University of South Bohemia in České Budějovice

Faculty of Science

Identification of Glycosylated Tick Proteins Containing Immunogenic Epitopes

Master thesis

Bc. Zuzana Dvorníková

Supervisor: RNDr. Ján Štěrba, Ph.D.

Consultant: RNDr. Jarmila Štěrbová

České Budějovice 2023

Dvorníková, Z., 2023: Identification of Glycosylated Tick Proteins Containing Immunogenic Epitopes. Mgr. Thesis, in English – 84 pp., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Annotation

Glycosylation is one of the most common post-translational modifications of proteins. The saccharide parts bound to the proteins then plays a role in triggering immune responses. Ticks are blood-sucking parasites that transfer different pathogens into the human body, causing very serious illnesses. However, transmission of the pathogens via ticks still hasn't been fully described in terms of its process. The aim of this thesis was to identify and purify proteins carrying potentially immunogenic α1,3-core fucose, sialic acid, and galactose-α-1,3-galactose epitopes from IRE/CTVM20, ISE18, and OME/CTVM22 tick cell lines using biochemical methods and subsequent mass spectrometry analysis. Furthermore, visualisation of individual epitopes in every cell line using fluorescence microscopy was done.

Declaration

I declare that I am the author of this qualification thesis and that in writing it I have used only the sources and literature displayed in the list of used sources.

České Budějovice, 2023

……

Bc. Zuzana Dvorníková

Acknowledgement

I would like to thank to my supervisor RNDr. Ján Štěrba, Ph.D., to kindly accept me in his lab, for his valuable advice and support. I would like to also thank my consultant RNDr. Jarmila Štěrbová for supervision. Furthermore, I would like to thank Mgr. Filip Dyčka, Ph.D. for helping with the proteomic part of this thesis. Big thank also goes to RNDr. Pavlína Věchtová, Ph.D for teaching me how to take care for our precious tick cell lines.

I am also very grateful for all my colleagues from the lab of Applied Biochemistry, for always helping me when I was lost. I am grateful to have my lab buddy such as Káťa J.

Finally, I would like to thank my family, and my boyfriend for being my biggest supporters with the biggest amount of patience.

Contents

List of abbreviations

Ara	D-arabinose
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BSA	bovine serum albumin
BT	bovine transferrin
CHCA	α-cyano-4-hydroxycinnamic acid
DTT	dithiotreitol
Endo H	endoglycosidase H
ER	endoplasmic reticulum
ESI	electrospray ionisation
FA	formic acid
Fuc	D-Fucose
Fuc-1-P	fucose-1-phosphate
FucNAc	N-acetyl-D-fucosamine
GA	Golgi apparatus
Gal	D-galactose
GalNAc	N-acetyl-D-galactosamine
GdnSCN	guanidine thiocyanate
Glc	D-glucose
$Glc_3Man_9GlcNAc_2$	$Glucose_3Mannose_9N\text{-}acetylglucosamine_2$
GlcNAc	N-acetyl-D-glucosamine
GlcNAc-1-P	N-acetyl-D-glucosaminephosphotransferase
GSL-B4	*Griffonia simplicifolia* lectin I, isolectin B4
HRP	horse radish peroxidase
IAA	indole-3-acetic acid
MAL II	*Maackia amurensis* lectin
MALDI	matrix assisted laser desorption/ionization
Man	D-Mannose
Man-1-P	mannose-1-phosphate
MS	mass spectrometry
NBT	nitro blue tetrazolium
$Neu5,9Ac_2$	N-acetyl-9-O-acetyl neuraminic acid

Neu5Ac	N-acetylneuraminic acid
Neu5Gc	N-glycolylneuraminic acid
OST	oligosaccharotransferase
PAGE	polyacrylamide gel electrophoresis
PNGase F	peptide:N-glycosidase F
PVDF	polyvinylidene difluoride
Rha	D-Rhamnose
SDS	sodium dodecyl sulfate
SNA	*Sambucus nigra* lectin
TFA	Trifluoroacetic acid
Xyl	D-xylose
Xyl-1-P	xylose-1-phosphate

1 Introduction

Protein glycosylation has crucial functions in protein maturation and interactions with other molecules, as well as influencing protein characteristics and activities. All eukaryotic cells require protein glycosylation, which is involved in a number of critical cellular activities. The significance of glycans in pathogen recognition and interactions between vectors and hosts is an intriguing element of glycan function. Many interactions in the immune system, which is made up of an intricate network of cells and molecules that work together to activate the host defence system, require certain carbohydrate structures and proteins that precisely detect and bind them (De Schutter & Van Damme, 2015; Helenius & Zurich, 2001).

One of the most prevalent blood-sucking ectoparasites of people and other animals are ticks alongside with mosquitoes. They are some of the most significant pathogen carriers and are responsible for a variety of illnesses, and we encounter them all over the world. The transmission of pathogens by ticks has still not been properly described and is studied by many scientists.

The aim of this work was to identify proteins carrying α1,3-core fucose, sialic acid, and galactose-α-1,3-galactose epitopes from IRE/CTVM20, ISE18, and OME/CTVM22 tick cell lines using SDS-PAGE and Western blot followed by mass spectrometry analysis. Another task was to purify these glycosylated proteins from the three tick cell lines using biochemical methods with subsequent mass analysis. The last goal was to visualise these epitopes in all three cell lines using fluorescence microscopy.

2 Literal review

2.1 Glycoproteins

Glycoproteins are formed by of cross- or post-translational modifications, in which oligosaccharide chains are covalently bound to amino acid side chains. This process is known as glycosylation. Secreted extracellular proteins are often glycosylated. Glycosylation is the most abundant post-translational modification in organisms, and it plays a key role in the function of these proteins (Seeberger, 2022).

The term “glycan” refers to the saccharide part of these molecules. Glycans are divided into oligosaccharides and polysaccharides. Oligosaccharides consist of 10 to 20 monosaccharides residues connected by glycosidic linkages. On the other hand, polysaccharides consist of more than 20 monosaccharides residues. Monosaccharides are the primary constituents of glycans and are often referred to as carbohydrates because they contain carbonyl groups (aldehyde or ketone) (Seeberger, 2022). Each monosaccharide is terminated at one end of the chain with hydroxymethylene unit, and the other end is terminated with a hydroxymethyl group. If a monosaccharide contains an aldehyde group, the saccharide is called aldose, and if it contains a ketone group, it is called ketose. For aldoses, the simplest one is glyceraldehyde, and for ketoses, dihydroxyacetone (Seeberger, 2022).

The main building blocks of proteins are amino acids. Proteins are made out of 20 basic amino acids, which share a basic structure made from one atom of carbon, to which is connected a carboxyl group (-COOH), an amino group ($-NH_2$), and one atom of hydrogen. Their differences are based on the final (fourth) group, which is different for every amino acid and defines its chemical properties. Together, they are connected by peptide bonds resulting in peptides (up to 50 amino acids in a long chain) or complex proteins (more than 50 amino acids) (Connie et al., 2013).

2.1.1 Functions of glycoproteins' glycans

Glycans have a wide range of biological functions, from very minor to essential for the growth, maintenance, or survival of the organism that produces them, similar to other important groups of macromolecules. Many glycans lack well-defined roles as of yet. Depending on the aglycone (protein or lipid) to which a certain glycan is bound, the glycan may also serve other purposes. More and more complicated glycans evolved because they are a liability for the organism that synthesises them since they are commonly bound by bacteria and microbial toxins. In general, certain biological activities inside the organism are more

likely to be mediated by terminal sequences, unusual structures, and glycan modifications (Seeberger, 2022).

