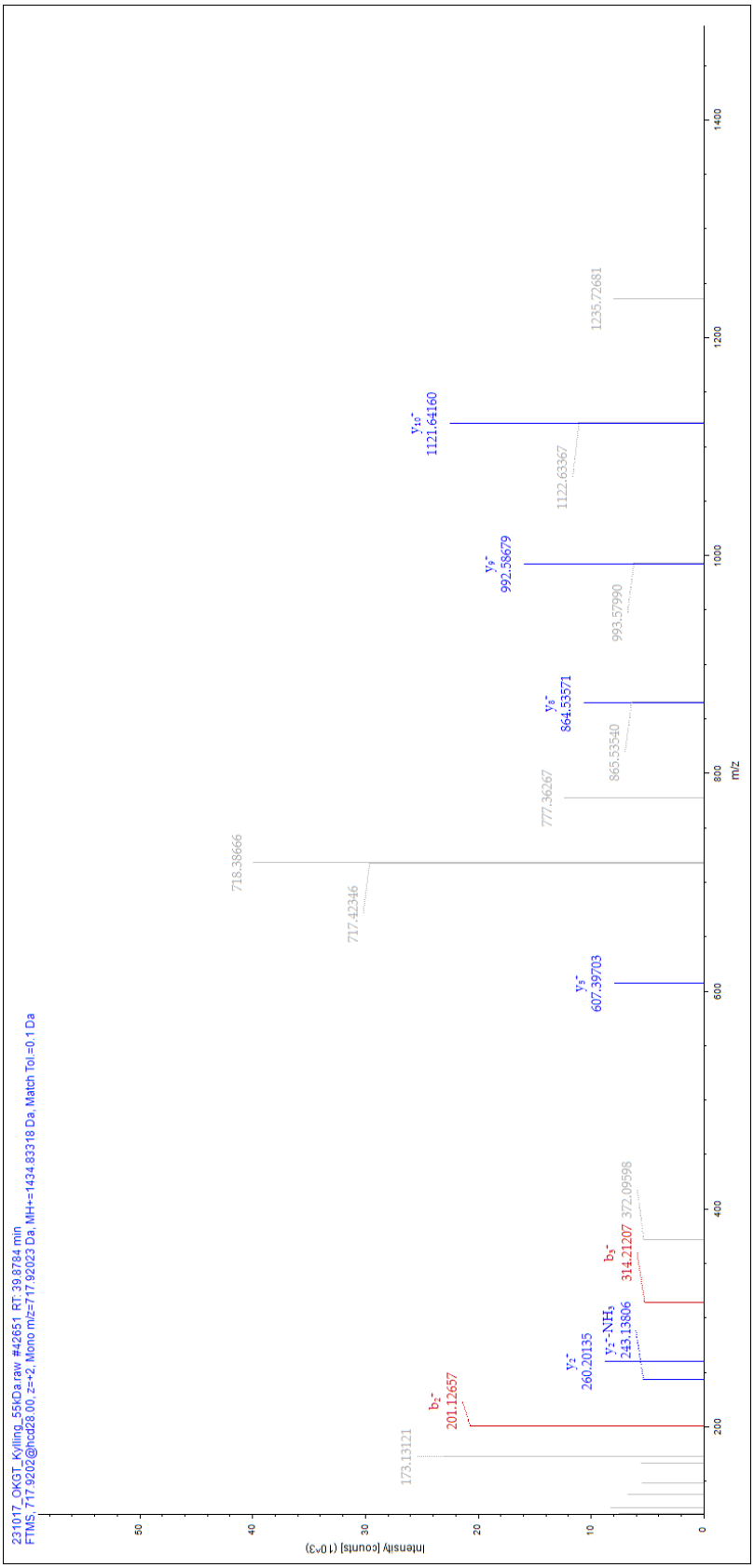


Figure 7.6: Chicken 55kDa #42651

Table 7.6: Chicken 55kDa #42651

m/z	Intensity [counts]		
120.07061	4940		
120.08217	7521		
126.05707	8318		
138.05797	6713		
148.33679	5463		
166.68120	5517		
173.13121	22984		
201.12657	20743	b_2^+	201.12657
243.13806	5360	y_2^+-NH_3	243.13806
260.20135	8771	y_2^+	260.20135
314.21207	5250	b_3^+	314.21207
372.09598	5342		
607.39703	7907	y_5^+	607.39703
717.42346	29632		
718.38666	39943		
777.36267	12437		
864.53571	10670	y_8^+	864.53571
865.53540	6422		
992.58679	15948	y_9^+	992.58679
993.57990	6214		
1121.64160	22552	y_{10}^+	1121.64160
1122.63367	11068		
1235.72681	8040		

231017_OKGT_Kylling_60kDa.raw #21699 RT: 28.5640 min: MLNVSQIK-GlcNAc2Man2

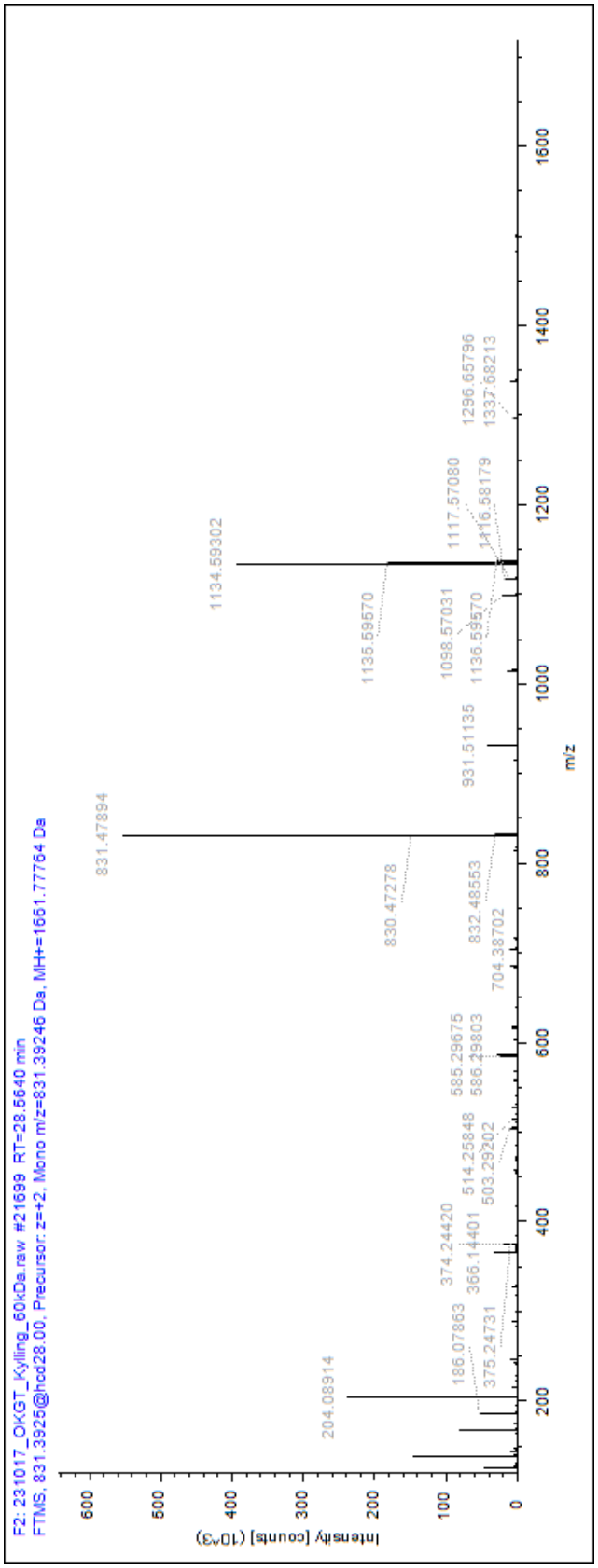


Figure 7.7: Chicken 60kDa #21699

Table 7.7: Chicken 60kDa #21699

m/z	Intensity [counts]	*m/z*	Intensity [counts]	*m/z*	Intensity [counts]
126.05617	47867	373.13992	1543	684.36328	5519
127.04025	4448	374.24420	20714	685.37097	8759
129.10349	3188	375.24731	9803	704.38702	10225
131.08754	2986	384.15869	2258	705.38074	3559
135.08084	1558	417.25348	1779	715.44330	3882
138.05673	145975	454.19806	2708	716.44769	4907
139.05983	6059	456.25220	1817	721.37024	1663
144.06714	9628	457.23911	4286	813.45532	6124
145.05130	2534	458.24371	2876	814.44873	4332
147.11470	5165	468.21280	2342	817.46985	2527
163.06230	2311	470.27060	2050	830.47278	148940
168.06776	81862	471.26843	1811	831.47894	553431
169.07137	3475	472.21072	2287	832.48553	31184
175.12102	6254	473.29086	1484	914.48584	5915
186.07863	53379	485.27515	2011	915.48474	2403
187.08127	1458	497.24741	1542	931.51135	41062
187.11188	3740	503.29202	11699	932.52002	14589
195.10342	1675	504.29031	7442	1014.55109	14230
204.08914	237612	505.26190	6892	1015.55957	7318
205.09229	14729	505.77130	1653	1039.53235	1732
215.10641	6860	511.27170	2560	1098.57031	21751
222.09897	2297	514.25848	8415	1099.57544	8351
228.13744	3361	515.25903	4419	1100.57019	1800
241.12170	2092	520.27539	3395	1116.58179	18409
243.13693	4295	528.19879	6726	1117.57080	11005
246.18422	8425	531.78143	3258	1118.58545	3577
247.18707	1783	540.78687	3978	1134.59302	394721
258.11362	1414	557.30176	6256	1135.59570	182547
283.17319	1486	558.30005	4097	1136.59570	29390
288.20709	7366	567.28058	3786	1156.60571	2014
300.15939	3974	568.28424	4641	1296.65796	7287
317.18085	2035	585.29675	25244	1297.63171	2570
317.22379	3653	586.29803	28241	1337.68213	9790
328.19186	7025	603.33484	4411	1338.67725	4261

Table 7.7 continued from previous page

329.18982	3015	616.37994	8460	1482.73535	1551
366.14401	33688	617.37811	7225	1499.73657	3102
367.14719	2775	639.34973	1871	1500.71777	1837
369.22989	1669	644.32520	1874		
371.19678	2681	657.36963	1841		

231017_OKGT_Sau_60kDa.raw #2928 RT: 9.7484 min: NESLIK-GlcNAc-1Da

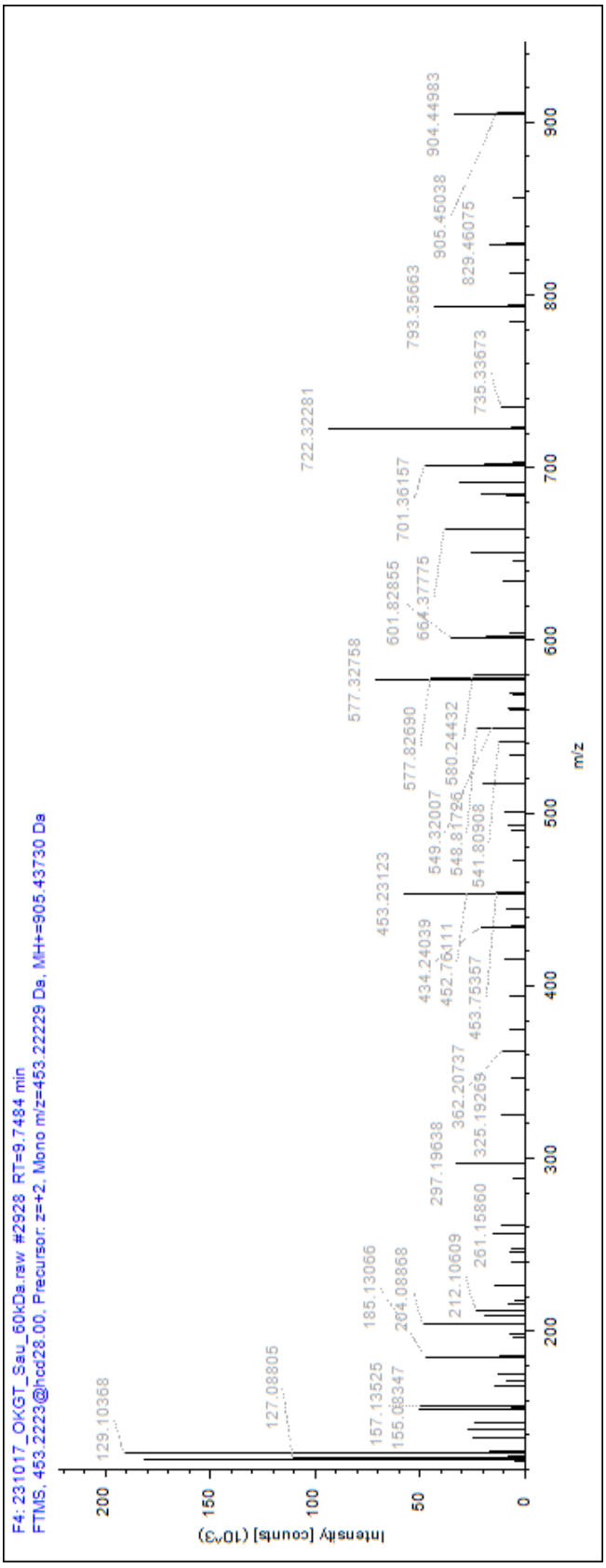


Figure 7.8: Sheep 60kDa #2928

Table 7.8: Sheep 60kDa #2928

m/z	Intensity [counts]	*m/z*	Intensity [counts]	*m/z*	Intensity [counts]
125.07330	5439	256.16898	15567	577.32758	71158
126.05637	182071	261.15860	11613	577.82690	45228
127.06015	9784	288.20477	5895	578.32983	14397
127.08805	110845	297.19638	32649	580.24432	24891
128.09161	8421	325.19269	11465	601.82855	35726
129.10368	191200	346.16818	6397	602.33228	18286
130.08832	16858	362.20737	10847	634.29022	10307
130.10744	7970	374.21292	7615	646.35474	5745
138.05600	24878	394.22699	7073	651.27985	26113
143.08249	5437	415.23483	9333	664.37775	38616
143.11922	27213	433.90918	13222	683.37195	8768
147.11444	24191	434.24039	20932	684.34021	21159
155.08347	50723	434.57990	6283	691.32141	31204
156.08701	6548	444.16083	8681	701.36157	48116
157.13525	49778	452.76111	27951	702.37207	19465
168.06796	14884	453.23123	57585	703.38824	5670
171.11453	9118	453.75357	13681	722.32281	94037
175.12198	12766	454.21387	8201	723.32684	6521
185.13066	47453	472.27292	6014	735.33673	11753
186.07877	12027	490.26419	6976	784.41937	7362
196.43607	5494	493.21759	8277	793.35663	43347
198.12701	7158	500.28690	8519	794.35815	8200
204.08868	48857	500.80002	9450	812.45288	7168
209.09505	19446	516.77875	19916	829.46075	17064
212.10609	23684	517.27618	10491	830.45581	9520
216.13818	7941	533.23633	7448	856.33789	5706
217.26770	5309	541.80908	12586	904.44983	33943
226.02315	6627	548.81726	23060	905.45038	13504
226.12170	14305	549.32007	16116		
240.17435	6852	559.81226	6967		
246.16077	7150	560.30658	8571		
248.11584	6575	568.81744	5619		
248.16325	6713	569.30975	7111		

230427_OKGTollersrud_OKGT_sau_S1.raw #46086 RT: 39.1235 min: ASLYPVYELNGTVYSQEWLV GlcNAc2FucMan3

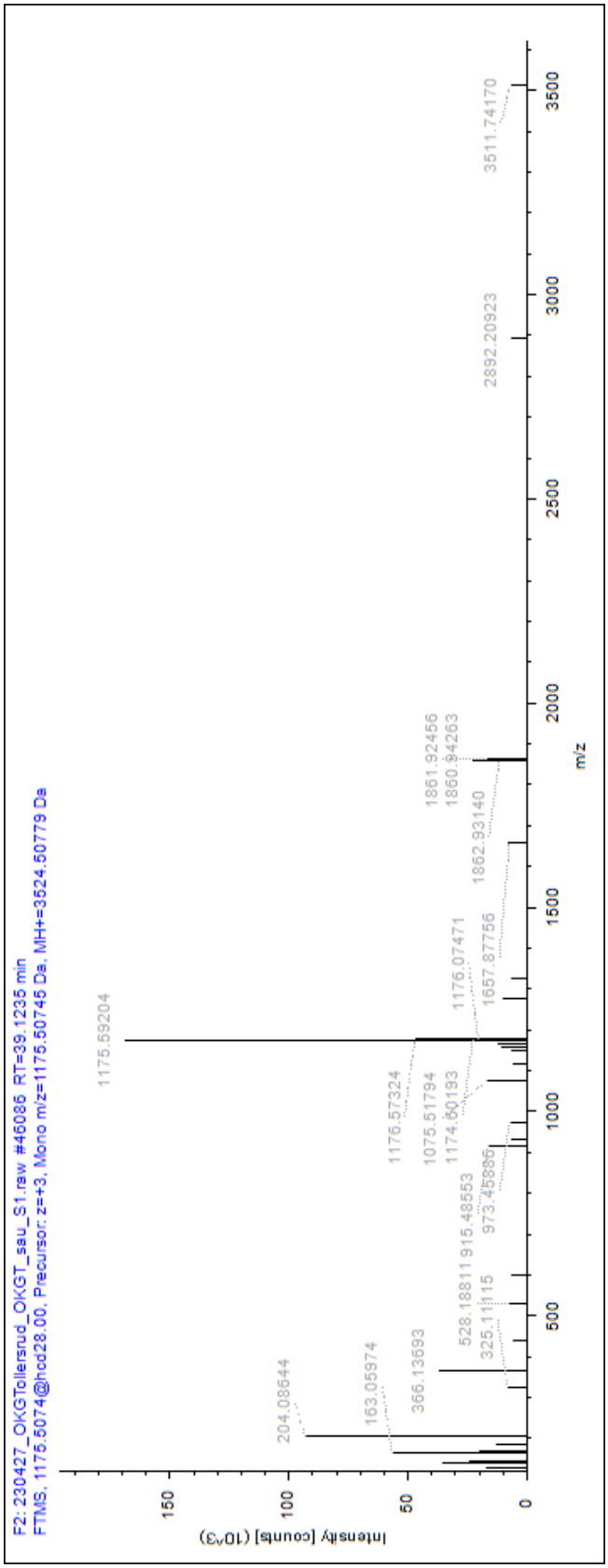


Figure 7.9: Sheep S1 60kDa #46086

Table 7.9: Sheep S1 60kDa #46086

m/z	Intensity [counts]	*m/z*	Intensity [counts]
126.05476	16981	1116.19666	5839
127.03922	15907	1151.53711	6882
138.05429	35803	1157.54700	10492
145.04950	24396	1166.55457	11882
163.05974	56624	1167.53687	7972
164.06361	6777	1174.60193	22824
168.06531	19749	1175.06616	11485
186.07658	13227	1175.59204	168822
204.08644	93185	1176.07471	20403
325.11115	8181	1176.57324	47128
366.13693	36892	1275.63147	9760
441.10608	5866	1327.15894	6255
528.18811	7488	1657.87756	7383
599.85846	6145	1658.83179	8023
915.48553	16402	1860.94263	22890
916.48242	8552	1861.92456	16657
930.99408	6511	1862.93140	11674
973.45886	7318	2892.20923	6725
1075.51794	16819	3511.74170	6931
1076.52527	12582		

230427_OKGTollersrud_OKGT_sau_S1.raw #6040 RT: 11.1966 min: TKNESLIK-GlcNAc wo/ H_2O

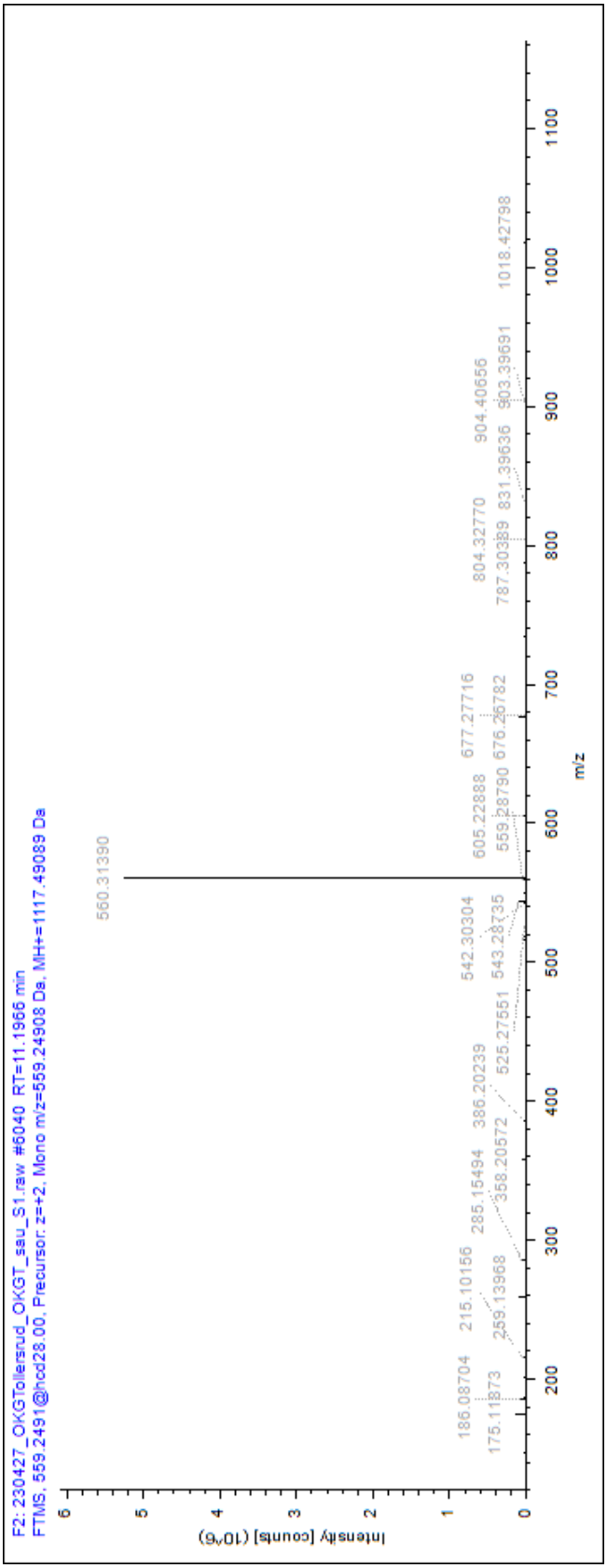


Figure 7.10: Sheep S1 60kDa #6040

Table 7.10: Sheep S1 60kDa #6040

m/z	Intensity [counts]	*m/z*	Intensity [counts]
129.10225	13707	471.24164	6560
130.08563	6611	491.18839	10320
147.11258	19724	494.75067	6042
158.09193	11997	508.25107	11621
169.06001	14498	512.65076	5862
169.09656	6756	516.28967	24260
173.09093	7236	518.29218	19802
173.13026	6257	525.27551	23897
175.11873	134770	526.26129	21962
186.08704	74269	542.30304	31056
187.10780	24976	543.28735	96054
201.12270	18687	549.75726	18304
204.13474	15771	550.26251	10293
215.10156	45668	558.77307	19877
228.13480	16243	559.28790	41916
243.13158	7729	560.31390	5253709
259.13968	85587	605.22888	36040
276.16641	13756	676.26782	84648
285.15494	52635	677.27716	20250
314.17477	7023	734.36536	11516
331.19189	7650	787.30389	29211
343.16101	9017	788.30206	13348
358.20572	56467	804.32770	28944
368.18967	6701	805.33130	10016
375.22882	7285	831.39636	17890
386.20239	32747	903.39691	16485
402.67258	8788	904.40656	12866
416.20523	10516	1018.42798	21066
471.24164	6560		

230427_OKGTollersrud_OKGT_sau_S1.raw #17262 RT: 19.2224 min: NLLDAVILNSTR-GlcNAc2FucMan-H_2O

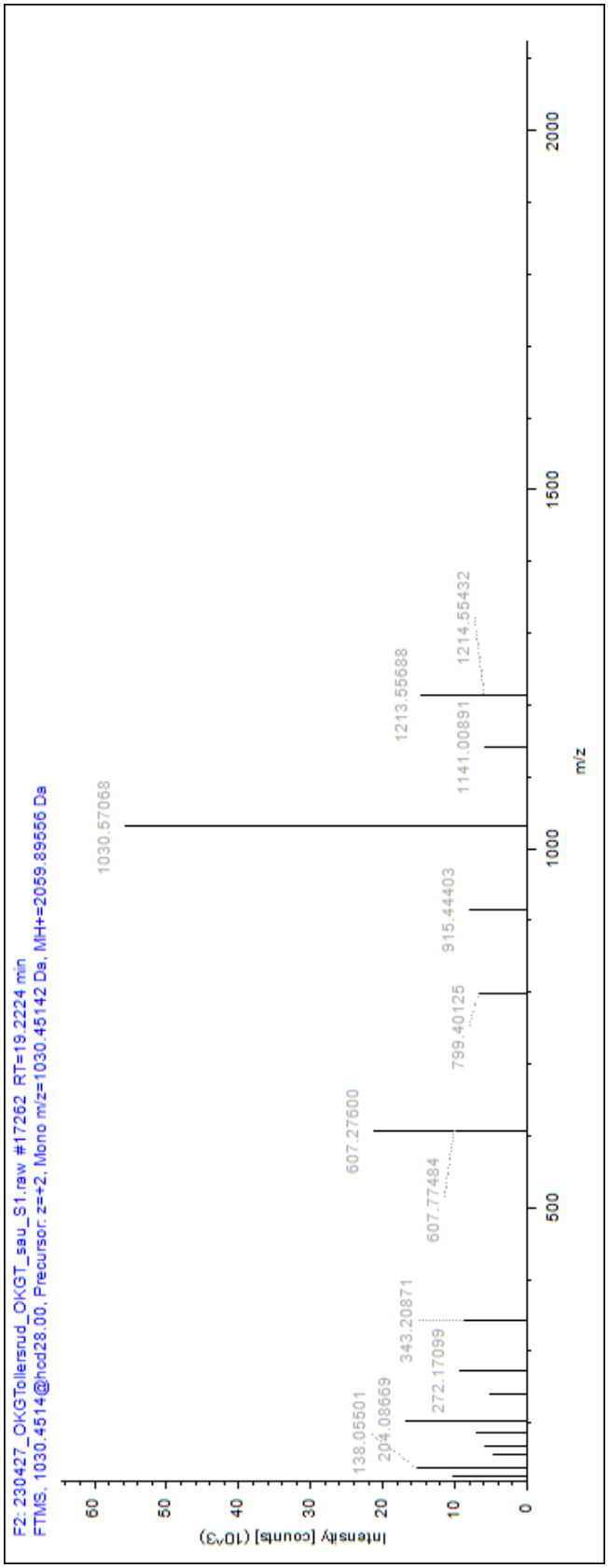


Figure 7.11: Sheep S1 60kDa #17262

Table 7.11: Sheep S1 60kDa #17262

m/z	Intensity [counts]
126.05509	10280
126.34061	5780
127.08691	8728
138.05501	15443
155.97653	4602
168.06525	5920
186.07423	6959
204.08669	16814
240.78970	5098
272.17099	9373
343.20871	8920
607.27600	21243
607.77484	10107
799.40125	6679
915.44403	8099
1030.57068	55787
1141.00891	5959
1213.55688	14622
1214.55432	5911