Numerous barriers, organizational, stabilizing, and protecting roles are provided by glycans. In plants, cellulose serves as the building block for cell walls, whereas chitin serves the same purpose in fungi and arthropods. Most glycoproteins have glycans on the surface, which can act as a general barrier, preventing proteases from recognising the underlying polypeptide (mucins), preventing antibodies from binding (viral glycoproteins), and even (as in the case of mucins) shielding entire tissue surfaces from microbial attachment. Proteoglycans, one kind of glycan found in matrix molecules, are crucial for maintaining the structure, porosity, and integrity of tissues in multicellular organisms (Seeberger, 2022).

The endoplasmic reticulum's role in folding freshly generated polypeptides and/or in the subsequent preservation of protein solubility and shape are two additional structural functions performed by glycans. Uncorrected glycosylation of certain proteins can prevent them from folding correctly or from leaving the endoplasmic reticulum, in fact. These misfolded glycoproteins are then transported to proteasomes, where they are degraded (Seeberger, 2022; Taylor & Drickamer, 2011).

In addition to their structural role, glycans serve as a protective storage for physiologically significant molecules. They also shield them from proteolysis, extends their active lifetimes, and enables their release under specified circumstances. (Kobata, 1992; Seeberger, 2022).

2.2 Glycosylation

As mentioned in the previous chapter, glycosylation is the process by which glycans are connected to proteins. The synthesis of glycoproteins is not entirely encoded in the genome. The primary structure of protein part is encoded in the genome, and through the process of transcription and translation of the mRNA template, a copy of the protein is synthesised. Protein activity is defined based on its three-dimensional structure. On the contrary, glycans do not have a genome-encoded template, and they are assembled through a series of enzymatic reactions performed by glycosyltransferases and glycosidases. However, glycosyltransferases are encoded in the genome. Glycan structure synthesis can be controlled by the transcription and translation of genes that generate glycosyltransferases and glycosidases (Taylor & Drickamer, 2011).

To date, there are four known distinct types of glycan-protein bonds. These are summarised in Tab. I. More than one glycan and more than one glycosylation type can occur at the same protein, it depends on the amino acid sequence of the protein and the availability and activity of enzymes (Spiro, 2002).

Table I: Summary of protein-glycan glycosidic bonds. Separated by the type of linkage (to which atom does the glycan structure bind in the protein), to which amino acid is the glycan linked and which saccharide is proximal (innermost), bound to the protein (Spiro, 2002).

Type of glycosylation	Linkage	
	Amino acid	Proximal saccharide
N-glycosylation	Asn	GlcNAc
		GalNAc
		Glc
		Rha
	Arg	Glc
O-glycosylation	Tyr	Gal
		Glc
	Hyl	Gal
	Ser/Thr	GlcNAc
		GalNAc
		Gal
		Man
		Fuc
	Hyp	GlcNAc
		Gal
		Ara
C-glycosylation (C-mannosyl)	Trp	Man
P-glycosylation	Ser	GlcNAc-1-P
		Man-1-P
		Fuc-1-P
		Xyl-1-P

The most abundant glycan types are those linked to proteins via N-glycosidic and O-glycosidic bonds. The variability of glycosylation is immense, whether between species, within one species, or in the biological structures of the same organism (Gagneux &Varki, 1999).

2.2.1 N-glycosylation

N-linked glycans are connected to the Asn side chain via its amide nitrogen (-NH_2) (Taylor & Drickamer, 2011). In eukaryotes, N-glycosylation occurs at an amino acid sequence called a sequon. Sequon is a triplet of amino acids Asn-Xxx-Ser/Thr, where Xxx is any amino acid except proline because it bends the polypeptide chain and thus, glycan could not be connected. This process is co-translational and is catalysed by oligosaccharyl transferase (Weerapana & Imperiali, 2006). According to the available data, N-glycosylation follows more structural rules than O-glycosylation, as O-glycan do not share one same core. To this day there are four known basic cores of O-glycans (Kobata, 1992).

The process of N-glycosylation will be described for Eukaryotes. N-glycosylation starts with the synthesis of a glycan precursor made up of 14 subunits, with the structure $Glc_3Man_9GlcNAc_2$ which is connected to a dolichol anchor (Taylor & Drickamer, 2011). Dolichol is a 14-24 units long isoprene chain that is found in the endoplasmic reticulum membrane (Carroll et al., 1992). In the process of the N-glycan precursor formation, GlcNAc is first linked to dolichol diphosphate forming diphosphate-dolichol-GlcNAc. Subsequently, another GlcNAc is connected, followed by five Mans. Until this step, the entire process proceeds on the cytoplasmatic side of the endoplasmic reticulum. After permutation into the lumen of the endoplasmic reticulum, four more Mans are connected, and finally three Glcs; thus, the whole precursor is made. OST is the enzyme responsible for the connection of the precursor to the protein. Both the dolichol-linked glycan precursor and the OST are anchored in the membrane on the luminal side; once the newly formed polypeptide chain produced by ribosomes is transported to the ER lumen, OST can transfer the precursor to an available sequon and dolichol is released and used again (Kobata, 1992; Taylor & Drickamer, 2011). The process of N-glycosylation is shown in Fig. 1.

The precursor, which has 14 units when transferred to the protein, is trimmed by ER glucosidases by three Glcs residues. Following trimming, glycoprotein folding occurs with the assistance of the calnexin/calreticulin complex, which acts as a quality control check to determine if glycoprotein folding is correct. The complex then determines if the now-trimmed precursor $Man_9GlcNAc_2$ will continue into the *cis*-GA, where it can be further processed until $Man_5GlcNAc_2$ is formed (Seeberger, 2022).

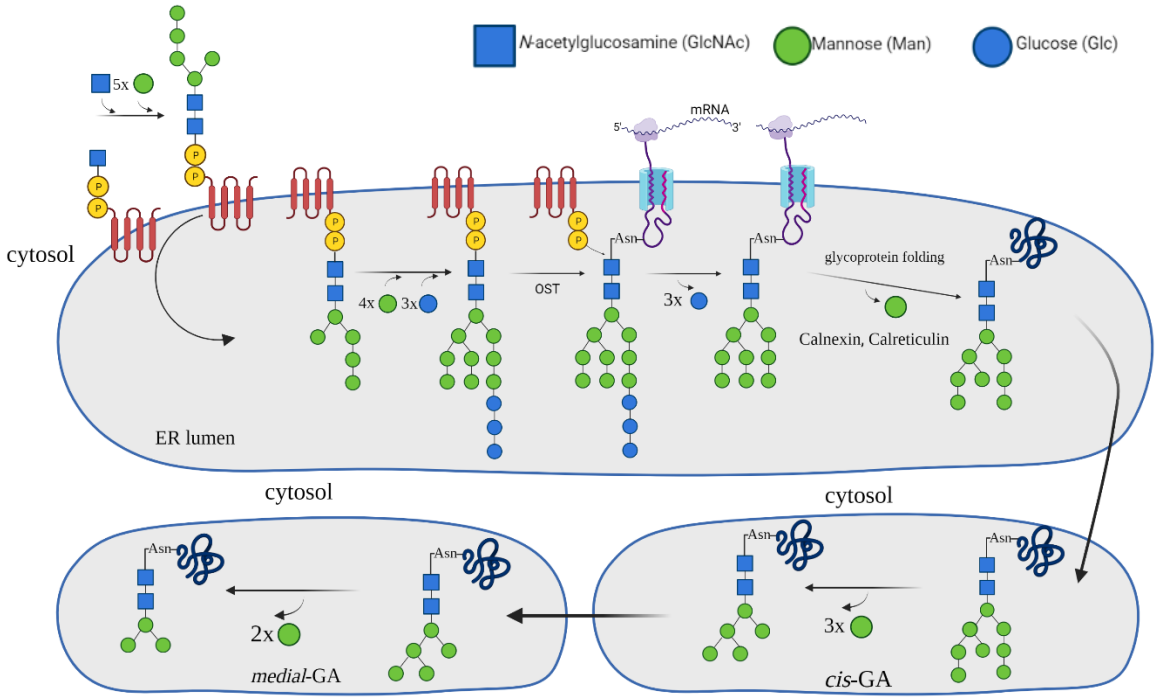


Figure 1. Process of N-glycosylation in ER and GA. Created with BioRender.com by the author according to sources (Esmail & Manolson, 2021; Seeberger, 2022).