230427_OKGTollersrud_OKGT_sau_S1.raw #4623 RT: 10.0450 min: EEELANQR

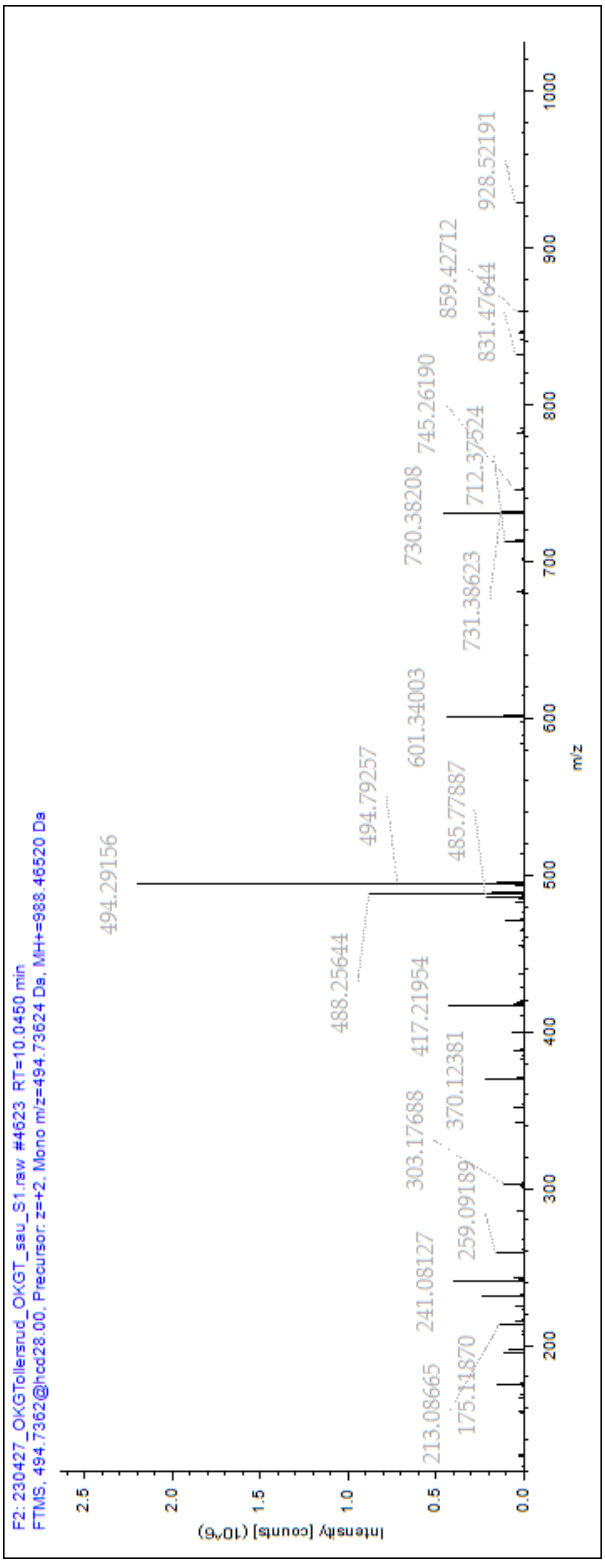


Figure 7.12: Sheep S1 60kDa #4623

230427_OKGTollersrud_OKGT_sau_S1.raw #9993 RT: 14.0520 min: SIQTAMGLR

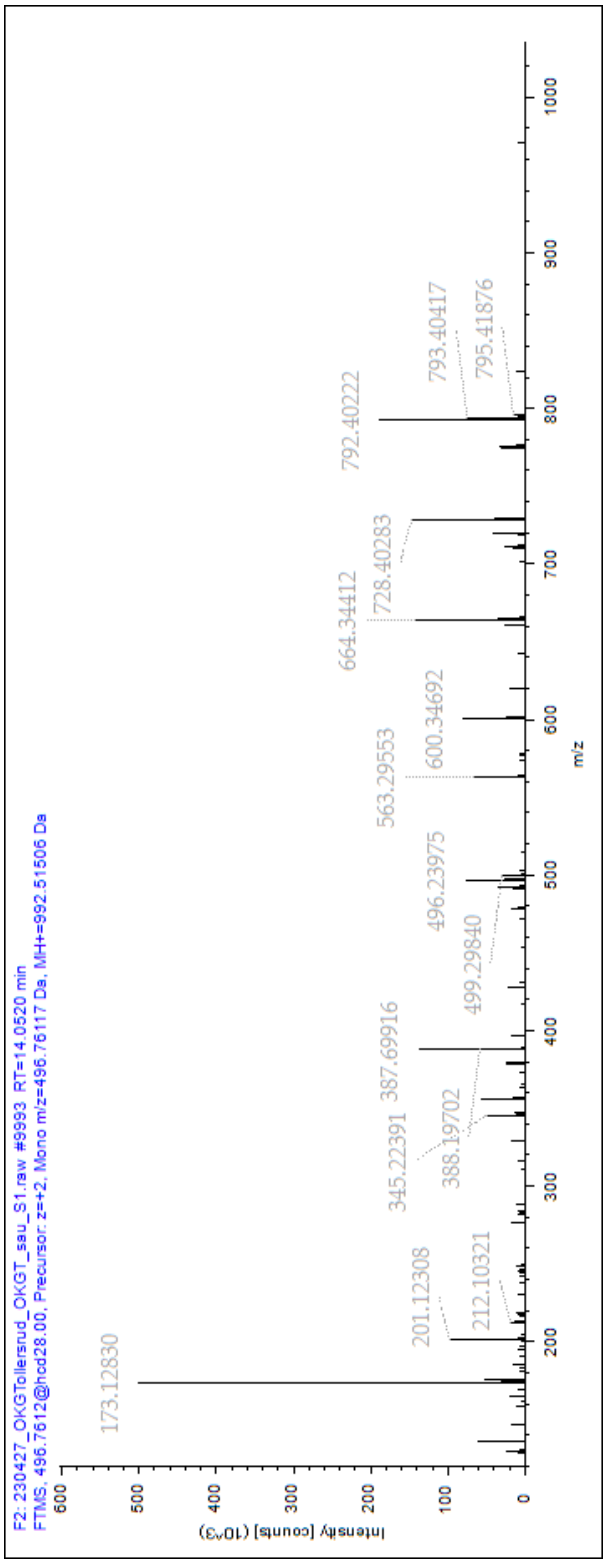


Figure 7.13: Sheep S1 60kDa #9993

230427_OKGTollersrud_OKGT_sau_S1.raw #10117 RT: 14.1389 min: SAMETWEER

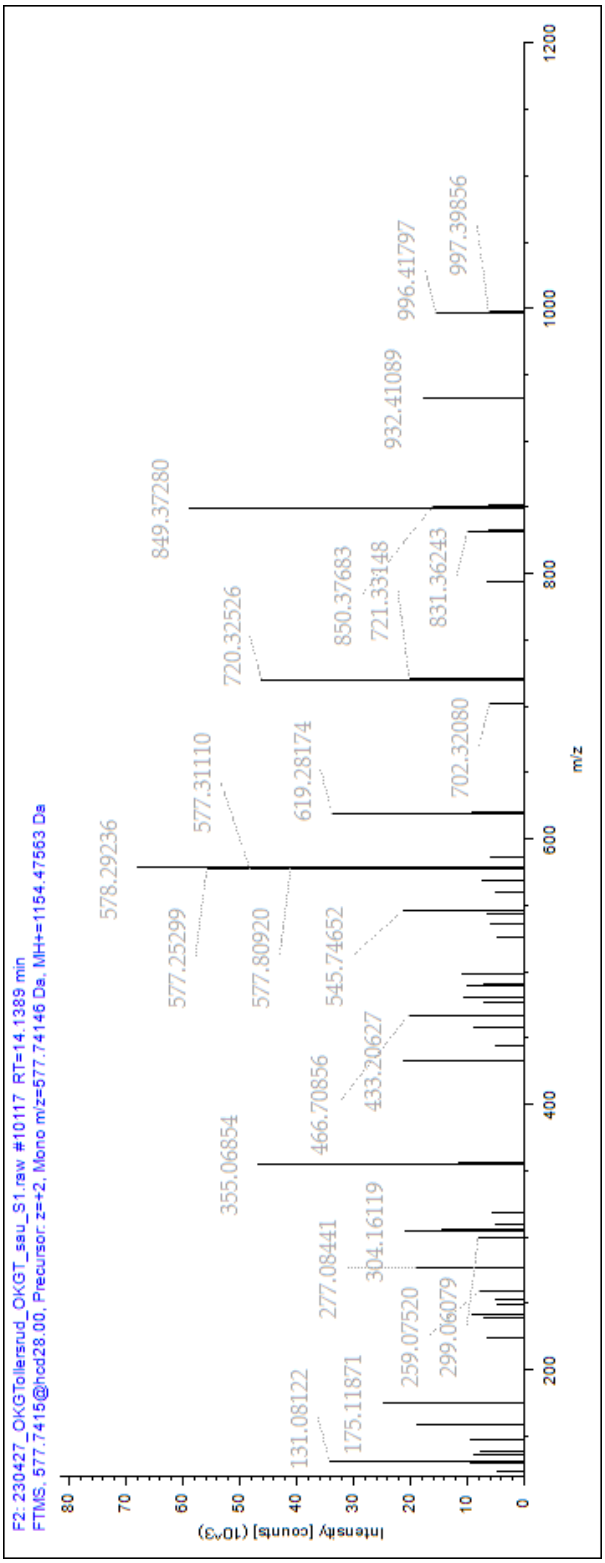


Figure 7.14: Sheep S1 60kDa #10117

230427_OKGTollersrud_OKGT_sau_S1.raw #12894 RT: 16.0995 min: YSTLSETMK

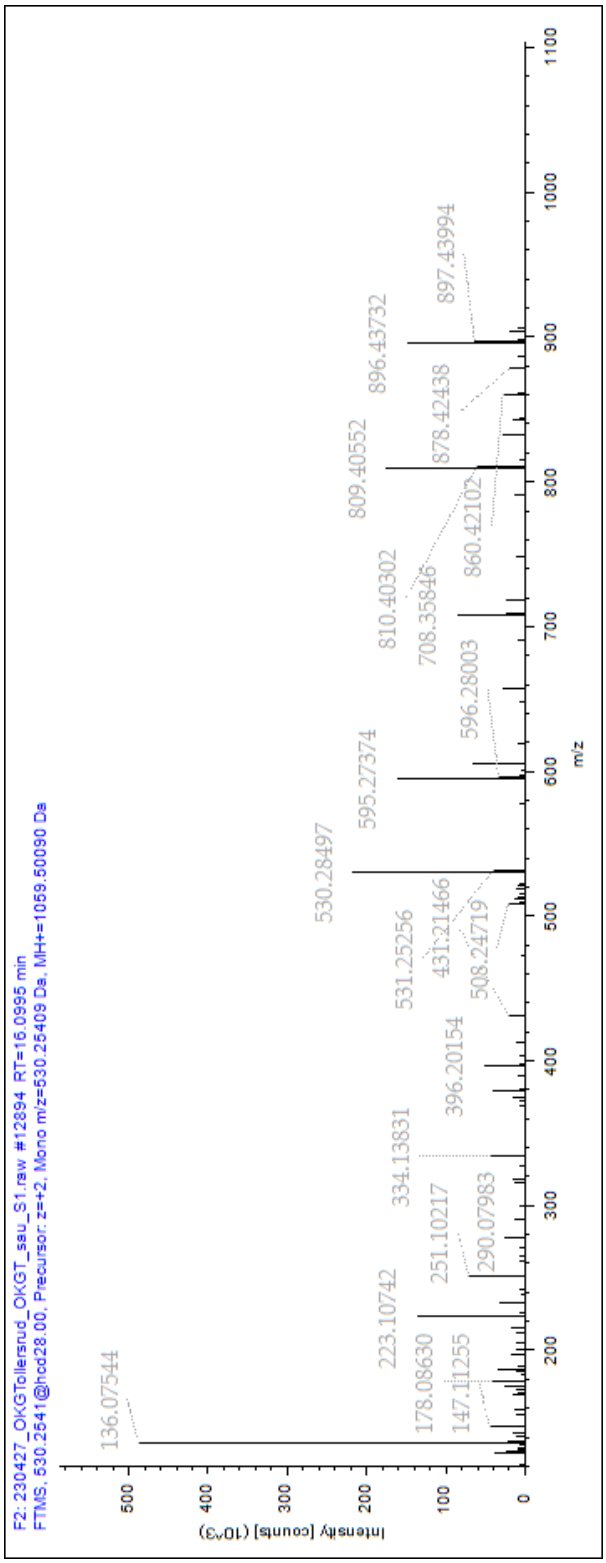


Figure 7.15: Sheep S1 60kDa #12894

230427_OKGTollersrud_OKGT_sau_S1.raw #13336 RT: 16.3991 min: TYEEVAHNFAK

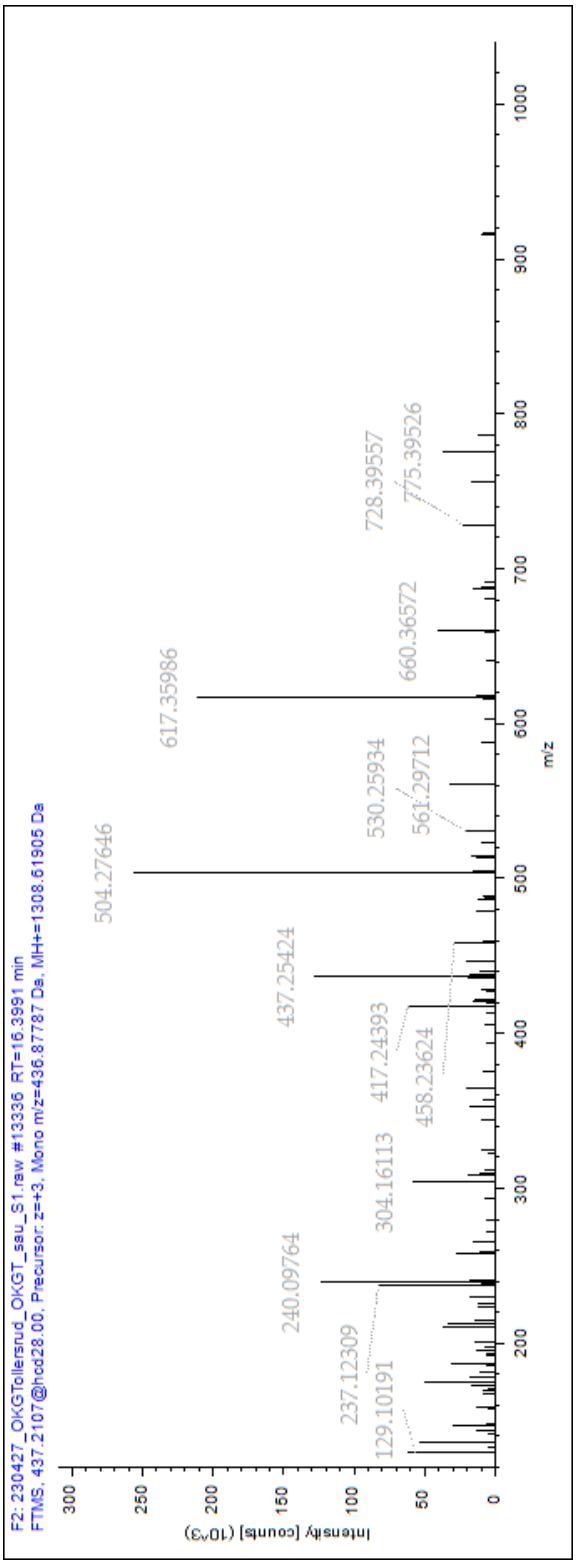


Figure 7.16: Sheep S1 60kDa #13336

230427_OKGTollersrud_OKGT_sau_S1.raw #13604 RT: 16.5932 min: TYEEVAHNFAK

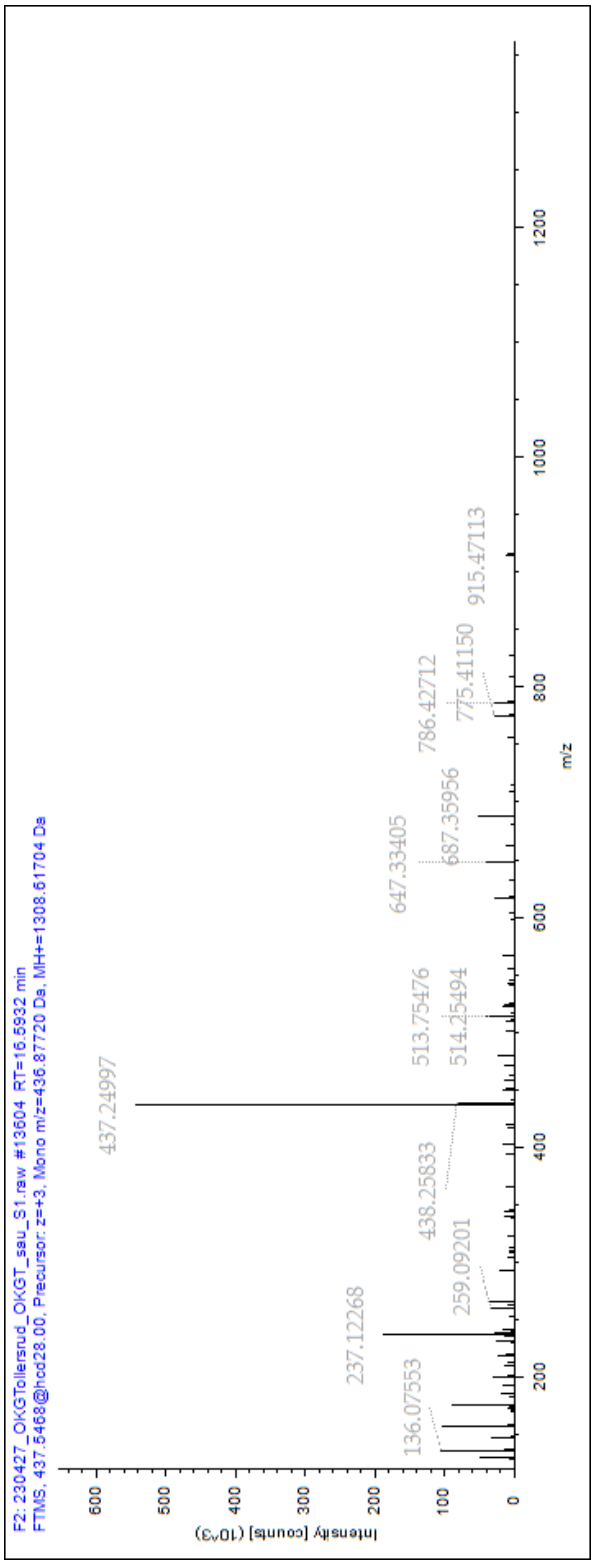


Figure 7.17: Sheep S1 60kDa #13604

230427_OKGTollersrud_OKGT_sau_S1.raw #14128 RT: 17.0518 min: EDTGHTLY-DYR

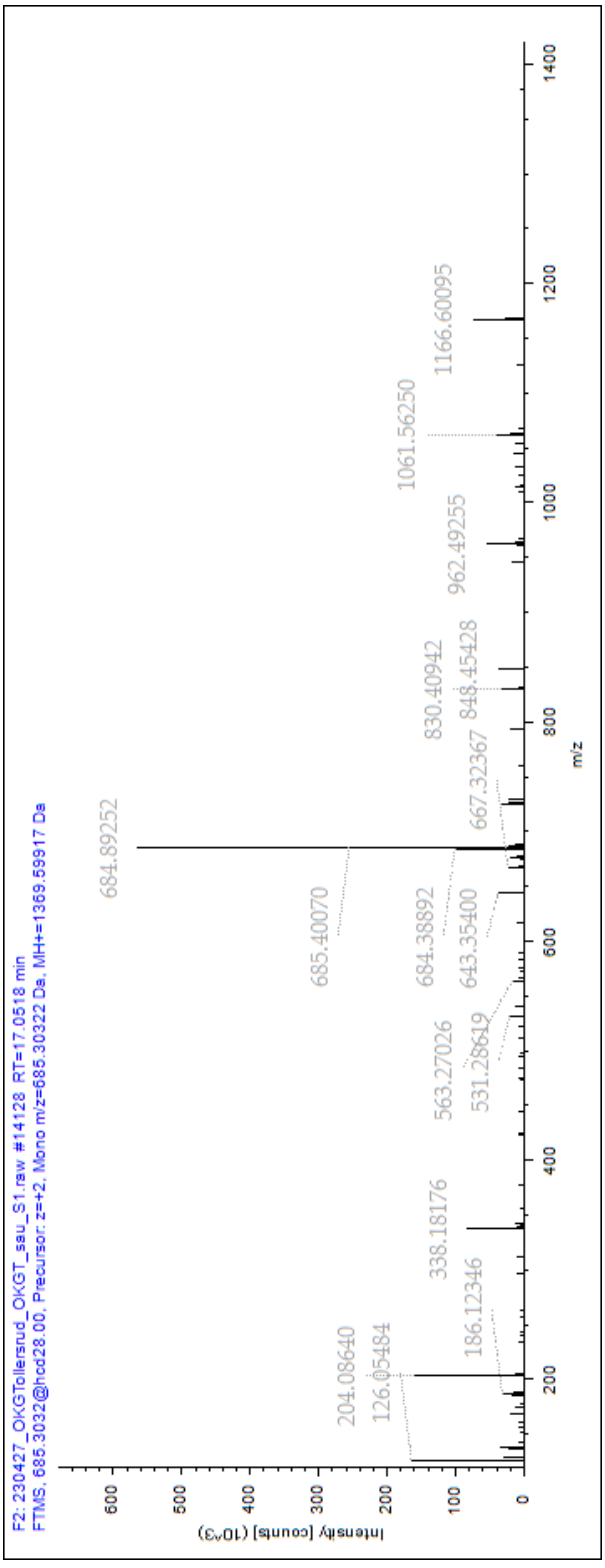


Figure 7.18: Sheep S1 60kDa #14128

230427_OKGTollersrud_OKGT_sau_S1.raw #15638 RT: 18.0929 min: FVAGLAR

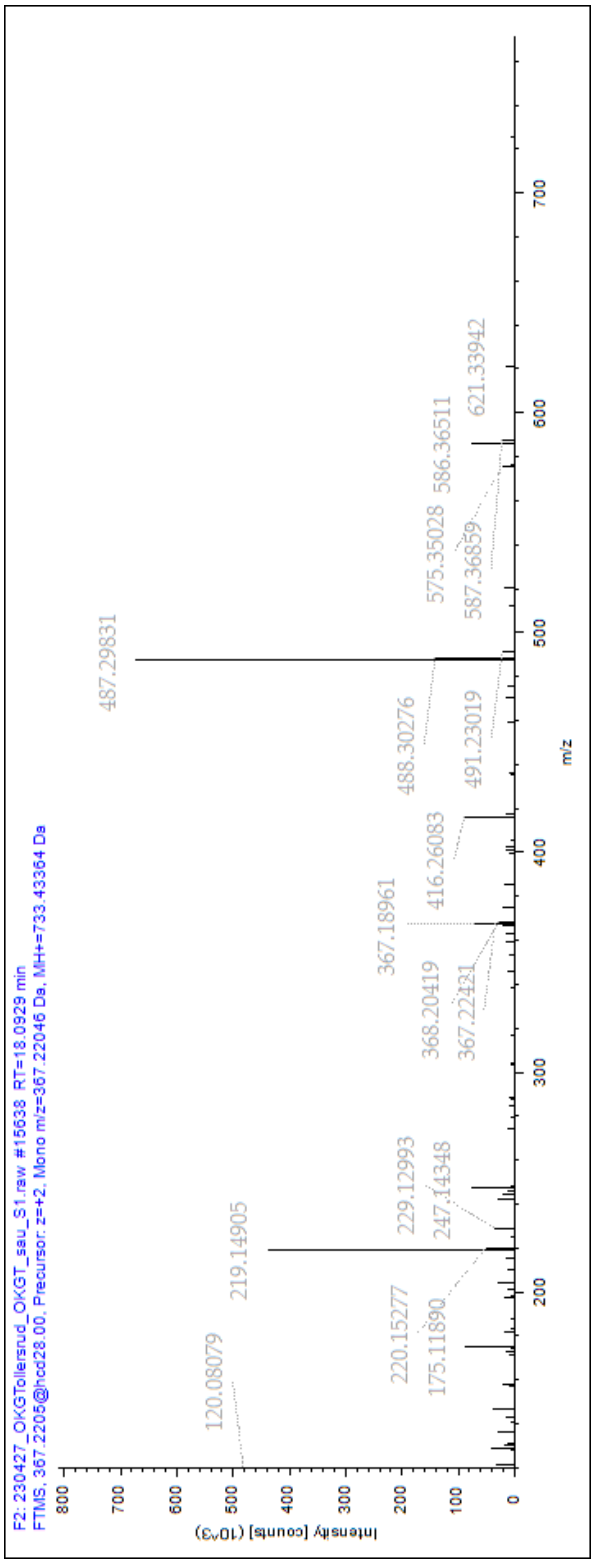


Figure 7.19: Sheep S1 60kDa #15638

230427_OKGTollersrud_OKGT_sau_S1.raw #16825 RT: 18.9186 min: SAMETWEER

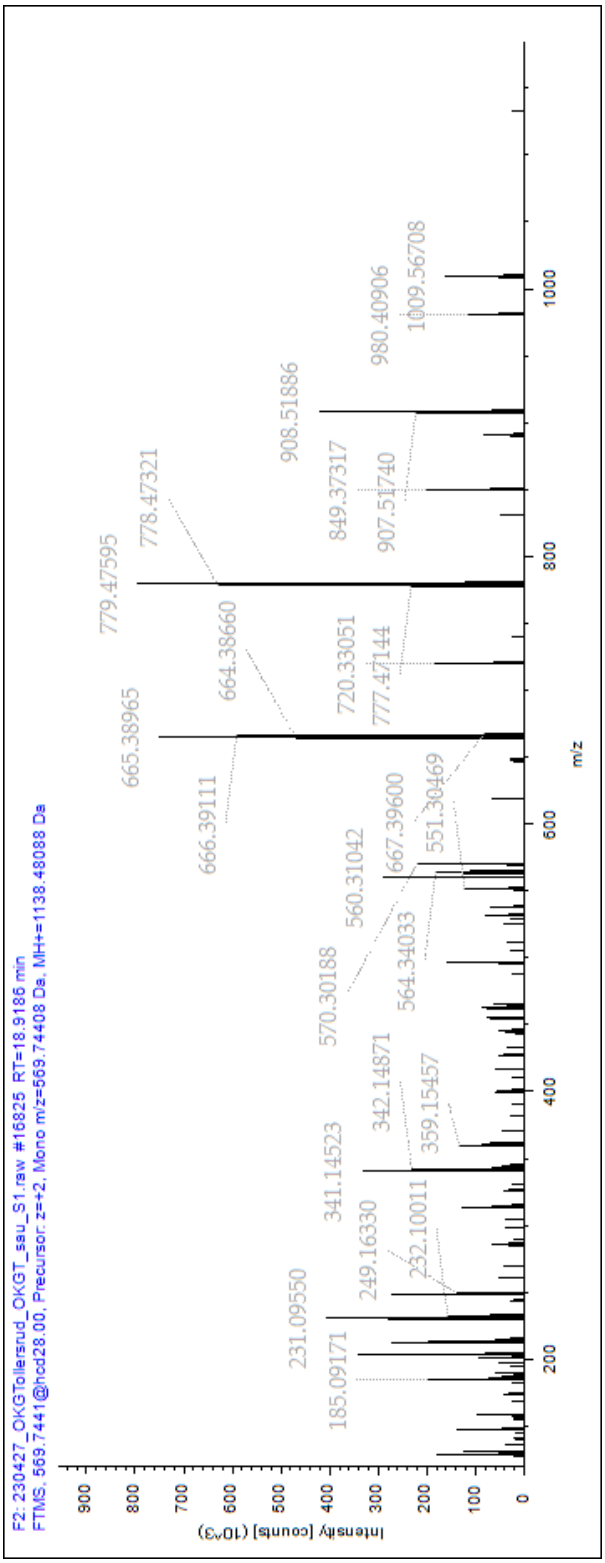


Figure 7.20: Sheep S1 60kDa #16825

230427_OKGTollersrud_OKGT_sau_S1.raw #19005 RT: 20.4971 min: VGGQLVLTR

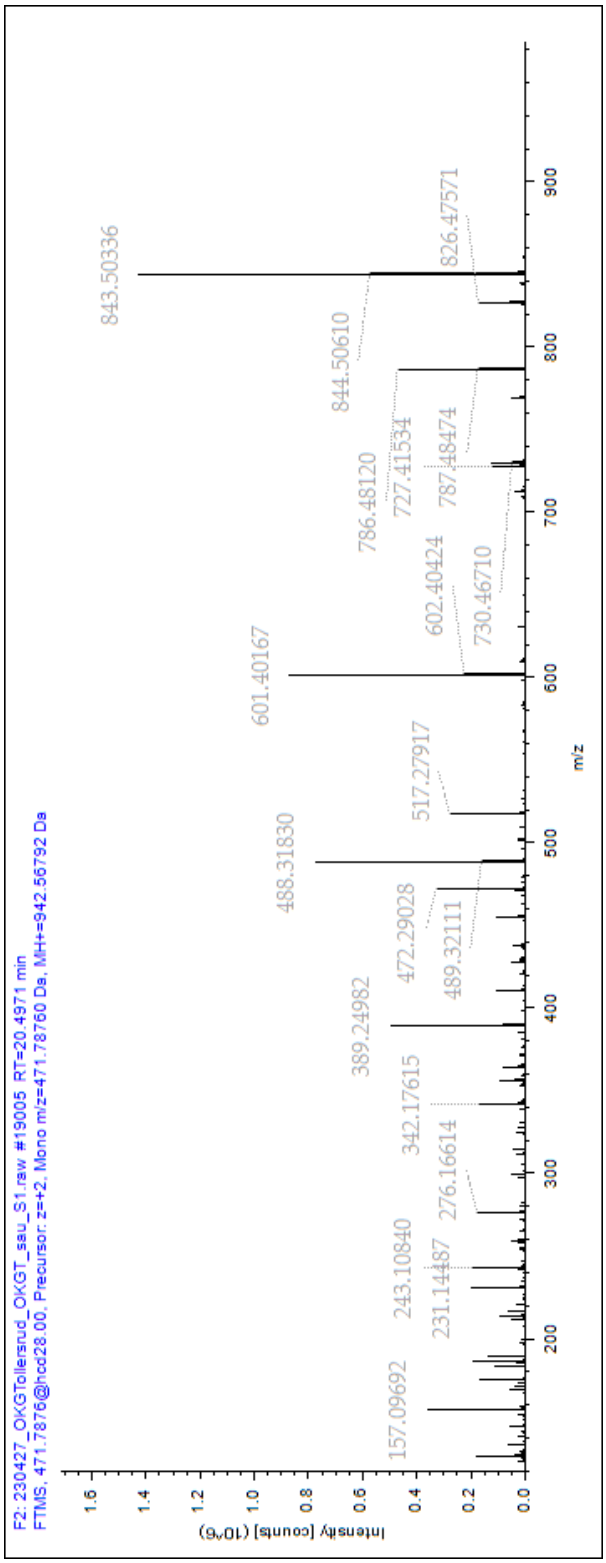


Figure 7.21: Sheep S1 60kDa #19005

230427_OKGTollersrud_OKGT_sau_S1.raw #19066 RT: 20.5391 min: QLAINSIR

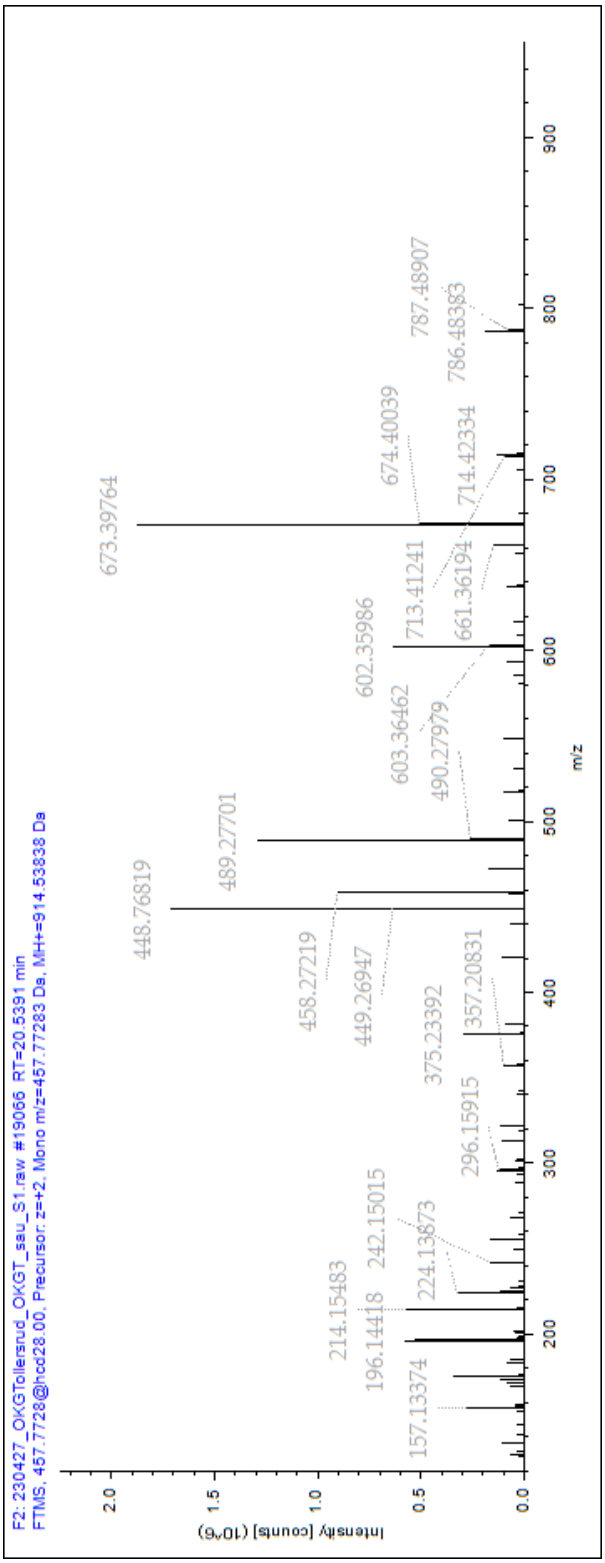


Figure 7.22: Sheep S1 60kDa #19066

230427_OKGTollersrud_OKGT_sau_S1.raw #20425 RT: 21.3553 min: SIQTAMGLR

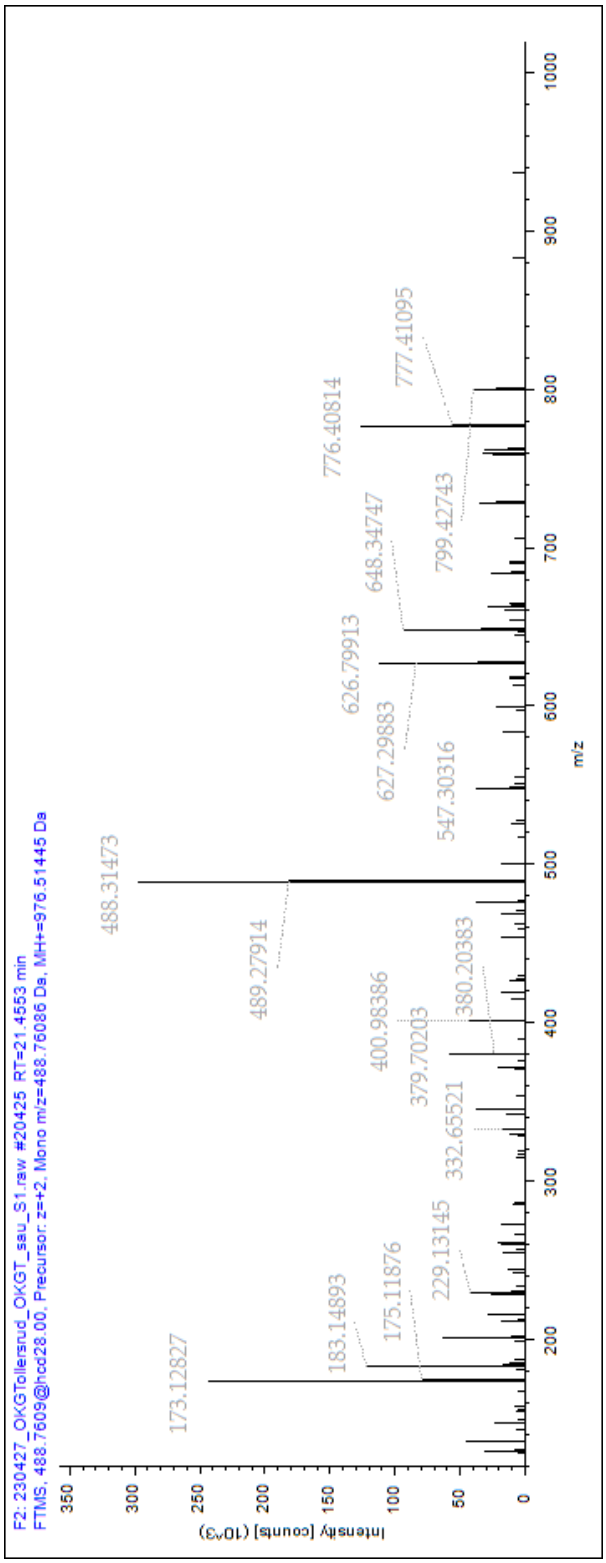


Figure 7.23: Sheep S1 60kDa #20425

230427_OKGTollersrud_OKGT_sau_S1.raw #20526 RT: 21.5262 min: SIQTAMGLR

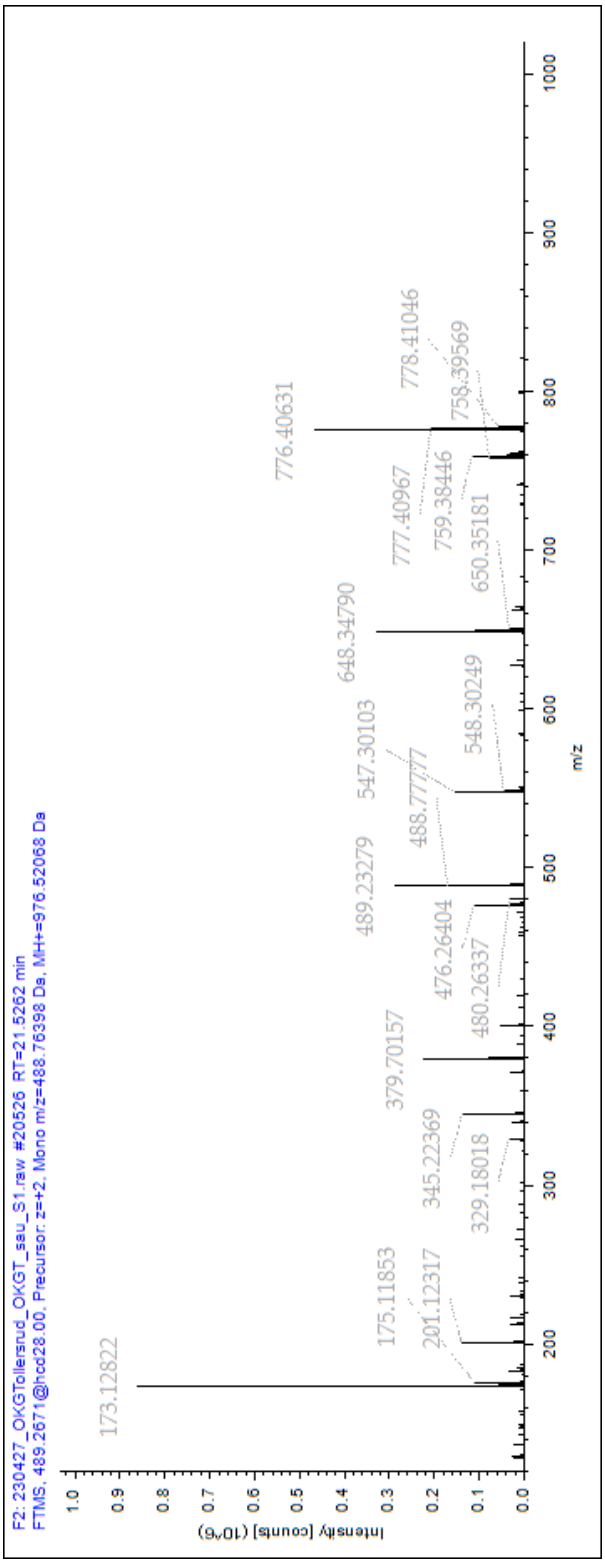


Figure 7.24: Sheep S1 60kDa #20526

230427_OKGTollersrud_OKGT_sau_S1.raw #27671 RT: 26.4467 min: TGSILPSMHF-FQAK

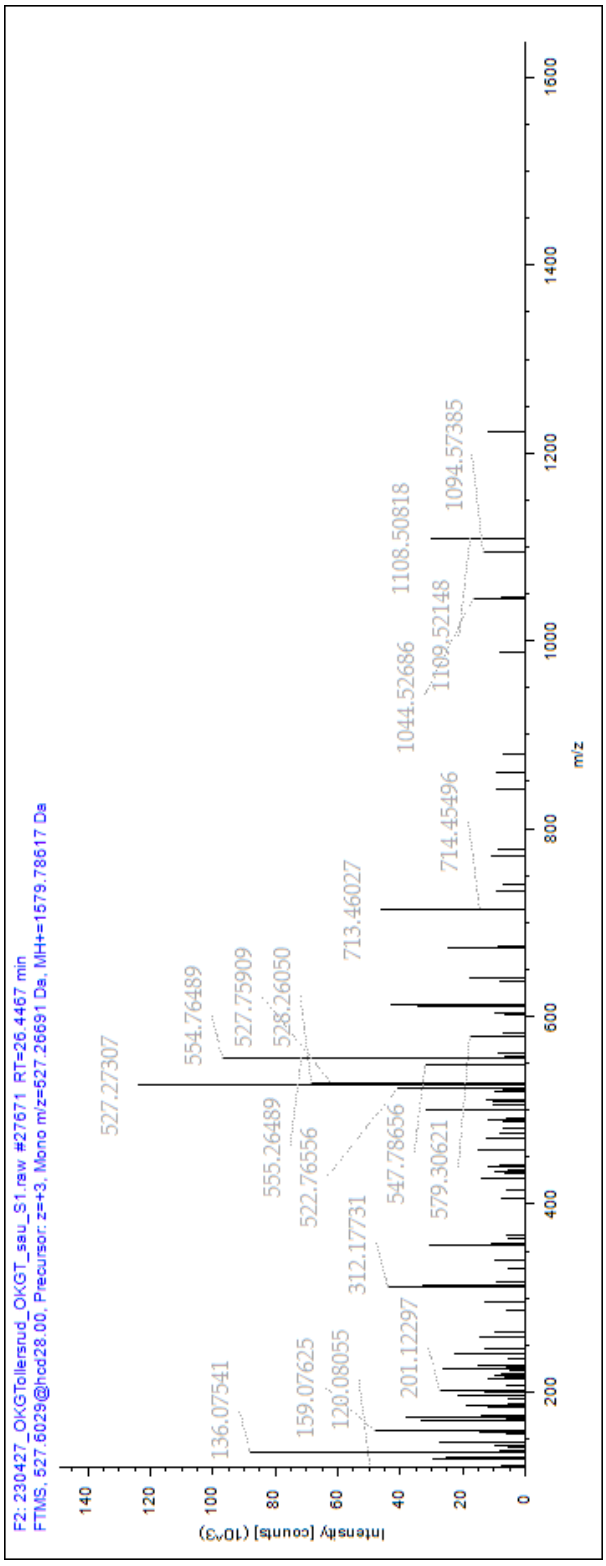


Figure 7.25: Sheep S1 60kDa #27671

230427_OKGTollersrud_OKGT_sau_S1.raw #36780 RT: 32.6970 min: TGSILPSMHF-FQAK

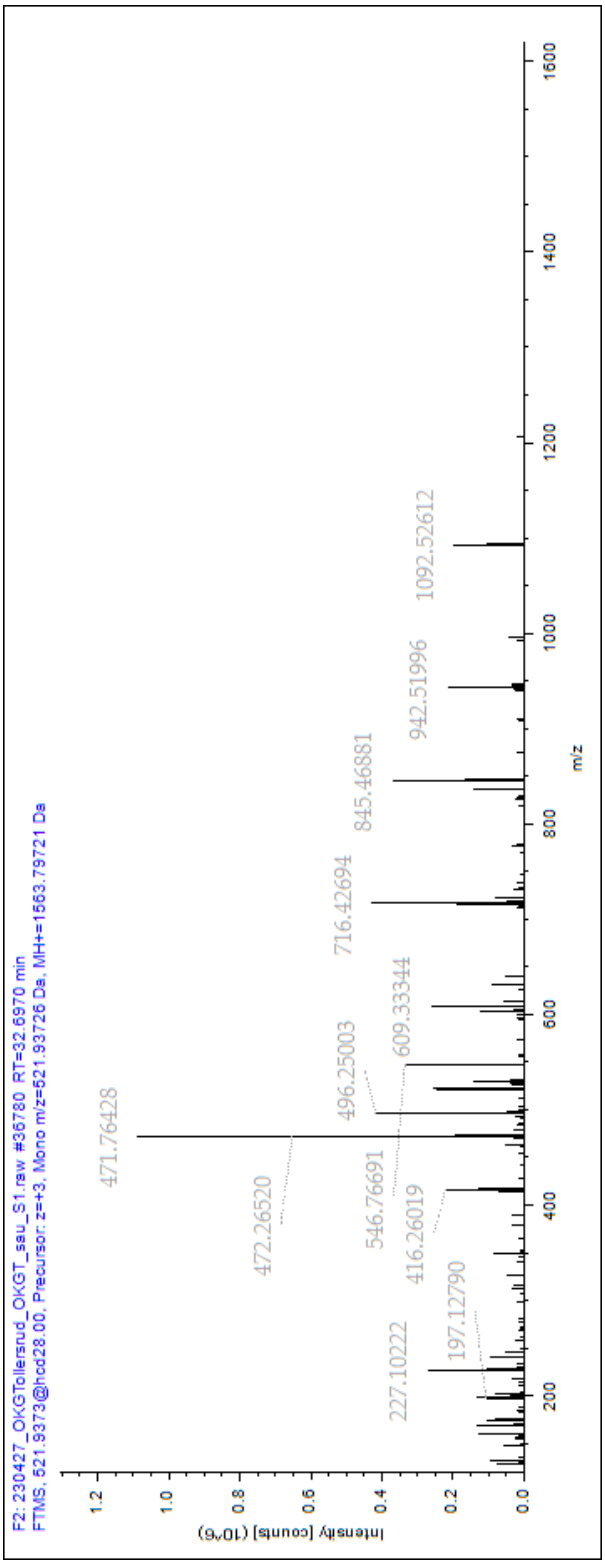


Figure 7.26: Sheep S1 60kDa #36780

230427_OKGTollersrud_OKGT_sau_S1.raw #42916 RT: 36.8598 min: TIFFVLR

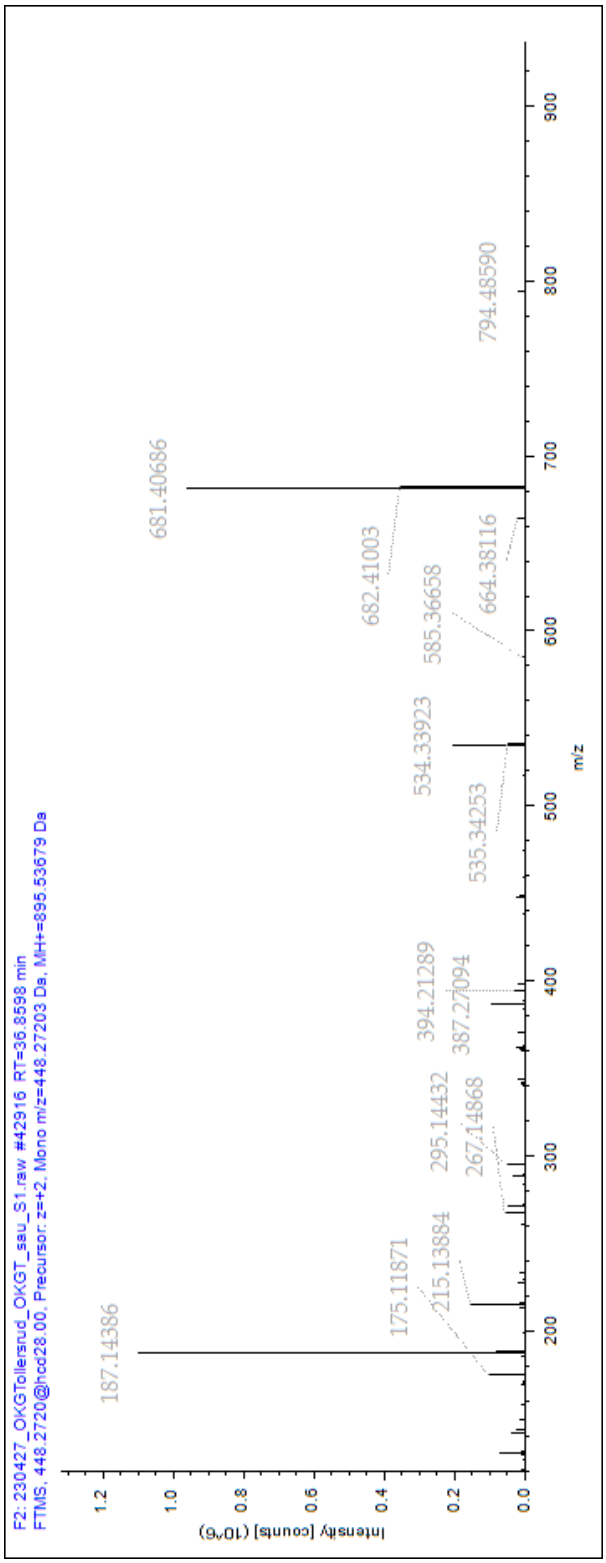


Figure 7.27: Sheep S1 60kDa #42916

230427_OKGTollersrud_OKGT_sau_S1.raw #47017 RT: 39.7553 min: VKFPGIVAGFDLVGR

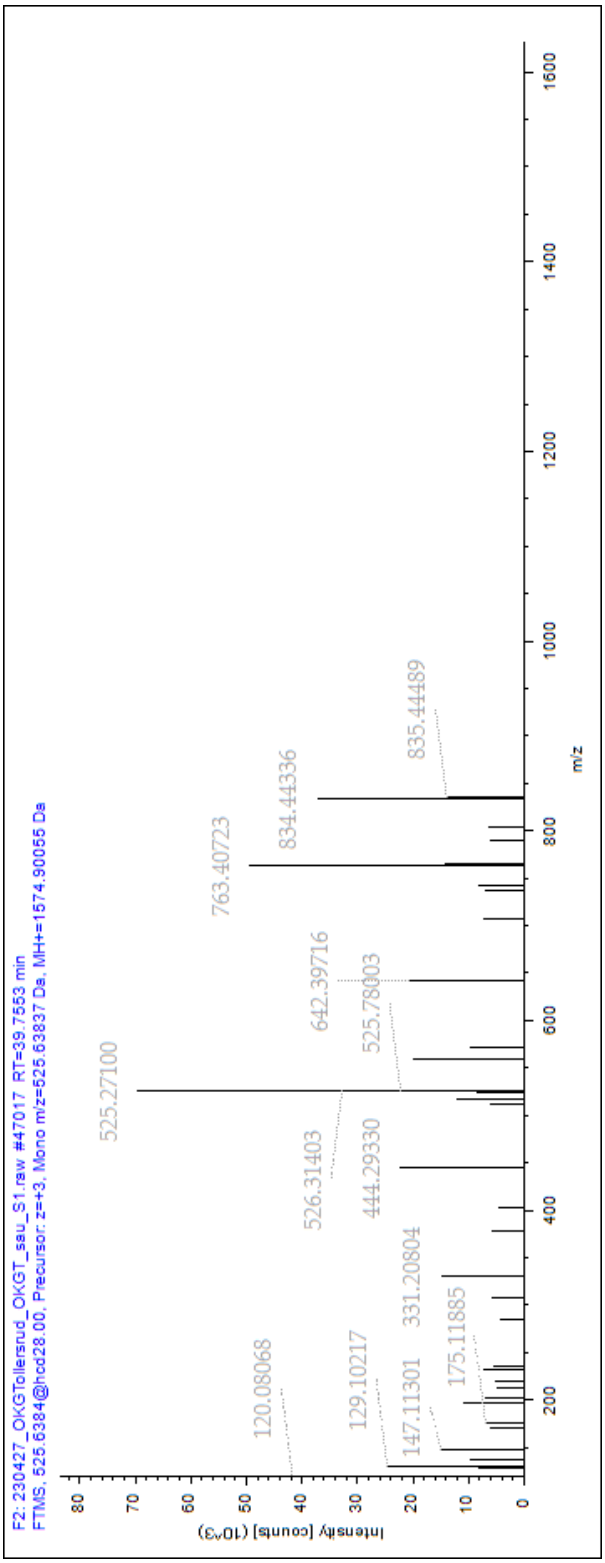


Figure 7.28: Sheep S1 60kDa #47017

230427_OKGTollersrud_OKGT_sau_S1.raw #48089 RT: 40.4934 min: DIPIEVCPIS-NQVLK

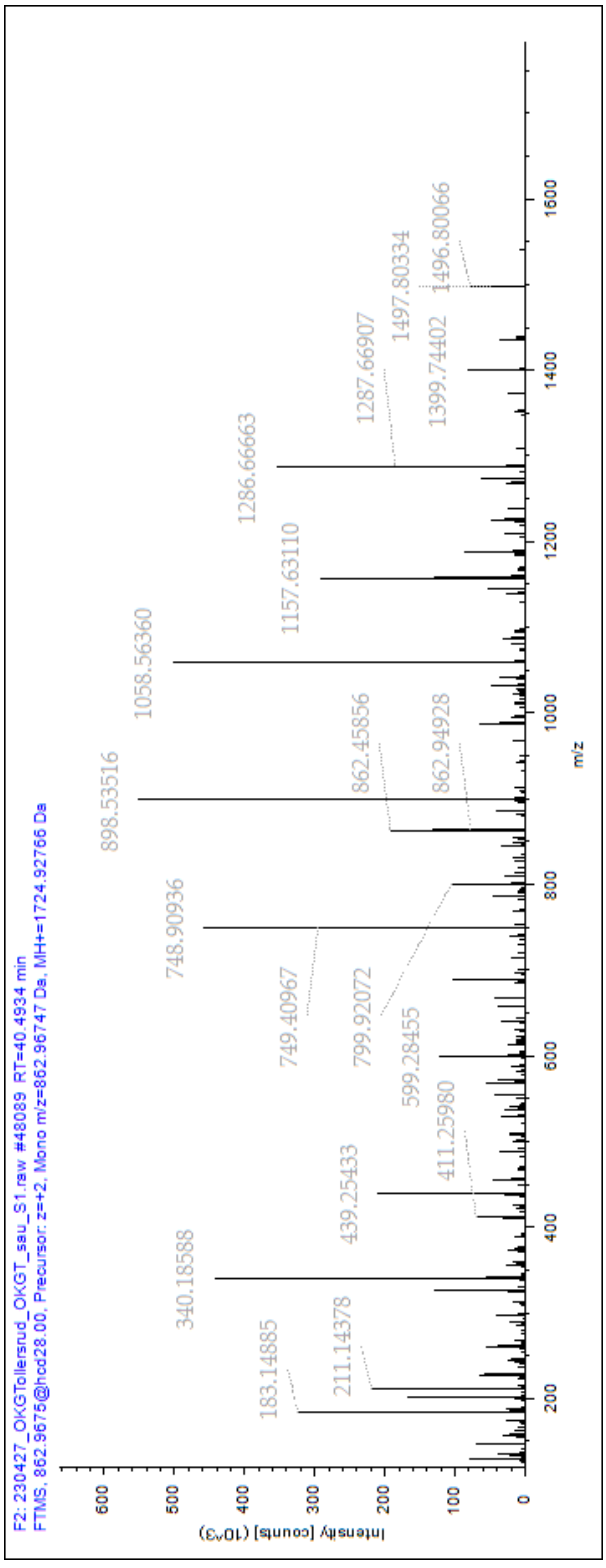


Figure 7.29: Sheep S1 60kDa #48089

230427_OKGTollersrud_OKGT_sau_S1.raw #53349 RT: 44.2358 min: FPGIVAGFDLVGR

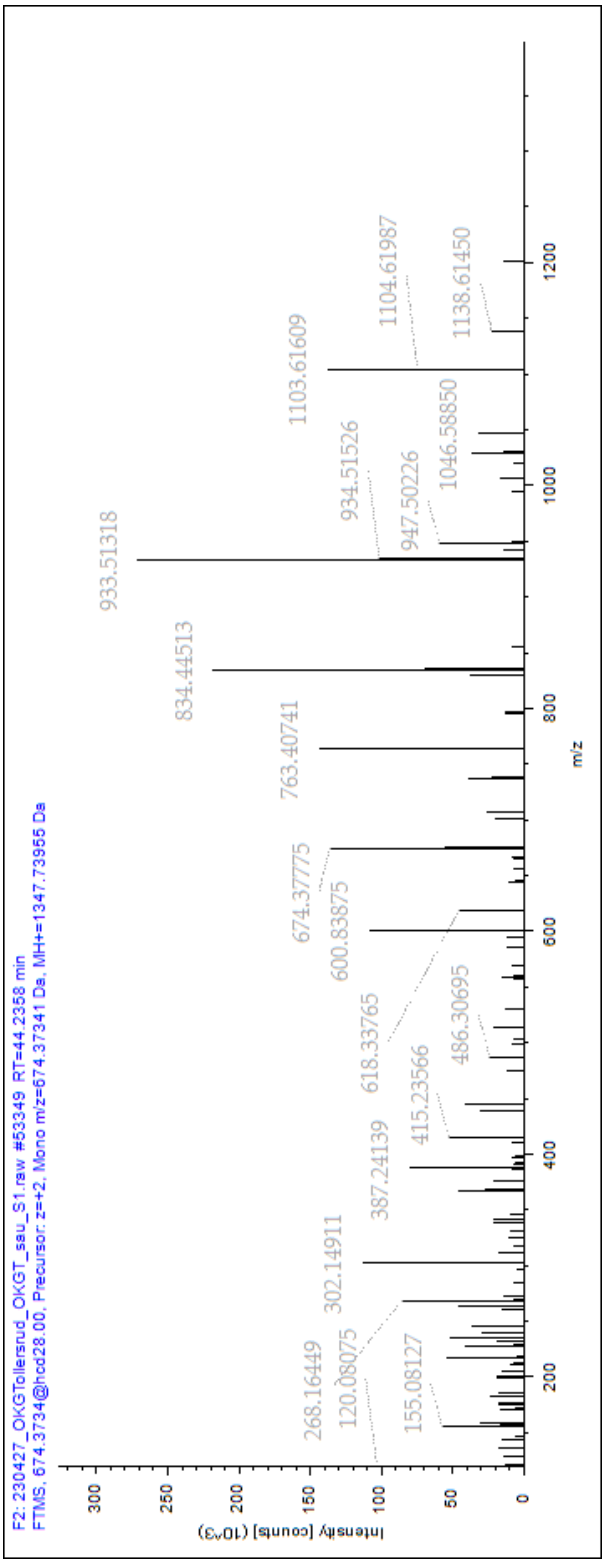


Figure 7.30: Sheep S1 60kDa #53349

230427_OKGTollersrud_OKGT_sau_S1.raw #53397 RT: 44.2732 min: FPGIVAGFDLVGR

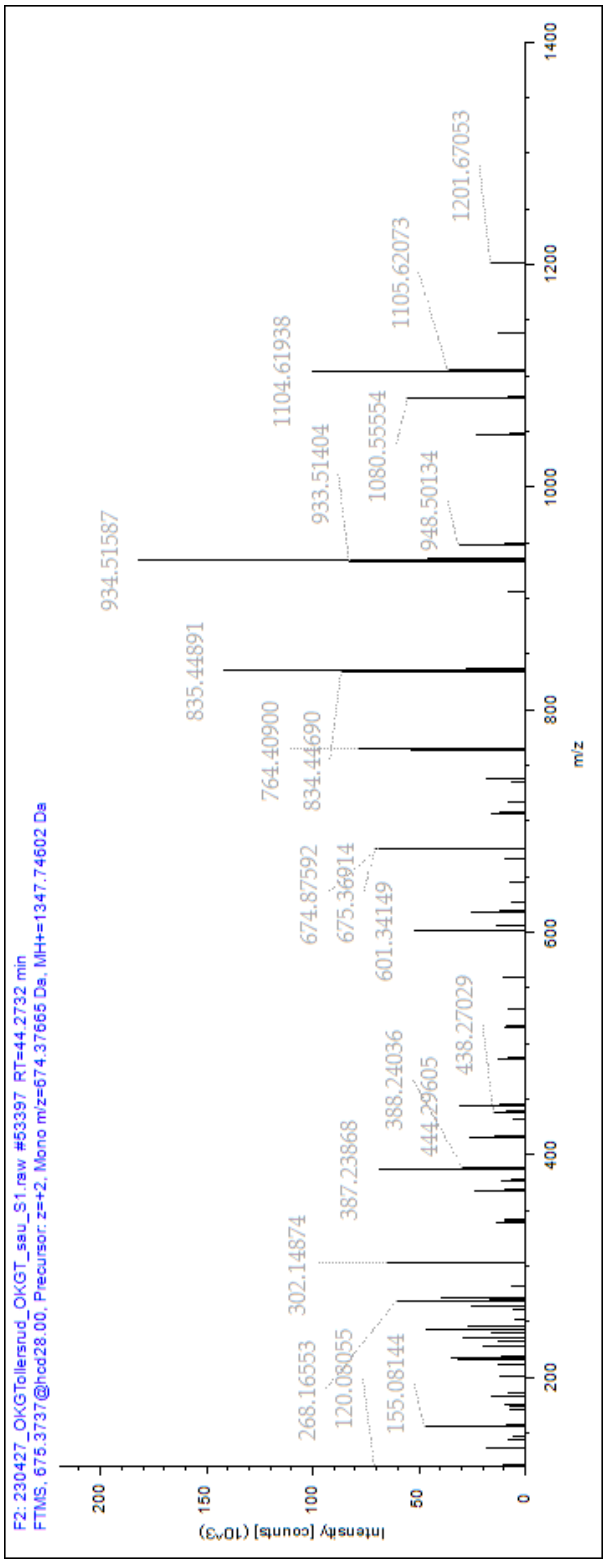


Figure 7.31: Sheep S1 60kDa #53397

Bibliography

[1] The DADA2 Foundation, "What is DADA2?" [Online]. Available: https://dada2.org/what-is-dada2/ 1.1, 1.4

[2] O. K. Greiner-Tollersrud, V. Boehler, M. Krausz, A. Polyzou, M. Seidl, A. Goschin, C. R. Smulski, E. Trompouki, R. Link, H. Ebersbach, H. Srinivas, M. Marchant, D. Staab, C. Vågbø, D. Guerini, S. Raieli, S. Baasch, P. Henneke, R. Geiger, X. Peng, B. Grimbacher, E. Bartok, I. Alseth, M. Warncke, and M. Proietti, "ADA2 is a Lysosomal Adenosine Deaminase Acting on DNA (LADAD) controlling DNA sensing," *Under review in Cell Report*, 8 2023. 1.1, 1.3.2, 4.3

[3] The DADA2 Foundation, "About Us | The DADA2 Foundation." [Online]. Available: https://dada2.org/about-us/ 1.1

[4] V. Kumar, A. K. Abbas, and J. C. Aster, "Inflammation and repair," in *Robbins Basic Pathology*, 10th ed., V. Kumar, A. K. Abbas, J. C. Aster, and J. A. Perkins, Eds. Philadelphia: Elsevier, 2018, ch. 3, pp. 57–96. 1.2.1, 1.3.2

[5] Z. L. Chang, "Important aspects of Toll-like receptors, ligands and their signaling pathways," *Inflammation Research*, vol. 59, no. 10, pp. 791–808, 10 2010. 1.2.1, 1.2.1

[6] A. S. Sameer and S. Nissar, "Toll-Like Receptors (TLRs): Structure, Functions, Signaling, and Role of Their Polymorphisms in Colorectal Cancer Susceptibility," *BioMed Research International*, vol. 2021, pp. 1–14, 9 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8452412/ 1.2.1, 1

[7] S. Uematsu and S. Akira, "Toll-like Receptors and Type I Interferons," *Journal of Biological Chemistry*, vol. 282, no. 21, pp. 15 319–15 323, 5 2007. 1.2.1

[8] Q. Chen, L. Sun, and Z. J. Chen, "Regulation and function of the cGAS–STING pathway of cytosolic DNA sensing," *Nature Immunology*,

vol. 17, no. 10, pp. 1142–1149, 10 2016. [Online]. Available: https://www.nature.com/articles/ni.3558 1.2.1, 1.2.1, 2

[9] J. Wu, L. Sun, X. Chen, F. Du, H. Shi, C. Chen, and Z. J. Chen, "Cyclic GMP-AMP Is an Endogenous Second Messenger in Innate Immune Signaling by Cytosolic DNA," *Science*, vol. 339, no. 6121, pp. 826–830, 2 2013. 1.2.1

[10] S. Pandey and T. Kawai, "Chapter 5 - Host DNA Induced Inflammation and Autoimmune Diseases," in *Biological DNA Sensor*, K. J. Ishii and C. K. Tang, Eds. Amsterdam: Academic Press, 2014, pp. 103–132. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780124047327000058 1.2.1

[11] W. Zhou, D. Richmond-Buccola, Q. Wang, and P. J. Kranzusch, "Structural basis of human TREX1 DNA degradation and autoimmune disease," *Nature Communications*, vol. 13, no. 1, 7 2022. [Online]. Available: https://www.nature.com/articles/s41467-022-32055-z 1.2.1