Mannosidases trims more Mans residues in the *medial*-GA, leaving only three Mans and two GlcNAcs (Fig. 1), which can be further modified into various structures. N-linked glycans share the core structure derived from the precursor, but they differ in terminal residues that extend from this core. This process is occurring in *trans*-GA (Seeberger, 2022; Taylor & Drickamer, 2011). The three main common structures are complex, high-mannose, and hybrid types, and their schematic structures are represented in Fig. 2.

The final steps of N-glycosylation occur in the Golgi apparatus, where more sugars are added to the core, elongating the branches by adding Gal or Fuc. Fucoses are added to the proximal GlcNAc by α1-6 Fuc linkage in higher organisms or also α1-3 Fuc linkage in lower Eukaryotes.

The action of many different enzymes in the process of chain elongation leads to a high level of variability in N-linked glycoproteins, which can differ in the length of individual branches as well as in their number, diversity, and number of added sugars (Seeberger, 2022).

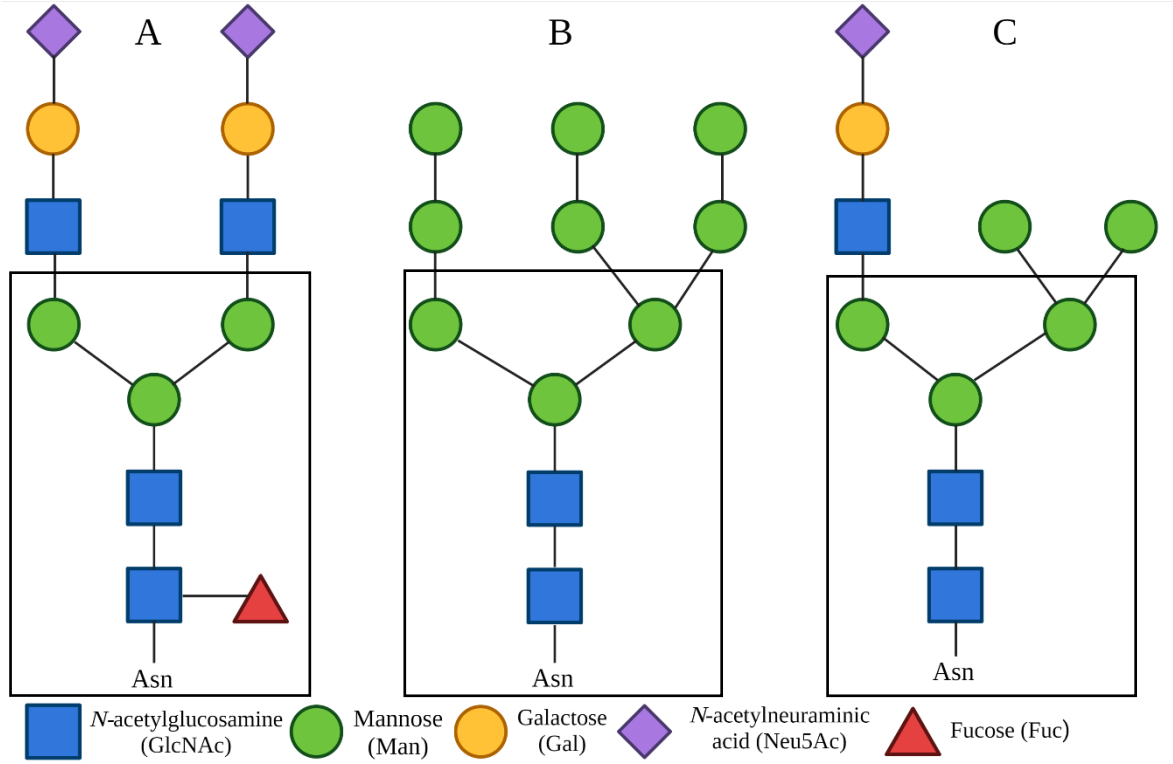


Figure 2: Types of N-glycans. N-glycans in eukaryotic cell connected to the sequon are of three general types: (A) complex, (B) high-mannose, or (C) hybrid. Each type includes a common core: $Man_3GlcNAc_2Asn$, the remaining part of the precursor (black squares). (A) The complex type: any sugars except mannose can be attached to the core, up to five branches. (B) The high-mannose type: Only mannoses can be attached, up to three branches. (C) The hybrid type: on one branch only mannoses can be attached, on the other branch any other sugar can be attached, except mannose. Created with BioRender.com by the author according to sources (Seeberger, 2022; Taylor & Drickamer, 2011).

2.2.2 O-glycosylation

O-glycans are attached to the hydroxyl groups of Ser or Thr. Compared to N-glycans, O-glycans are usually shorter, and their occurrence is not as frequent; however, O-glycans are more variable. O-glycan interactions with proteins are substantially less clear-cut. First, there is no shared glycan core, despite the fact that Ser and Thr (and occasionally hydroxylysine) serve as the only O-glycosylation sites (Rudd et al., 1998; Saldova & Wilkinson, 2020).

There are various O-glycan types, distinguished by the initial sugar that connects them to the protein. These include O-GalNAc (mucin-type), O-GlcNAc, O-Mannose, O-Fucose, O-Glucose, and O-Xylose (Saldova & Wilkinson, 2020). Normally, an N-acetylgalactosaminyltransferase that transfers GalNAc residue to the side chain

of a Ser or a Thr residue initiates O-linked glycosylation in the GA or ER. Afterwards, numerous core structures are produced by stepwise enzymatic elongation by certain transferases. These core structures are then further extended or changed by sialylation, sulfation, acetylation, fucosylation, and polylactosamine-extension (Rudd et al., 1998).

O-GalNAc glycans are also known as mucin-type glycans because clusters of this kind of linkage are thought to be a distinguishing feature of mucins (Sadler, 1984; Spiro, 1973). Nearly every biological process in Eukaryotes, including cell-cell communication, cell adhesion, signal transduction, immunological surveillance, epithelial cell defence, and host-pathogen interactions, depends on O-GalNAc glycans (Seeberger, 2022).

The addition of O-linked GlcNAc to the Ser and Thr of cytoplasmic and nuclear proteins is another kind of O-linked glycosylation. The O-GlcNAc transferase, which is highly conserved, catalyses this addition (Kreppel et al., 1997; Lubas et al., 1997).

The O-linked Mannosyl type of glycosylation in mammalian cells is another peculiar kind. The importance of protein O-Mannosylation in unicellular eukaryotic organisms has been well demonstrated (Gentzsch & Tanner, 1996). Without it, the structure and stiffness of the cell wall may be adversely damaged. O-Mannosyl glycans may be significant players in the cellular recognition stage of mammalian development, in addition to the potential effects of O-Mannosylation on the stability and transport of proteins (Endo, 1999).

The O-Fucose and O-Glucose connections can be considered together. In mature glycoproteins, Fuc can be found either by itself or as the inner part of short oligosaccharides. It has been discovered that Ser is the only residue where Glc may connect, and in some cases, this sugar residue serves as the attachment site for one or two Xyl residues (Harris & Spellman, 1993).

2.2.3 Glycosyltransferases and glycosidases

Glycosyltransferases are enzymes that attach glycans to proteins, while glycosidases are enzymes that cleave parts of glycans. Glycosidases are generally hydrolases, i.e., they catalyse a breakdown reaction where the agent of breakdown is water.

A wide family of enzymes known as glycosyltransferases (GTs) is responsible for the production of oligosaccharides, polysaccharides, and glycoconjugates. Both prokaryotes and eukaryotes have these enzymes, and they often exhibit remarkable selectivity for both the glycosyl donor substrates and the acceptor substrates. The majority of the glycosylation processes that give rise to the variety of glycoproteins found in eukaryotic cells take place

in the GA. The GTs that generate glycosidic bonds by transferring a sugar residue from an activated nucleotide sugar donor to a acceptor molecule are particularly prevalent (Breton et al., 2006).