[12] S. R. Simpson, W. O. Hemphill, T. Hudson, and F. W. Perrino, "TREX1 – Apex predator of cytosolic DNA metabolism," *DNA Repair*, vol. 94, p. 102894, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568786420301427 1.2.1

[13] K. Kawane, K. Motani, and S. Nagata, "DNA Degradation and Its Defects," *Cold Spring Harbor Perspectives in Biology*, vol. 6, 2014. 1.2.1, 3

[14] E. Cunill-Semanat and J. Salgado, "Spontaneous and Stress-Induced Pore Formation in Membranes: Theory, Experiments and Simulations," *The Journal of Membrane Biology*, vol. 252, no. 4-5, pp. 241–260, 10 2019. 1.2.1, 3, 4.3

[15] N. Niyonzima, S. S. Bakke, I. Gregersen, S. Holm, Sandanger, H. L. Orrem, B. Sporsheim, L. Ryan, X. Y. Kong, T. B. Dahl, M. Skjelland, K. K. Sørensen, A. M. Rokstad, A. Yndestad, E. Latz, L. Gullestad, G. Andersen, J. K. Damås, P. Aukrust, T. E. Mollnes, B. Halvorsen, and T. Espevik, "Cholesterol crystals use complement to increase NLRP3 signaling pathways in coronary and carotid atherosclerosis," *EBioMedicine*, vol. 60, 10 2020. [Online]. Available: /pmc/articles/PMC7494683//pmc/articles/PMC7494683/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7494683/ 1.2.1

[16] P. Duewell, H. Kono, K. J. Rayner, C. M. Sirois, G. Vladimer, F. G. Bauernfeind, G. S. Abela, L. Franchi, G. Nũez, M. Schnurr, T. Espevik, E. Lien, K. A.

Fitzgerald, K. L. Rock, K. J. Moore, S. D. Wright, V. Hornung, and E. Latz, "NLRP3 inflammasomes are required for atherogenesis and activated by cholesterol crystals," *Nature 2010 464:7293*, vol. 464, no. 7293, pp. 1357–1361, 4 2010. [Online]. Available: https://www.nature.com/articles/nature08938 1.2.1

[17] National Institute of Arthritis and Musculoskeletal and Skin Diseases, "Autoinflammatory Diseases," 1 2017. [Online]. Available: https://www.niams.nih.gov/health-topics/autoinflammatory-diseases 1.2.2

[18] J. Krainer, S. Siebenhandl, and A. Weinhäusel, "Systemic autoinflammatory diseases," *Journal of Autoimmunity*, vol. 109, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0896841120300329 1.2.2

[19] A. Simon, H. Park, R. Maddipati, A. A. Lobito, A. C. Bulua, A. J. Jackson, J. J. Chae, R. Ettinger, H. D. De Koning, A. C. Cruz, D. L. Kastner, H. Komarow, R. M. Siegel, and C. A. Dinarello, "Concerted action of wild-type and mutant TNF receptors enhances inflammation in TNF receptor 1-associated periodic fever syndrome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 21, pp. 9801–9806, 2010. [Online]. Available: www.pnas.org/cgi/doi/10.1073/pnas.0914118107 1.2.2

[20] I. Touitou and I. Koné-Paut, "Autoinflammatory diseases," *Best Practice & Research Clinical Rheumatology*, vol. 22, no. 5, pp. 811–829, 10 2008. 1.2.2

[21] Merriam-Webster, "Vasculitis." [Online]. Available: https://www.merriam-webster.com/dictionary/vasculitis 1.2.2

[22] I. Meyts and I. Aksentijevich, "Deficiency of Adenosine Deaminase 2 (DADA2): Updates on the Phenotype, Genetics, Pathogenesis, and Treatment," *Journal of Clinical Immunology*, vol. 38, no. 5, p. 578, 6 2018. [Online]. Available: https://link.springer.com/article/10.1007/s10875-018-0525-8 1.2.2, 1.3.2, 1.4.1, 4.3

[23] Centers for Disease Control and Prevention, "Rheumatic Diseases and Pain," 8 2022. [Online]. Available: https://www.cdc.gov/arthritis/communications/features/rheumatic-diseases-and-pain.html 1.2.2

[24] J. E. Hanks and D. Levine, "Rheumatic conditions," *A Comprehensive Guide to Geriatric Rehabilitation, Third Edition*, pp. 134–140, 1 2014. [Online]. Available: https://www.sciencedirect.com/topics/medicine-and-dentistry/rheumatic-disease 1.2.2

[25] Z. Szekanecz, I. B. McInnes, G. Schett, S. Szamosi, S. Benkő, and G. Szűcs, "Autoinflammation and autoimmunity across rheumatic and musculoskeletal diseases," *Nature Reviews Rheumatology*, vol. 17, pp. 585–595, 2021. [Online]. Available: https://www.nature.com/articles/s41584-021-00652-9.pdf 1.2.2

[26] P. C. Trivedi, J. J. Bartlett, and T. Pulinilkunnil, "Lysosomal Biology and Function: Modern View of Cellular Debris Bin," *Cells*, vol. 9, no. 5, 5 2020. [Online]. Available: /pmc/articles/PMC7290337//pmc/articles/PMC7290337/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7290337/ 1.2.3, 1.2.3, 4.3

[27] M. B. Rao, A. M. Tanksale, M. S. Ghatge, and V. V. Deshpande, "Molecular and Biotechnological Aspects of Microbial Proteases," *Microbiology and Molecular Biology Reviews*, vol. 62, no. 3, pp. 597–635, 9 1998. 1.2.3

[28] L. Zhang and L. J. Reha-Krantz, "Nuclease," in *Brenner's Encyclopedia of Genetics: Second Edition*. Academic Press, 1 2013, pp. 118–123. 1.2.3

[29] F. Labroussaa, S. Torres-Puig, and J. Jores, "Nucleases," in *Methods in Microbiology*. Academic Press, 1 2023, vol. 52, ch. 1.4.6.5, pp. 3–32. 1.2.3

[30] B. Winchester, "Lysosomal metabolism of glycoproteins," *Glycobiology*, vol. 15, no. 6, pp. 1R–15R, 1 2005. [Online]. Available: https://dx.doi.org/10.1093/glycob/cwi041 1.2.3

[31] C. J. Evans and R. J. Aguilera, "DNase II: genes, enzymes and function," *Gene*, vol. 322, pp. 1–15, 12 2003. 1.2.3

[32] Y. Y. Lan, D. Londoño, R. Bouley, M. S. Rooney, and N. Hacohen, "Dnase2a deficiency uncovers lysosomal clearance of damaged nuclear DNA via autophagy," *Cell reports*, vol. 9, no. 1, p. 192, 10 2014. [Online]. Available: /pmc/articles/PMC4555847//pmc/articles/PMC4555847/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4555847/ 1.2.3, 4.3