Enzymes called glycosidases cleave saccharide molecules. They can be divided into exoglycosidases, which exclusively release terminal sugars, and endoglycosidases, which cleave either at the proximal sugar, separating the glycoprotein into its protein and glycan components, or in the middle of the glycan.

In research, endoglycosidases are often used to analyse glycoproteins or glycan structures. Endoglycosidase H (Endo H, EC 3.2.1.96) cleaves between two GlcNAc residues in the core of N-glycan, thus leaving one GlcNAc connected to a protein of hybrid or high-mannose N-glycans; endoglycosidase F (Endo F, EC 3.1.1.96) cleaves at the same site but mainly cleaves high-mannose type and biantennary N-glycans (Seeberger, 2022).

Peptide-N-glycosidase F (PNGase F, EC 3.5.1.52) and Peptide-N-glycosidase A (PNGase A, EC 3.5.1.52) both cleave complex, high-mannose, and hybrid type N-glycans between Asn and proximal GlcNAc. Their main functional difference is that PNGase F might not function if the core is α1-3-fucosylated, whereas PNGase A will function even in these situations. The effectiveness of their use in mammalian glycoproteins is not well understood yet (Seeberger, 2022).

2.3 Ticks

Ticks belong to the class *Arachnida*, subclass *Acarina*, superorder *Parasitiformes*, order *Ixodida*, superfamily *Ixodoidea*. *Ixodoidea* can be separated into three main families: *Ixodidae* (hard ticks), *Argasidae* (soft ticks), and *Nuttalliellidae*, which contain only one rare species *Nuttalliella namaqua*. These families include about 900 species, divided into 18 genera.

Ticks are ectoparasites (external parasites) that feed on the blood of mammals, birds, and rarely reptiles and amphibians to sustain their haematophagic existence. Due to their blood feeding behaviour, they are vectors of various pathogens, Lyme disease being one of the many illnesses that ticks are spreading.

The two main groups of ticks are the *Argasidae*, or soft ticks (Fig. 3), which have a membranous exterior surface, and the *Ixodidae*, or hard ticks (Fig. 4), which have thick outer shells composed of chitin. Hard ticks stay attached to a host's skin and feed for an extended

amount of time, but soft ticks often reside in crevices and emerge quickly to feed several times for shorter period. Adult ticks have eight legs; however, like the majority of other arachnids, lower developmental stages have six legs. Tick bites resemble mosquito bites in appearance; however, they can occasionally bleed or have a bull's eye appearance (Sonenshine & Roe, 2013).

Ticks, as was already noted, are blood-sucking parasites that need to ingest on various animals in order to mature. Each stage of the tick's life cycle – larva, nymph, and adult – feeds by sucking blood. Every stage typically transitions into the following stage after feeding on a different host. Typically, males don't feed; instead, they attach to a host animal just to mate with the female. In the case of females, the adult then continues to feed until it is prepared to lay eggs (Sonenshine & Roe, 2013).

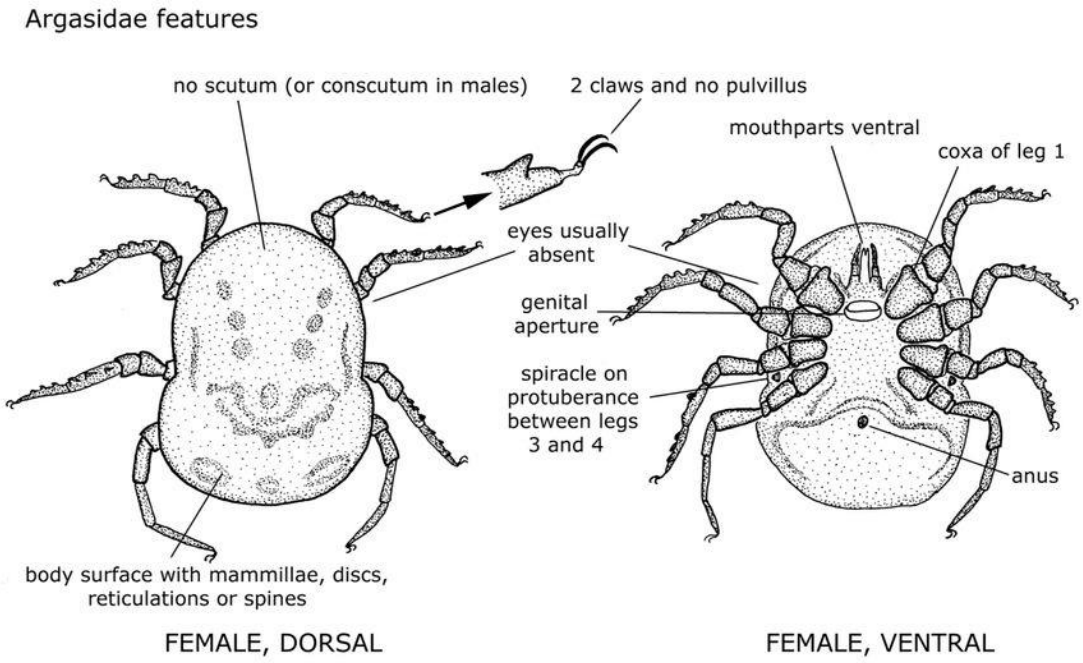


Figure 3: Soft tick membranous exterior surface lacks a scutum. Their body seems to be wrinkled, and males and females have about the same size. The tick's mouthparts are situated on the underside of the body, where they are concealed by the front half of the tick's body (Barker & Walker, 2014).

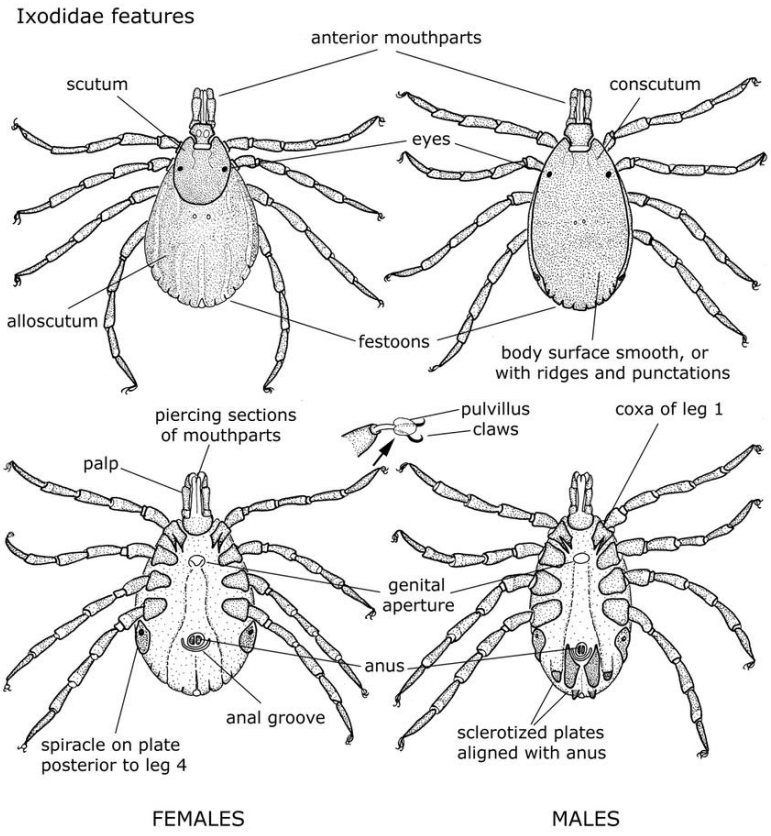


Figure 4: Hard ticks have thick outer chitin shells. Adult females of hard ticks are somewhat bigger than males, and males and females have varied colouring. Scutums, or "plates," are seen on the backs of hard ticks. The mouthparts of hard ticks are also visible when the tick is seen from above (Barker & Walker, 2014).

In this thesis, cells lines originating in three tick species were used; below, these three species will be shortly described.

2.3.1 *Ixodes ricinus* tick (hard European sheep tick)

I. ricinus is one of the most common ticks in Europe, also most abundant in the Czech Republic. Although this tick is indigenous to most of Europe, it is rare or absent in many Mediterranean regions due to the hot, dry weather. Additionally, it has been seen in parts of the Middle East and northern Africa; Fig. 5 shows a schematic representation of the distribution of *I. ricinus* on the world map. It typically lives in cold, somewhat humid, shrubby, or forested regions. It can be found in open regions as well as deciduous and mixed woods where there is a lot of vegetation and moisture (Rovid Spickler, 2022).

The palpi of *I. ricinus* ticks are longer than broad, and they lack eyes. Adults are red-brown, while engorged female ticks are light grey. Males are around 2.5–3 mm long, whereas females are 3–4 mm long before feeding. The females can grow up to 1 cm in length when they are engorged (Rovid Spickler, 2022).

I. ricinus is a three-host tick with a variable lifespan of 2 to 6 years. In the spring or fall, the larvae seek for a host, feed three to six days, then stop feeding and moult. The next spring or autumn, the nymphs search for a host, engorge for three to five days, then drop off and moult. The next spring or autumn, the adults look for a host. Females need anything between 5 and 14 days to finish eating. 500 to 2000 eggs, which hatch by late spring to late fall, are laid by fertilised females. Birds are attacked by immatures, but all three life stages may be hosted by sheep (Capinera, 2008).

2.3.2 *Ixodes scapularis* tick (hard American deer tick)

North America is where *I. scapularis* is the most common, primarily in the eastern U.S. coast, westward into Texas, and northward into Minnesota. Fig. 5 shows a schematic representation of the distribution of *I. scapularis* on the world map. Because the environment in this region is favourable for their primary host, the deer, they are widespread throughout most of the deer-populated parts of the United States. This tick appears to flourish in humid conditions (Rand et al., 2003).

The length of *I. scapularis* is around 3 mm. Males are totally dark brown or black. The eight legs of both sexes are black. The dorsal shield and head of the female are both black, while their belly is dark red. The nymph's body is transparent, and its head is black at the larval stage. The nymph is smaller than the adult and has only one to two millimetres in length (Drummond, 2013).

In the wild, *I. scapularis* has a lifespan of around two years. Typically, females die after laying their eggs, males typically die after mating with one or more females. Springtime egg laying results in summertime egg hatching. (Kiszewski & Spielman, 2002). Engorged larvae leave their initial host after three to five days of feeding, descend to the ground, and stay there for most of the spring and winter. Following this time, larvae change into nymphs, which feed for three to four days on a variety of animals, mostly small mammals. Engorged nymphs moult into adults in the same way that it happens to larvae. From October to the end of the winter, this adult stage is active. While adult male ticks hardly ever feed at all, adult female ticks will feed for five to seven days on their host (Anderson & Magnarelli, 1980; Troughton & Levin, 2007).

2.3.3 *Ornithodoros moubata* tick (soft African tick, African hut tampan)

O. moubata's ventrally positioned head, which is unseen from the dorsal side, and the absence of a chitin dorsal shield (scutum), are significant physical features of the nymph or adult stage. These ticks have a leather-like surface that is wrinkled and covered in tiny

nodules. They also lack eyes, unlike many ticks of the same family. The male is 8 mm long, while the female measures roughly 10 mm (Saari et al., 2019).

Soft ticks have more life stages than hard ticks (egg, larva, 2–5 nymphal stages, and adult). In general, soft ticks are multi-host ticks. After engorgement, females stop feeding quickly and leave the host. Fully engorged females lay eggs mostly in cracks and crevices on the ground. Female ticks lay small numbers of eggs (100–500) up to six times during their lives (Basu & Charles, 2017). Fig. 5 shows a schematic representation of the distribution of *O. moubata* on the world map.

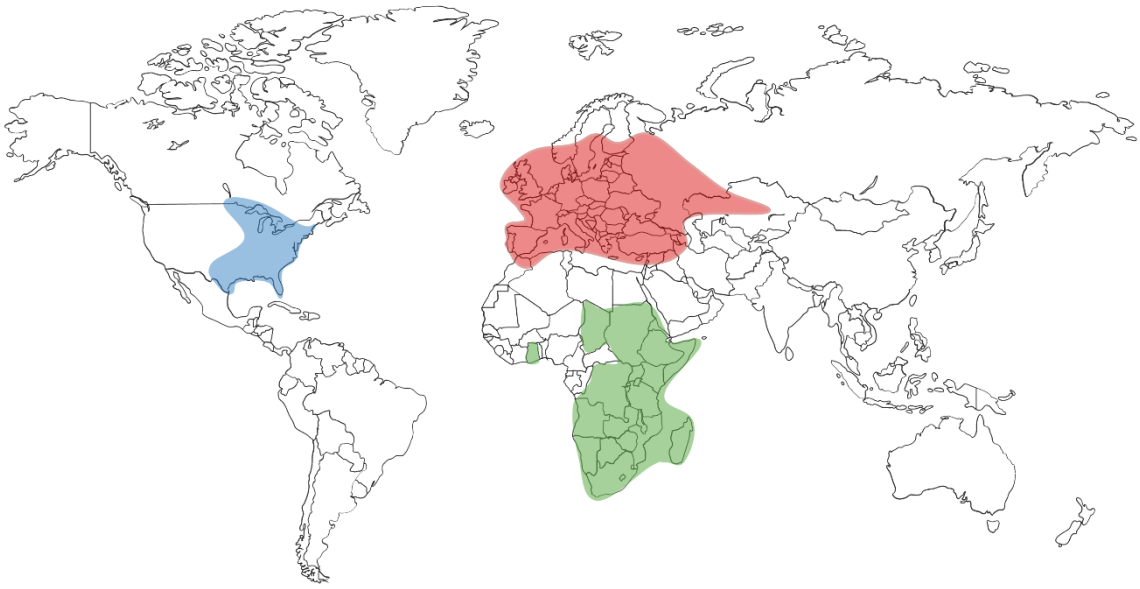


Figure 5: Geographical distribution represented on a world map of I. ricinus − red colour; I. scapularis − blue colour; O. moubata − green colour. Created with BioRender.com by the author according to sources (Frant et al., 2017; Schotthoefer & Frost, 2015).

2.4 Carbohydrate immunogenic epitopes

A glycoprotein's glycan component has the power to trigger cross-reactive immune reactions in some cases. These develop when IgE antibodies respond to antigens that are distinct from the initial immunogen but share a common glycan structure (Commins & Platts-Mills, 2009). Vegetables, including peanuts, celery, potatoes, and tomatoes, have glycoproteins that include these so-called cross-reactive carbohydrate determinants (CCDs). Furthermore, N-glycans found in seafood and insect venoms are also recognised by carbohydrates-specific IgE in addition to glycan groups produced from plants (Van der Veen et al., 1997).

Ishihara et al. (1979) described the first carbohydrate structure of CCDs as an N-linked oligomannose-type chain with β1,2-Xyl and α1-3-Fuc connected to the glycan core, found on bromelain, a thiol protease found in pineapples. This is usually considered to be the beginning point for determining CCDs and their importance (Ishihara et al., 1979). These epitopes are found in insects and plants, but not in mammals.

2.4.1 Sialic acids

Unusual sugars known as sialic acids (Sia), which have a common nine-carbon backbone (Fig. 6), are widely expressed on the cell surfaces of higher Eukaryotes, mainly as terminal components of glycoprotein. They have significant biophysical impacts, such as supplying a significant amount of negative charge repulsion between cells, which can change the biophysical characteristics of cellular interactions (Varki & Gagneux, 2012). In addition, several pathogens and parasites use sialic acid as a shield which protects them from the immune system of their hosts (Nagamune et al., 2004).