[33] M. P. Rodero, A. Tesser, E. Bartok, G. I. Rice, E. Della Mina, M. Depp, B. Beitz, V. Bondet, N. Cagnard, D. Duffy, M. Dussiot, M. L. Frémond, M. Gattorno, F. Guillem, N. Kitabayashi, F. Porcheray, F. Rieux-Laucat, L. Seabra, C. Uggenti, S. Volpi, L. A. Zeef, M. A. Alyanakian, J. Beltrand, A. M. Bianco, N. Boddaert, C. Brouzes, S. Candon, R. Caorsi, M. Charbit, M. Fabre, F. Faletra, M. Girard, A. Harroche, E. Hartmann, D. Lasne, A. Marcuzzi, B. Neven, P. Nitschke, T. Pascreau, S. Pastore, C. Picard,

P. Picco, E. Piscianz, M. Polak, P. Quartier, M. Rabant, G. Stocco, A. Taddio, F. Uettwiller, E. Valencic, D. Vozzi, G. Hartmann, W. Barchet, O. Hermine, B. Bader-Meunier, A. Tommasini, and Y. J. Crow, "Type I interferon-mediated autoinflammation due to DNase II deficiency," *Nature Communications 2017 8:1*, vol. 8, no. 1, 12 2017. [Online]. Available: https://www.nature.com/articles/s41467-017-01932-3 1.2.3, 4.3

[34] O. K. Greiner-Tollersrud, T. Berg, P. Healy, G. Evjen, U. Ramachandran, and Nilssen, "Purification of bovine lysosomal alpha-mannosidase, characterization of its gene and determination of two mutations that cause alpha-mannosidosis," *European journal of biochemistry*, vol. 246, no. 2, pp. 410–419, 6 1997. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/9208932/ 1.2.3, 2.2.3

[35] P. F. Daniel, B. Winchester, and C. D. Warren, "Mammalian α-mannosidases—multiple forms but a common purpose?" *Glycobiology*, vol. 4, no. 5, pp. 551–566, 1994. 1.2.3

[36] D. Malm and Nilssen, "Alpha-mannosidosis," *Orphanet Journal of Rare Diseases*, vol. 3, no. 1, p. 21, 2008. [Online]. Available: https://doi.org/10.1186/1750-1172-3-21https://ojrd.biomedcentral.com/articles/10.1186/1750-1172-3-21#Abs1 1.2.3

[37] European Medicines Agency, "Lamzede | European Medicines Agency," 2 2023. [Online]. Available: https://www.ema.europa.eu/en/medicines/human/EPAR/lamzede 1.2.3

[38] A. Varki and S. Kornfeld, "Historical Background and Overview," in *Essentials of Glycobiology*, 4th ed., A. Varki, R. Cummings, J. Esko, and et al., Eds. NY: Cold Spring Harbor Laboratory Press, 2022, ch. 1. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK579927/ 1.2.4, 1.2.4

[39] Karen J. Colley, Ajit Varki, Robert S. Haltiwanger, and Taroh Kinoshita, "Cellular Organization of Glycosylation," in *Essentials of Glycobiology*, 4th ed., Varki A, Cummings RD, Esko JD, and Et Al, Eds. NY: Cold Spring Harbor Laboratory Press, 2022, ch. 4. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK579926/ 1.2.4, 1.2.4

[40] A. Varki, R. D. Cummings, M. Aebi, N. H. Packer, P. H. Seeberger, J. D. Esko, P. Stanley, G. Hart, A. Darvill, T. Kinoshita, J. J. Prestegard, R. L. Schnaar, H. H. Freeze, J. D. Marth, C. R. Bertozzi, M. E. Etzler, M. Frank, J. F. Vliegenthart, T. Lütteke, S. Perez, E. Bolton, P. Rudd, J. Paulson, M. Kanehisa, P. Toukach,

K. F. Aoki-Kinoshita, A. Dell, H. Narimatsu, W. York, N. Taniguchi, and S. Kornfeld, "Symbol nomenclature for graphical representations of glycans," *Glycobiology*, vol. 25, no. 12, pp. 1323–1324, 12 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/glycans/snfg.html 1.2.4

[41] P. M. Stanley, Kelley W. Lewis, Nathan E. Taniguchi, and M. Naoyuki Aebi, "N-Glycans," in *Essentials of Glycobiology*, 4th ed., Varki A, Cummings R, Esko JD, and et al., Eds. NY: Cold Spring Harbor Laboratory Press, 2022, ch. 9. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK579964/ 1.2.4, 1.2.4, 1.2.4, 5, 6, 1.2.4, 7, 1.2.4, 1.2.4

[42] Q. Chao, Y. Ding, Z.-H. Chen, M.-H. Xiang, N. Wang, and X.-D. Gao, "Recent Progress in Chemo-Enzymatic Methods for the Synthesis of N-Glycans," *Frontiers in Chemistry*, vol. 8, p. 513, 1 2020. [Online]. Available: https://www.researchgate.net/figure/Structures-and-symbols-of-common-monosaccharides-in-N-glycans_fig1_342217529 4

[43] N. Dahms, T. Braulke, and A. Varki, "P-Type Lectins," in *Essentials of Glycobiology*, 4th ed. NY: Cold Spring Harbor Laboratory Press, 2022, ch. 33. 1.2.4

[44] P. Gagneux, T. Hennet, and A. Varki, "Biological Functions of Glycans," in *Essentials of Glycobiology*, 4th ed. NY: Cold Spring Harbor Laboratory Press, 2022, ch. 7. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK579984/ 1.2.4

[45] R. L. Schnaar, "Glycobiology simplified: diverse roles of glycan recognition in inflammation," *Journal of Leukocyte Biology*, vol. 99, no. 6, pp. 825–838, 6 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4952015/ 1.2.4

[46] O. K. Greiner-Tollersrud, V. Venkatakrishnan, M. Wuhrer, and N. G. Karlsson, "Introduction of «tomi» N-glycans as a new term to group lysosomal type N-glycan degradation products," *Under review in Glycobiology*, 2024. 1.2.4, 8, 2.3.1

[47] D. M. SWALLOW, L. F. WEST, and A. V. ELSEN, "The role of lysosomal sialidase and β-galactosidase in processing the complex carbohydrate chains on lysosomal enzymes and possibly other glycoproteins," *Annals of Human Genetics*, vol. 48, no. 3, pp. 215–221, 1984. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/6431895/ 1.2.4

[48] M. Damme, W. Morelle, B. Schmidt, C. Andersson, J. Fogh, J.-C. Michalski, and T. Lübke, "Impaired Lysosomal Trimming of N-Linked Oligosaccharides Leads to Hyperglycosylation of Native Lysosomal Proteins in Mice with α-Mannosidosis," *Molecular and Cellular Biology*, vol. 30, no. 1, p. 283, 1 2010. [Online]. Available: /pmc/articles/PMC2798300//pmc/articles/PMC2798300/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC2798300/ 1.2.4, 2.3.1

[49] A. V. Zavialov, X. Yu, D. Spillmann, G. Lauvau, and A. V. Zavialov, "Structural Basis for the Growth Factor Activity of Human Adenosine Deaminase ADA2," *Journal of Biological Chemistry*, vol. 285, no. 16, pp. 12 367–12 377, 4 2010. 1.3, 1.3.2, 1.3.2, 11, 12, 1.3.4, 4.3

[50] A. Zavialov and Engström, "Human ADA2 belongs to a new family of growth factors with adenosine deaminase activity," *Biochemical Journal*, vol. 391, no. 1, pp. 51–57, 10 2005. 1.3, 1.3, 1.3.2

[51] G. Haskó, J. Linden, B. Cronstein, and P. Pacher, "Adenosine receptors: therapeutic aspects for inflammatory and immune diseases," *Nature reviews. Drug discovery*, vol. 7, no. 9, p. 770, 9 2008. [Online]. Available: /pmc/articles/PMC2568887//pmc/articles/PMC2568887/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC2568887/ 1.3.1, 4.3

[52] P. A. Borea, S. Gessi, S. Merighi, F. Vincenzi, and K. Varani, "Pathological overproduction: the bad side of adenosine," *British Journal of Pharmacology*, vol. 174, no. 13, pp. 1945–1960, 7 2017. 1.3.1

[53] H. Liu and Y. Xia, "Beneficial and detrimental role of adenosine signaling in diseases and therapy," *Journal of Applied Physiology*, vol. 119, no. 10, pp. 1173–1182, 11 2015. 1.3.1

[54] S. Iwaki-Egawa, C. Namiki, and Y. Watanabe, "Adenosine deaminase 2 from chicken liver: purification, characterization, and N-terminal amino acid sequence," *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, vol. 137, no. 2, pp. 247–254, 2 2004. 1.3.2, 1.3.4, 2.2.2, 8, 4.3

[55] The Human Protein Atlas, "ADA2 protein expression summary - The Human Protein Atlas." [Online]. Available: https://www.proteinatlas.org/ENSG00000093072-ADA2 1.3.2

[56] T. K. Tarrant, S. J. Kelly, and M. S. Hershfield, "Elucidating the pathogenesis of adenosine deaminase 2 deficiency: current status and unmet needs," *Expert Opinion on Orphan Drugs*, vol. 9, no. 11-12, pp. 257–264, 12 2021. 1.3.2, 11, 1.3.5, 1.4.1, 18, 4.3

[57] Uniprot.org, "Biophysicochemical properties," 4 2022. [Online]. Available: https://www.uniprot.org/help/biophysicochemical_properties 1.3.2

[58] ——, "Q9NZK5 · ADA2_HUMAN." [Online]. Available: https://www.uniprot.org/uniprotkb/Q9NZK5/entry 1.3.2, 13, 14

[59] ——, "P00813 · ADA_HUMAN." [Online]. Available: https://www.uniprot.org/uniprotkb/P00813/entry 1.3.2

[60] C. Wu and C. H. Fry, "The effects of extracellular and intracellular pH on intracellular Ca2+ regulation in guinea-pig detrusor smooth muscle," *The Journal of Physiology*, vol. 508, no. Pt 1, p. 143, 4 1998. [Online]. Available: /pmc/articles/PMC2230873//pmc/articles/PMC2230873/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC2230873/ 1.3.2, 4.3

[61] Z. Chang and X. Fu, "Biogenesis of Secretory Proteins in Eukaryotic and Prokaryotic Cells," in *Encyclopedia of Cell Biology (Second Edition)*, second edition ed., R. A. Bradshaw, G. W. Hart, and P. D. Stahl, Eds. Oxford: Academic Press, 2023, pp. 689–702. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128216187000481 1.3.2

[62] M. Ito, Y. Maejima, K. Nishimura, Y. Nakae, A. Ono, and S. Iwaki-Egawa, "A role for N-glycosylation in active adenosine deaminase 2 production," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1866, no. 12, p. 130237, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304416522001556 1.3.3

[63] I. Aksentijevich, N. S. Moura, and K. Barron, "Adenosine Deaminase 2 Deficiency," *GeneReviews®*, 8 2019. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK544951/http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6339927 1.3.4

[64] M. D. Contreras-Aguilar, A. Tvarijonaviciute, I. Monkeviciene, M. Martín-Cuervo, L. G. González-Arostegui, L. Franco-Martínez, J. J. Cerón, F. Tecles, and D. Escribano, "Characterization of total adenosine deaminase activity (ADA) and its isoenzymes in saliva and serum in health and inflammatory

conditions in four different species: an analytical and clinical validation pilot study," *BMC Veterinary Research*, vol. 16, no. 384, 10 2020. [Online]. Available: /pmc/articles/PMC7549231//pmc/articles/PMC7549231/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7549231/ 1.3.4

[65] O. K. Greiner-Tollersrud, V. Boehler, E. Bartok, M. Krausz, A. Polyzou, J. Schepp, M. Seidl, J. O. Olsen, C. R. Smulski, S. Raieli, K. Hübscher, E. Trompouki, R. Link, H. Ebersbach, H. Srinivas, M. Marchant, D. Staab, D. Guerini, S. Baasch, P. Henneke, G. Kochs, G. Hartmann, R. Geiger, B. Grimbacher, M. Warncke, and M. Proietti, "ADA2 is a lysosomal DNase regulating the type-I interferon response," *bioRxiv*, no. 1, 6 2020. [Online]. Available: https://www.biorxiv.org/content/10.1101/2020.06.21.162990v2https://www.biorxiv.org/content/10.1101/2020.06.21.162990v2.abstract 1.3.4, 1.3.5, 4.1, 4.3

[66] Uniprot.org, "Q2VQV8 · Q2VQV8_CHICK." [Online]. Available: https://www.uniprot.org/uniprotkb/Q2VQV8/entry 14

[67] ——, "W5PV50 · W5PV50_SHEEP." [Online]. Available: https://www.uniprot.org/uniprotkb/W5PV50/entry 14

[68] A. V. Zavialov, E. Gracia, N. Glaichenhaus, R. Franco, A. V. Zavialov, and G. Lauvau, "Human adenosine deaminase 2 induces differentiation of monocytes into macrophages and stimulates proliferation of T helper cells and macrophages," *Journal of Leukocyte Biology*, vol. 88, no. 2, pp. 279–290, 5 2010. 1.3.5

[69] C. Carmona-Rivera, S. S. Khaznadar, K. W. Shwin, J. A. Irizarry-Caro, L. J. O'Neil, Y. Liu, K. A. Jacobson, A. K. Ombrello, D. L. Stone, W. L. Tsai, D. L. Kastner, I. Aksentijevich, M. J. Kaplan, and P. C. Grayson, "Deficiency of adenosine deaminase 2 triggers adenosine-mediated NETosis and TNF production in patients with DADA2," *Blood*, vol. 134, no. 4, p. 406, 7 2019. [Online]. Available: /pmc/articles/PMC6659253//pmc/articles/PMC6659253/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6659253/ 1.3.5

[70] R. Dhanwani, M. Takahashi, I. T. Mathews, C. Lenzi, A. Romanov, J. D. Watrous, B. Pieters, C. C. Hedrick, C. A. Benedict, J. Linden, R. Nilsson, M. Jain, and S. Sharma, "Cellular sensing of extracellular purine nucleosides triggers an innate IFN-β response," *Science Advances*, vol. 6, no. 30, 7 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7375821/ 1.3.5

[71] Z. G. Jiang, B. Sandhu, L. Feldbrügge, E. U. Yee, E. Csizmadia, S. Mitsuhashi, J. Huang, N. H. Afdhal, S. C. Robson, and M. Lai, "Serum Activity of Macrophage-Derived Adenosine Deaminase 2 Is Associated With Liver Fibrosis in Nonalcoholic Fatty Liver Disease," *Clinical Gastroenterology and Hepatology*, vol. 16, no. 7, pp. 1170–1172, 7 2018. 1.3.5

[72] S. Nanthapisal, C. Murphy, E. Omoyinmi, Y. Hong, A. Standing, S. Berg, M. Ekelund, S. Jolles, L. Harper, T. Youngstein, K. Gilmour, N. J. Klein, D. Eleftheriou, and P. A. Brogan, "Deficiency of Adenosine Deaminase Type 2: A Description of Phenotype and Genotype in Fifteen Cases," *Arthritis & Rheumatology*, vol. 68, no. 9, pp. 2314–2322, 9 2016. [Online]. Available: https://acrjournals.onlinelibrary.wiley.com/doi/epdf/10.1002/art.39699 1.4, 16, 17, 4.3

[73] P. Navon Elkan, S. B. Pierce, R. Segel, T. Walsh, J. Barash, S. Padeh, A. Zlotogorski, Y. Berkun, J. J. Press, M. Mukamel, I. Voth, P. J. Hashkes, L. Harel, V. Hoffer, E. Ling, F. Yalcinkaya, O. Kasapcopur, M. K. Lee, R. E. Klevit, P. Renbaum, A. Weinberg-Shukron, E. F. Sener, B. Schormair, S. Zeligson, D. Marek-Yagel, T. M. Strom, M. Shohat, A. Singer, A. Rubinow, E. Pras, J. Winkelmann, M. Tekin, Y. Anikster, M.-C. King, and E. Levy-Lahad, "Mutant Adenosine Deaminase 2 in a Polyarteritis Nodosa Vasculopathy," *New England Journal of Medicine*, vol. 370, no. 10, pp. 921–931, 3 2014. [Online]. Available: https://www.nejm.org/doi/10.1056/NEJMoa1307362 1.4, 1.2, 1.4, 17, 4.3