The group of Sia contains about fifty members. The most abundant one is N-acetyl-9-O-acetyl neuraminic acid (Neu5,9Ac_2), followed by N-acetylneuraminic acid (Neu5Ac), and the less usual N-glycolylneuraminic acid (Neu5Gc) (Seeberger, 2022).

While Sia is not prevalent in arthropods, it is frequently seen in mammals. The *Drosophila melangoster* fly, which is the most extensively researched model organism, has Sia in the embryo, where it is important for the development of the nervous system (Roth et al., 1992). Come et al. (2014) imply that the presence of Sia has a function in mediating dengue infection in *Aedes aegypti* mosquitoes (Cime-Castillo et al., 2014), thus participating in the transmission of pathogens by arthropods.

It can thus be expected that many animal pathogens have developed to target Sia given its prevalence on cell surfaces. Several different viral hemagglutinins, bacterial adhesins, bacterial toxins, and parasite host-binding proteins are among the Sia-specific binding proteins that are important for infectious process. Similar to Drosophila, ticks are thought to self-produce Sia in the ovary, eggs, and later in the larval stage. However, the precise role of Sia at this stage has not yet been determined (Seeberger, 2022; Vancova et al., 2012).

The most common Sias are Neu5Ac and Neu5Gc. Humans cannot synthesise Neu5Gc because of a unique genetic defect, however Neu5Gc is higly present in mammals. Humans produce polyclonal antibodies against various Neu5Gc-glycans from consumed

meal originating from mammals, such as red meat and dairy. Human immune system then recognizes the molecule as foreign, leading to a production of anti-Neu5Gc antibodies. In several studies was suggested this can be a cause of chronic inflammation-mediated illnesses. It was also reported that Neu5Gc was found in human cancer tissue (Varki, 2010; Yehuda & Padler-Karavani, 2020).

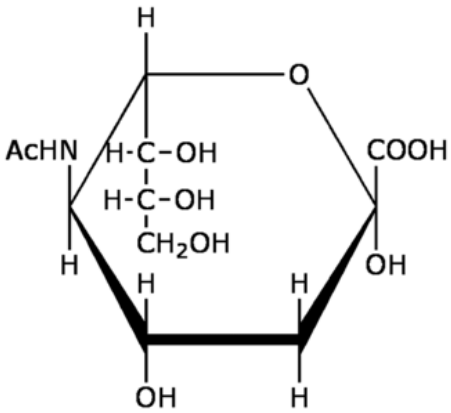


Figure 6: Sialic acid structure. The common nine-carbon backbone (Kamerling & Gerwig, 2006).

2.4.2 α1,3-core fucose

Unlike vertebrates, which mostly add Fuc in a α1,6-core linkage, plants and arthropods are adding Fuc also in α1,3-linkage to the proximal GlcNAc. α1,3-Fuc linked to the proximal GlcNAc causes the development of IgE antibodies against the so-called horseradish peroxidase (HRP) epitope in e.g., humans or rabbits (Seeberger, 2022).

It's interesting to note that, despite ticks producing 1,3-core fucose epitope-containing proteins, a tick bite does not trigger an immunological reaction. Other monosaccharide residues that weaken the immune response or the requirement for the presence of several epitopes to induce an allergic reaction may be the cause (Vancová & Nebesářová, 2015; Věchtová et al., 2018).

2.4.3 Galactose-α-1,3-galactose (αGal epitope)

Another immunogenic epitope linked to the production of IgE antibodies is a terminal galactose-α-1,3-galactose (αGal epitope), the structure of which is show in Fig. 7. The αGal epitope is not expressed in humans, Old World monkeys (*Cercopithecidae*) or apes, but it is extensively expressed in other mammals. αGal deficiency in both humans and Old World monkeys is caused by α1,3-galactosyltransferase inactivation (EC 2.4.1.151) (Macher & Galili, 2008).

With individuals reporting severe allergic responses, anaphylaxis, urticaria (skin rash), or angioedema (swelling caused by an allergic reaction) after consuming red meat, αGal has been suggested as a potential target in food allergy research. In Australia, research done by Nunen and colleagues (2009) documented an association between severe local reactions to tick bites from the tick *I. holocyclus* and red meat allergy and proposed a pathogenic relationship (Nunen et al., 2009). Commins and colleagues (2009) found a connection between sensitization to αGal and allergic responses to cetuximab (an epidermal growth factor receptor inhibiting monoclonal chimeric mouse/human antibody) (Commins et al., 2009). αGal sensitization, tick exposure, red meat (beef, pork, and lamb) allergy, and meat-derived gelatine allergy have been linked, according to research by Mullins and colleagues (Mullins et al., 2012).

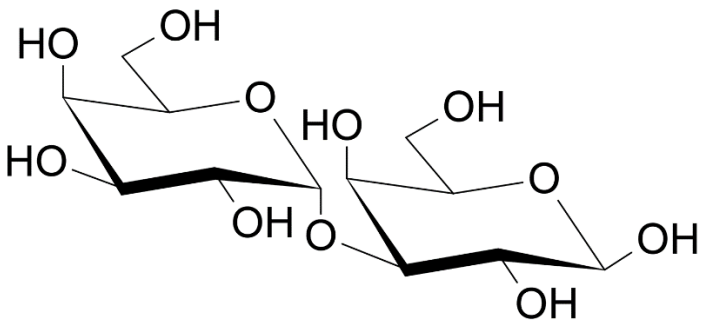


Figure 7: Galactose-α-1,3-galactose (αGal epitope) structure (Anwar et al., 2019)

2.5 Glycoprotein detection methods

A variety of methods are used for the detection of glycoproteins, ranging from simple chemical reaction to, for example, nuclear magnetic resonance (NMR). Because of their complexity, glycoproteins can be classified into several classes depending on the region of the glycoprotein that interests us. For example, determining the protein part (i.e., the amino acid sequence), determining the glycan part (i.e., identifying monosaccharide units and their connection), or determining the attachment of the protein and glycan parts. Therefore, selecting the best approach or procedure for identification is crucial and depends on a variety of parameters. Here, I proceed to outline the principles of procedures that were applied to in this thesis.

2.5.1 SDS-PAGE, lectinoblotting or immunoblotting

The investigation of proteins can be done using electrophoresis, which is a technique used to separate complex protein mixtures (such as those from cells or immunoprecipitations). Additionally, it can be used to purify proteins for later use. The rate at which a protein migrates depends on its charge, size, and shape as well as the size of the gel pores; pore size decreases as acrylamide concentration increases (applies only for PAGE). Through pores in the gel matrix, proteins move during PAGE as a result of an applied electrical field. The molecular size, quantity, and size of subunits of proteins may all be determined using PAGE. The carbohydrate content of glycoproteins may be determined by comparing their mobility before and after deglycosylation (Gallagher, 2006). Proteins are often denatured with SDS to make them linear and give them a similar negative charge/length, allowing them to migrate in the direction of the anode (positive pole) (Raymond & Weintraub, 1959).

Proteins are first separated by mass using SDS-PAGE, and the next step to visualise only the proteins that are glycosylated can be lectin blotting or immune blotting using a suitable carbohydrate-binding antibody. Lectins are carbohydrate-binding proteins that could interact with free carbohydrates, as well as carbohydrates found in glycoproteins and glycolipids, with remarkable specificity (Cao et al., 2013; Mahmood & Yang, 2012). The proteins from the gel are electrophoretically transferred onto membranes (nitrocellulose or PVDF) in both applications of this technique. After that, the membrane is incubated with either lectins (for lectinoblotting) or antibodies (for immunoblotting). Unbound lectin or antibody is washed away, and the bound lectin/antibody is detected. By developing the membrane, the glycoproteins can be identified (Mahmood & Yang, 2012).

These two methods (SDS-PAGE and Western blotting) can then be combined so that the visualised bands on the membrane correspond to the stained bands on the gel. These bands can be cut out of the gel identified by mass spectrometry.

2.5.2 Mass spectrometry

The molecular weight of the particles may be determined using mass spectrometry, an analytical method that separates ionised particles like atoms, molecules, and clusters by measuring the charge-to-mass ratios (m/z).

MALDI (matrix assisted laser desorption/ionization) is one of various ionisation techniques. By encasing a large number of biomolecules in a matrix and evaporating and ionising them with a pulsed laser beam, MALDI enables the study of a wide variety of biomolecules, ranging in size from small metabolite molecules to considerably bigger

proteins. Proteins and DNA are examples of high-molecular-weight compounds that are brittle and prone to fragmentation when ionised using traditional ionisation techniques but can be ionised using MALDI. In order to get high-quality mass spectra, the matrix must be chosen carefully (Domon & Aebersold, 2006).

By applying a high voltage and temperature to an aerosol, the ESI method is a mass spectrometry technique used to evaporate the sample and effectively produce ions in gaseous phase. Since it overcomes the tendency of these molecules to fragment when ionised, it is especially effective in creating ions from macromolecules. Due to its potential to generate multiple-charged ions, ESI differs from MALDI in that it may effectively increase the mass range of the analyser (Ho et al., 2003).

As an ion passes by or collides with a surface, the detector in MS either records the charge generated or the current created. With a scanning instrument, a mass spectrum − record of ions as a function of m/z − is formed by the signal generated in the detector (Ho et al., 2003).

2.5.3 Immunofluorescence

Distribution of the target molecule inside the cells and/or tissues can be performed using fluorescent dyes conjugated to a suitable antibody targeting a specific biomolecule using the sensitive and adaptable technique known as immunofluorescence. Target protein labelling is frequently carried out on immobilised, preserved specimens in order to identify the proteins later using fluorescence microscopy (Wardyn & Jeyasekharan, 2018).

3 Aims of the thesis

- Detection of α1,3-core fucose, sialic acid, and galactose-α-1,3-galactose epitopes in three tick cell lines (IRE/CTVM20, ISE18 and OME/CTVM22) using SDS-PAGE followed by Western blot and later identification using MS.
- Purification of proteins containing α1,3-core fucose, sialic acid or galactose-α-1,3-galactose epitope by using ultrafiltration or magnetic/agarose beads, in three tick cells lines (IRE/CTVM20, ISE18 and OME/CTVM22). Identification of purified proteins by MS.
- Visualization of glycoproteins containing α1,3-core fucose, sialic acid or galactose-α-1,3-galactose epitope by fluorescence microscopy in three tick cells lines (IRE/CTVM20, ISE18, OME/CTVM22).

4 Materials and methods

4.1 Materials

4.1.1 Culture cell lines

IRE/CTVM20 (Bell-Sakyi et al., 2007), ISE18 (Munderloh et al., 1994), OME/CTVM22 (Bell-Sakyi et al., 2009), provided by Dr. Lesley Bell-Sakyi, University of Liverpool
L-15 medium supplemented with 20 % BOFES (bovine fetal serum), 10 % TPB (tryptose phosphate broth), 1 % glutamine, 100 $\mu g \cdot ml^{-1}$ penicillin, 100 $\mu g \cdot ml^{-1}$ streptomycin
1x PBS
4 % SDS in 1x PBS
Protease inhibitor mix HP (Serva, Heidelberg, Germany)

4.1.2 SDS-PAGE, Western blot, developing membranes

SDS-PAGE

ROTIPHORESE®Gel 30, 37.5:1 acrylamide:bisacrylamide (Carl Roth GmbH, Karlshruhe, Germany)
4x separation buffer (1.5 M TRIS, 0.4 % SDS, pH 8.8)
4x stacking buffer (1,0 M TRIS, 0.4 % SDS, pH 6.8)
10 % APS
TEMED (Serva, Heidelberg, Germany)
TCE (Sigma-Aldrich, Saint Louis, MO, USA)
Protein Marker VI (10-245) prestained (AppliChem, Darmstadt, Germany), for detection
PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa (Thermo Fisher Scientific, Waltham, MA, USA), for purification
5x reducing loading buffer (0.2 % bromophenol blue, 40 % glycerol, 8 % SDS, 0.2 M Tris, pH 6.8, 100 mM DTT)
1x Running buffer (0.25 M Tris-base, 1.92 M glycine, 1 % SDS)
Mini-PROTEAN®TGX Stain-Free™ Precast Gels, 8-16 %, 12-well comb (Bio-Rad Hercules, CA, USA), for purification of αGal epitope

Western blot

PVDF Transfer membrane (Thermo Fisher Scientific, Waltham, MA, USA)
1x Blotting buffer (20 % methanol, 25 mM Tris-base, 192 mM glycine)
1x PBS-T (1x PBS with 0.05 % Tween-20)
Blocking solution (5 % non-fat milk powder in 1x PBS-T)

Anti-Peroxidase, made in Rabbit (Sigma-Aldrich, Saint Louis, MO, USA)
MALII, Biotinylated; SNA, Biotinylated; GSL I, isolectin B4, bBotinylated; Alkaline phosphatase anti-rabbit IgG (H+L), made in goat; Alkaline phosphatase streptavidin (Vector Laboratories, Burlingame, CA, USA)

Developing membranes

Alkaline phosphatase developing buffer (100 mM TRIS, 0.1 mM NaCl, 10 mM $MgCl_2$)
BCIP (50 $mg \cdot ml^{-1}$ in 100 % DMSO)
NBT (50 $mg \cdot ml^{-1}$ in 70 % DMSO)

4.1.3 Detection of glycosylated proteins

Blue Silver fixation solution (50 % ethanol, 2 % H_3PO_4)
Blue Silver staining solution (10 % H_3PO_4, 10 % $(NH_4)_2SO_4$, 0.12 % Coomassie G-250, 20 % methanol)

4.1.4 Optimalization of trypsin digestion

Apo-Transferrin bovine (Sigma-Aldrich, Saint Louis, MO, USA)
Peroxidase from horseradish (Boehringer Mannheim, Mannheim, Germany)
Trypsin from bovine pancreas (Sigma-Aldrich, Saint Louis, MO, USA), 1 mg of trypsin was dissolved in 1 ml of 1 mM HCl
1 mM HCl
6 M GdnSCN
8 M Urea
500 mM DTT
1 M IAA
1x PBS

4.1.5 Chloroform-methanol extraction

Methanol (Rotisolv® HPLC Gradient grade)
Chloroform (99+ %, for HPLC, stabilized with ethanol)

4.1.6 Ultrafiltration optimalization

Nanosep 10K Omega (Pall Corporation, Port Washington, New York, USA)
Nanosep 30K Omega (Pall Corporation, Port Washington, New York, USA)
4 % SDS in 1x PBS
100 mM $MnCl_2$, $MgCl_2$, $CaCl_2$ (incubation buffer)

PNGase F (BioLabs, Shirley, NY, USA)

Endo H (Roche Diagnostic GmbH, Basel, Switzerland)

4.1.7 Purification of glycosylated proteins using magnetic/agarose beads from culture cell lines

1x PBS (0.137 M NaCl, 2.68 mM KCl, 2.57 mM Na_2HPO_4, 1.76 mM KH_2PO_4, pH 7.4)

4 % SDS in 1x PBS

Anti-Peroxidase, made in Rabbit (Sigma-Aldrich, Saint Louis, MO, USA)

MALII, Biotinylated; SNA, Biotinylated; GSL I, isolectin B4, Biotinylated; Goat anti-biotin unconjugated (Vector Laboratories, Burlingame, CA, USA)

Dynabeads™ Protein A; Dynabeads™ Protein G (Invitrogen, Waltham, MA, USA)

Streptavidin-agarose (Click Chemistry Tools, Scottsdale, AZ, USA)

4.1.8 Sample preparation for fluorescence microscopy

All solutions used for fluorescence microscopy were filtered through a 0.22 μm filter before use (except for fixation solution).