[74] Q. Zhou, D. Yang, A. K. Ombrello, A. V. Zavialov, C. Toro, A. V. Zavialov, D. L. Stone, J. J. Chae, S. D. Rosenzweig, K. Bishop, K. S. Barron, H. S. Kuehn, P. Hoffmann, A. Negro, W. L. Tsai, E. W. Cowen, W. Pei, J. D. Milner, C. Silvin, T. Heller, D. T. Chin, N. J. Patronas, J. S. Barber, C.-C. R. Lee, G. M. Wood, A. Ling, S. J. Kelly, D. E. Kleiner, J. C. Mullikin, N. J. Ganson, H. H. Kong, S. Hambleton, F. Candotti, M. M. Quezado, K. R. Calvo, H. Alao, B. K. Barham, A. Jones, J. F. Meschia, B. B. Worrall, S. E. Kasner, S. S. Rich, R. Goldbach-Mansky, M. Abinun, E. Chalom, A. C. Gotte, M. Punaro, V. Pascual, J. W. Verbsky, T. R. Torgerson, N. G. Singer, T. R. Gershon, S. Ozen, O. Karadag, T. A. Fleisher, E. F. Remmers, S. M. Burgess, S. L. Moir, M. Gadina, R. Sood, M. S. Hershfield, M. Boehm, D. L. Kastner, and I. Aksentijevich, "Early-Onset Stroke and Vasculopathy Associated with Mutations in ADA2," *New England Journal*

of Medicine, vol. 370, no. 10, pp. 911–920, 3 2014. [Online]. Available: https://www.nejm.org/doi/10.1056/NEJMoa1307361 1.4, 1.2, 17, 4.3

[75] E. D. Batu, O. Karadag, E. Z. Taskiran, U. Kalyoncu, I. Aksentijevich, M. Alikasifoglu, and S. Ozen, "A Case Series of Adenosine Deaminase 2-deficient Patients Emphasizing Treatment and Genotype-phenotype Correlations," *The Journal of Rheumatology*, vol. 42, no. 8, pp. 1532–1534, 8 2015. [Online]. Available: https://www.jrheum.org/content/42/8/1532https://www.jrheum.org/content/42/8/1532.abstract 1.4, 1.2, 1.4, 17, 4.3

[76] The DADA2 Foundation, "What are the Signs & Symptoms of DADA2?" [Online]. Available: https://dada2.org/what-is-dada2/symptoms/#skin 15

[77] H. A. Simmonds, D. R. Webster, D. Perrett, S. Reiter, and R. J. Levinsky, "Formation and degradation of deoxyadenosine nucleotides in inherited adenosine deaminase deficiency," *Bioscience Reports*, vol. 2, no. 5, pp. 303–314, 5 1982. 1.4.1, 4.3

[78] A. Aiuti, M. G. Roncarolo, and L. Naldini, "Gene therapy for ADA-SCID, the first marketing approval of an ex vivo gene therapy in Europe: paving the road for the next generation of advanced therapy medicinal products," *EMBO molecular medicine*, vol. 9, no. 6, pp. 737–740, 6 2017. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/28396566/ 1.4.1

[79] S. S. I. Baraem and Nielsen, *Basic Principles of Chromatography*. Springer US, 2010, pp. 473–498. [Online]. Available: https://doi.org/10.1007/978-1-4419-1478-1_27 2.1.1

[80] D. S. Hage, M. Bian, R. Burks, E. Karle, C. Ohnmacht, and C. Wa, "Bioaffinity Chromatography ," in *Handbook of Affinity Chromatography*, 2nd ed., D. S. Hage, Ed. Licoln, Nebraska: CRC Press, 7 2005, ch. 5, pp. 101–126. 2.1.1

[81] FluoProbes, "Fluoprobes ® ft-ms902z," 2010. [Online]. Available: www.interchim.com/ 2.1.1

[82] S. Parkin, B. Rupp, and H. Hope, "Atomic Resolution Structure of Concanavalin A at 120 K," *Acta Crystallographica Section D Biological Crystallography*, vol. 52, no. 6, pp. 1161–1168, 11 1996. 1

[83] D. C. Harris, "Ion-Exchange Chromatography," in *Quantitative Chemical Analysis*, 8th ed. New York: W.H. Freeman and Co, 2010, ch. 25-1, pp. 635–641. 2.1.2, 2.1.2

[84] ——, "Conducting Capillary Electrophoresis," in *Quantitative Chemical Analysis*, 8th ed. New York: W. H. Freeman and Co, 2010, ch. 25-7, pp. 657–665. 2.1.3

[85] ——, "Principles of Capillary Electrophoresis," in *Quantitative Chemical Analysis*, 8th ed. New York: W. H. Freeman and Co, 2010, ch. 25-6, pp. 650–657. 2.1.3, 2.1.3

[86] A. B. Nowakowski, W. J. Wobig, and D. H. Petering, "Native SDS-PAGE: high resolution electrophoretic separation of proteins with retention of native properties including bound metal ions," *Metallomics*, vol. 6, no. 5, pp. 1068–1078, 2014. 2.1.3

[87] T. F. Scientific, "Pierce™ dtt (dithiothreitol), no-weigh™ format," 2023. [Online]. Available: https://www.thermofisher.com/order/catalog/product/A39255?gclid=CjwKCAjwyY6pBhA9EiwAMzmfwcMhF8zjwUqfiiVxIIV_xqNdA65JNbc4gjp7vlAtUrtzSyQjOA6_7xoC9mwQAvD_BwE&ef_id=CjwKCAjwyY6pBhA9EiwAMzmfwcMhF8zjwUqfiiVxIIV_xqNdA65JNbc4gjp7vlAtUrtzSyQjOA6_7xoC9mwQAvD_BwE:G:s&s_kwcid=AL!3652!3!656414123971!!!g!!!12650672775!120562016536&cid=bid_pca_msp_r01_co_cp1359_pjt0000_bid00000_0se_gaw_dy_pur_con 2.1.3

[88] B. Alberts, A. Johnaon, J. Lewis, D. Morgan, M. Raff, K. Roberts, and P. Walter, "Analyzing proteins," in *Molecular biology of the cell*, 6th ed. New York: Garland Science, Taylor and Francis Group, 2015, ch. 8-3, pp. 452–463. 2.1.3, 2.1.3

[89] D. C. Harris, *Fundamentals of Spectrophotometry*, 8th ed. W. H. Freeman and Co, 2010, pp. 393–414. 2.2.1, 2.2.1, 2.2.1, 2.2.1, 5

[90] N. Shanker and S. L. Bane, "Basic Aspects of Absorption and Fluorescence Spectroscopy and Resonance Energy Transfer Methods," in *Biophysical Tools for Biologists, Volume One: In Vitro Techniques*, ser. Methods in Cell Biology. Academic Press, 2008, vol. 84, pp. 213–242. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0091679X07840088 2.2.1

[91] S. Carrara, "Towards new efficient nanostructured hybrid materials for ECL applications," Ph.D. dissertation, Marseille, 4 2017. [Online]. Available: https://www.researchgate.net/publication/321823164_Towards_new_efficient_nanostructured_hybrid_materials_for_ECL_applications 6

[92] Agilent, "UV-Vis Spectrophotometer, Routine UV-Vis, Cary 60 | Agilent." [Online]. Available: https://www.agilent.com/en/product/molecular-spectroscopy/uv-vis-uv-vis-nir-spectroscopy/uv-vis-uv-vis-nir-systems/cary-60-uv-vis-spectrophotometer 7

[93] H. Bisswanger, "Enzyme assays," *Perspectives in Science*, vol. 1, no. 1-6, pp. 41–55, 5 2014. 2.2.2

[94] S. Iwaki-Egawa, WatanabeY, and H. Matsuno, "Correlations between matrix metalloproteinase-9 and adenosine deaminase isozymes in synovial fluid from patients with rheumatoid arthritis." *The Journal of Rheumatology*, vol. 28, no. 3, pp. 485–489, 2001. [Online]. Available: https://www.jrheum.org/content/28/3/485 2.2.2

[95] Y. F. Liao, A. Lal, and K. W. Moremen, "Cloning, Expression, Purification, and Characterization of the Human Broad Specificity Lysosomal Acid α-Mannosidase," *Journal of Biological Chemistry*, vol. 271, no. 45, pp. 28 348–28 358, 11 1996. 2.2.3

[96] G. Hansen, T. Berg, H. M. Riise Stensland, P. Heikinheimo, H. Klenow, G. Evjen, Nilssen, and O. K. Tollersrud, "Intracellular transport of human lysosomal α-mannosidase and α-mannosidosis-related mutants," *Biochemical Journal*, vol. 381, no. Pt 2, p. 537, 7 2004. [Online]. Available: /pmc/articles/PMC1133862//pmc/articles/PMC1133862/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC1133862/ 2.2.3

[97] D. C. Harris, " What Is Mass Spectrometry?" in *Quantitave Chemical Analysis*, 8th ed., J. Fiorillo, D. Quinn, K. Treadway, T. Szczepanski, and M. L. Byrd, Eds. New York: W. H. Freeman and Company, 2010, ch. 21-1, pp. 502–506. 2.3

[98] ——, "Chromatography–Mass Spectrometry," in *Quantitave Chemical Analysis*, 8th ed., J. Fiorillo, D. Quinn, K. Treadway, T. Szczepanski, and M. L. Byrd, Eds. New York: W. H. Freeman and Company, 2010, ch. 21-4, pp. 519–529. 2.3

[99] W. B. Dunn, "Mass Spectrometry in Systems Biology: An Introduction," *Methods in Enzymology*, vol. 500, pp. 15–35, 9 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780123851185000025?ref=pdf_download&fr=RR-2&rr=85c05a8e4919b515 2.3

[100] A.-N. Neagu, M. Jayathirtha, E. Baxter, M. Donnelly, B. A. Petre, and C. C. Darie, "Applications of Tandem Mass Spectrometry (MS/MS) in Protein Analysis for Biomedical Research," *Molecules*, vol. 27, no. 8, p. 2411, 4 2022. 2.3

[101] N. Mehta, M. Porterfield, W. B. Struwe, C. Heiss, P. Azadi, P. M. Rudd, M. Tiemeyer, and K. Aoki, "Mass Spectrometric Quantification of N-Linked Glycans by Reference to Exogenous Standards," *Journal of proteome research*, vol. 15, no. 9, p. 2980, 9 2016. [Online]. Available: /pmc/articles/PMC5501980//pmc/articles/PMC5501980/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5501980/ 2.3.1

[102] UniProt.org, "VNN2 - Vanin 2 - Gallus gallus (Chicken) | UniProtKB | UniProt." [Online]. Available: https://www.uniprot.org/uniprotkb/A0A8V0XM46/entry#function 4.2.1

[103] ——, "LOC107056139 - CN hydrolase domain-containing protein - Gallus gallus (Chicken) | UniProtKB | UniProt." [Online]. Available: https://www.uniprot.org/uniprotkb/A0A8V0XW73/entry 4.2.1

[104] ——, "GNS - N-acetylglucosamine-6-sulfatase - Gallus gallus (Chicken) | UniProtKB | UniProt." [Online]. Available: https://www.uniprot.org/uniprotkb/F1NI04/entry#function 4.2.1

[105] ——, "VNN2 - Pantetheine hydrolase VNN2 - Homo sapiens (Human) | UniProtKB | UniProt." [Online]. Available: https://www.uniprot.org/uniprotkb/O95498/entry#subcellular_location 4.2.1

[106] ——, "GNS - N-acetylglucosamine-6-sulfatase - Ovis aries (Sheep) | UniProtKB | UniProt." [Online]. Available: https://www.uniprot.org/uniprotkb/W5NZ62/entry#function 4.2.2

[107] ——, "SGSH - N-sulfoglucosamine sulfohydrolase - Ovis aries (Sheep) | UniProtKB | UniProt." [Online]. Available: https://www.uniprot.org/uniprotkb/W5NUI6/entry#function 4.2.2

[108] ——, "GALNS - N-acetylgalactosamine-6-sulfatase - Ovis aries (Sheep) | UniProtKB | UniProt." [Online]. Available: https://www.uniprot.org/uniprotkb/W5PVT3/entry#function 4.2.2

[109] ——, "LOC101123672 - NTR domain-containing protein - Ovis aries (Sheep) | UniProtKB | UniProt." [Online]. Available: https://www.uniprot.org/uniprotkb/W5NUX8/entry#function 4.2.2

[110] ——, "LOC101122940 - Alpha-2-macroglobulin - Ovis aries (Sheep) | UniProtKB | UniProt." [Online]. Available: https://www.uniprot.org/uniprotkb/W5NSA6/entry#names_and_taxonomy 4.2.2

[111] A. Riemann, H. Wußling, H. Loppnow, H. Fu, S. Reime, and O. Thews, "Acidosis differently modulates the inflammatory program in monocytes and macrophages," *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, vol. 1862, no. 1, pp. 72–81, 1 2016. 4.3

[112] A. J. van der Meer, A. Kroeze, A. J. Hoogendijk, A. A. Soussan, C. E. van der Schoot, W. A. Wuillemin, C. Voermans, T. van der Poll, and S. Zeerleder, "Systemic inflammation induces release of cell-free DNA from hematopoietic and parenchymal cells in mice and humans," *Blood Advances*, vol. 3, no. 5, p. 728, 3 2019. [Online]. Available: /pmc/articles/PMC6418504//pmc/articles/PMC6418504/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6418504/ 4.3

[113] M. Moras, S. D. Lefevre, and M. A. Ostuni, "From erythroblasts to mature red blood cells: Organelle clearance in mammals," *Frontiers in Physiology*, vol. 8, 12 2017. [Online]. Available: www.frontiersin.orghttps://www.frontiersin.org/journals/physiology/articles/10.3389/fphys.2017.01076/full 4.3

[114] V. V. Iyengar, A. Chougule, V. Gowri, P. Taur, S. Prabhu, M. Bodhanwala, and M. M. Desai, "DADA2 presenting as nonimmune hemolytic anemia with recurrent macrophage activation syndrome," *Pediatric Blood & Cancer*, vol. 69, no. 6, 11 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/pbc.29461https://onlinelibrary.wiley.com/doi/abs/10.1002/pbc.29461https://onlinelibrary.wiley.com/doi/10.1002/pbc.29461 4.3

[115] S. Özen, E. D. Batu, E. Z. Taşkiran, H. A. Özkara, Ünal, N. Güleray, A. Erden, Karadağ, F. Gümrük, M. Çetin, H. E. Sönmez, Y. Bilginer, D. Ayvaz, and I. Tezcan, "A Monogenic Disease with a Variety of Phenotypes: Deficiency of Adenosine Deaminase 2," *The Journal of Rheumatology*, vol. 47, no. 1, pp. 117–125, 1 2020. [Online]. Available: https://www.jrheum.org/content/47/1/117https://www.jrheum.org/content/47/1/117.abstract 4.3

[116] P. Y. Lee, E. S. Kellner, Y. Huang, E. Furutani, Z. Huang, W. Bainter, M. F. Alosaimi, K. Stafstrom, C. D. Platt, T. Stauber, S. Raz, I. Tirosh, A. Weiss, M. B. Jordan, C. Krupski, D. Eleftheriou, P. Brogan, A. Sobh, Z. Baz, G. Lefranc, C. Irani, S. S. Kilic, R. El-Owaidy, M. R. Lokeshwar, P. Pimpale, R. Khubchandani, E. P. Chambers, J. Chou, R. S. Geha, P. A. Nigrovic, and Q. Zhou, "Genotype and functional correlates of disease phenotype in deficiency of adenosine deaminase 2 (DADA2)," *The Journal of allergy and clinical immunology*, vol. 145, no. 6, p. 1672, 1 2020.

[Online]. Available: /pmc/articles/PMC7282972//pmc/articles/PMC7282972/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7282972/ 4.3

[117] D. C. Barral, L. Staiano, C. Guimas Almeida, D. F. Cutler, E. R. Eden, C. E. Futter, A. Galione, A. R. Marques, D. L. Medina, G. Napolitano, C. Settembre, O. V. Vieira, J. M. Aerts, P. Atakpa-Adaji, G. Bruno, A. Capuozzo, E. De Leonibus, C. Di Malta, C. Escrevente, A. Esposito, P. Grumati, M. J. Hall, R. O. Teodoro, S. S. Lopes, J. P. Luzio, J. Monfregola, S. Montefusco, F. M. Platt, R. Polishchuck, M. De Risi, I. Sambri, C. Soldati, and M. C. Seabra, "Current methods to analyze lysosome morphology, positioning, motility and function," *Traffic (Copenhagen, Denmark)*, vol. 23, no. 5, p. 269, 4 2022. [Online]. Available: /pmc/articles/PMC9323414//pmc/articles/PMC9323414/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC9323414/ 4.3

[118] K. Loell and V. Nanda, "Marginal protein stability drives subcellular proteome isoelectric point," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, no. 46, pp. 11 778–11 783, 11 2018. [Online]. Available: /pmc/articles/PMC6243250//pmc/articles/PMC6243250/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6243250/ 4.3

[119] C. Park, L. Meng, L. H. Stanton, R. E. Collins, S. W. Mast, X. Yi, H. Strachan, and K. W. Moremen, "Characterization of a Human Core-specific Lysosomal α1,6-Mannosidase Involved in N-Glycan Catabolism," *The Journal of biological chemistry*, vol. 280, no. 44, p. 37216, 11 2005. [Online]. Available: /pmc/articles/PMC1351102//pmc/articles/PMC1351102/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC1351102/ 4.3