Trypan blue (0.5 %)

70 % ethanol in DEPC water

Fixation solution: ROTI®Histofix 4 % formaldehyde, phosphate buffer, pH 7.0 (Carl Roth GmbH, Karlshruhe, Germany)

Permeabilization solution: 0.25 % Triton X-100 in 1x PBS

1x PBS

1 % and 3 % BSA in 1x PBS

50 mM NH_4Cl in 1 % BSA

Vectashield®Vibrance™ Antifade Mounting Medium with DAPI; MALII, Biotinylated; SNA, Biotinylated; GSL I, isolectin B4, Biotinylated; DyLight® 549 Streptavidin; DyLight® 488 Anti-Rabbit IgG (H+L), made in horse (Vector Laboratories, Burlingame, CA, USA)

Anti-Peroxidase, made in Rabbit (Sigma-Aldrich, Saint Louis, MO, USA)

4.2 Methods

4.2.1 Tick cell lines

Cell lines were regularly passaged every 14 days. All procedures were done in a laminar flow box. For IRE/CTVM20 (hereinafter IRE20) and ISE18, cells were blown using a sterile pipette and pipetting balloon. Next, approximately 1.5 ml of cell suspension was left in the culture tubes, which was topped up to 4.5 ml with new fresh medium. As for OME/CTVM22 (hereinafter OME22), which is non-adherent, using a sterile pipette and pipettor. The contents of the culture tubes were mixed and approximately 2 ml of cell suspension was left in the culture bottle and topped up to 4.5 ml with new fresh medium. Culture tubes were replaced with new ones every 10 passages.

Cells were aliquoted into 2 ml tubes and centrifuged for 3 min at 1,500 rpm. The supernatant was discarded, and the pellet was subsequently resuspended in 1 ml of 1x PBS. The tubes were centrifuged for 3 min at 1,500 rpm. This washing was performed a total of three times. Then the washed pellet was stored at -20 °C for later use.

The cell lysis was important for later identification or purification and was accomplished as follows: the cell pellet was dissolved in 400 µl of 4 % SDS in 1x PBS. The protease inhibitor cocktail mix was added in a 1:100 ratio. Following that the sample was vortexed for 30 min, sonicated in a water bath for 30 min, and vortexed once more for 30 min. Then the sample was centrifuged for 20 min at 14,000 rpm. Finally, the supernatant was collected and stored at -20 °C for later use. The concentration of the cell supernatant was measured using a NanoPhotometer (Implen, München, Germany) with 4 % SDS in 1x PBS as a blank.

4.2.2 SDS-PAGE, Western blot

SDS-PAGE

For this thesis, 0.75 mm thick 12 % and 15 % polyacrylamide gels with 10 or 15 wells were prepared. Separation and stacking gels were prepared using the relevant chemicals listed in Tab. IV. Moreover, a protein marker was always loaded with the samples at a volume of 5 µl. The electrophoretic apparatus was assembled, and a 100-volt current was set for 90 min.

Table II: List of chemicals and their volumes for preparation of 1 gel (12 % or 15 %) and for preparation of 1 stacking gel.

Chemical	12 % separation	15 % separation	5 % stacking
30 % polyacrylamide	2 ml	2.5 ml	165 µl
4x separation buffer	1.25 ml	1.25 ml	-
4x stacking buffer	-	-	250 µl
ddH_2O	1.7 ml	1.2 ml	575 µl
10 % APS	50 µl	50 µl	10 µl
TEMED	2 µl	2 µl	1 µl
TCE	25 µl	25 µl	5 µl

Western blot

A blotting sandwich was created in a blotting device (Trans-Blot®Turbo™ Transfer System, Bio-Rad, Hercules, CA, USA). Blotting sheets were first soaked in 1x blotting buffer while the PVDF membrane was first activated in methanol. From anode to cathode, the sandwich was constructed as follows: two blotting papers, a PVDF membrane, polyacrylamide gel, and two blotting papers. It was made sure there were no bubbles. The equipment was closed, and the transferring was done at 25 V for 30 min.

After the transfer, the membranes were incubated for 1 h at room temperature or at 4 °C overnight in 5 % skimmed milk in 1x PBS-T. Depending on which epitope required visualization, the membranes were next incubated with primary antibodies for either a whole night at 4 °C or an hour at room temperature. In order to identify α1,3-core fucose and sialic acid, the antibodies were added to a 5 % skimmed milk solution in a ratio of 1:1,000, and in a ratio of 1:500, to identify galactose-α-1,3-galactose. Tab. V includes a list of primary antibodies. The membranes were then washed three times in 1x PBS-T for 10 min each time, at room temperature. Thereafter, 5 % skimmed milk was mixed with the secondary antibodies (listed in Tab. V) in a 1:1,000 ratio and was incubated at room temperature for an hour. The membranes were then washed three more times in 1x PBS-T for 10 min following the incubation. After then, membranes were developed.

Developing membranes

Using an AP developing buffer and solutions of NBT and BCIP, membranes were colorimetrically developed. For every 10 ml, 66 µl of NBT ($c = 50\ mg \cdot ml^{-1}$) were used first

and mixed, then 33 µl of BCIP (c = 50 mg·ml^{-1}) were added and also mixed. Membranes were then developed for about an hour and documented.

Table III: A table of used primary and secondary antibodies in Western Blot.

	α1,3-core fucose	sialic acid	galactose-α-1,3-galactose
primary antibody	anti-HRP made in rabbit	SNA/MALII	GSL-B4
secondary antibody	AP anti-rabbit	AP streptavidin	AP streptavidin

4.2.3 Detection of glycosylated proteins

All cell lysates were suitably diluted after their concentrations were measured such that the sample volume put onto the gel contained approximately 1, 2, 4, 6, 8, and 10 µg of protein for each well. A corresponding amount of 5x loading buffer with DTT was added to the diluted cell lysates. Samples were then denaturated for 10 min at 60 °C.

For information on how gels were prepared, see Chapter 4.2.2. Separate cell lysates were put onto separate gels; thus, each gel represented a single cell line.

The gels were then cut in half, and one half was visualised by Silver Blue staining, while the other half of proteins from gels were transferred to membranes. The membranes were then labelled with appropriate primary and secondary antibodies for each immunogenic epitope (see Chapter 4.2.2).

The Blue Silver staining was performed as follows: the gels were incubated for 20 min in fixation solution. Then they were washed two times in distilled water, for 20 min each time. Finally, the staining solution was added overnight. The gels were then washed in distilled water and documented. The corresponding protein bands that appeared on the membrane after membrane development was cut out of the gel and MS analysis was performed.

4.2.4 Optimalization of trypsin digestion

Before the actual purification of glycoproteins from cell lysates, it was first important to cleave them into peptides. Purification is easier with smaller peptides than whole large proteins. Because the cell samples were very limited, we decided to optimise the trypsin digestion by using model proteins: BT and HRP, as they are commonly available.

To obtain the desired result, the protocol had to be optimised beforehand. The conditions that were optimised were: the amount of digested protein; the concentration

and amount of trypsin; other chemicals for better digestion; the time of digestion; and the final dilution so that the samples were suitable for MS analysis.

The final protocol for cell lysate digestion after optimalization was designed as follows. Before every reaction, each solution was freshly made. Into the volume of 200 µl of cell lysate containing approximately 600 µg of proteins, 220 µl of 6 M GdnSCN were added and the mixture was incubated for 10 min at room temperature. The reaction was incubated for 10 min at 60 °C and 20 µl of 500 mM DTT was added. After that, 20 µl of 1 M IAA was added, and the solution was then left at room temperature in the dark for 30 min. Following that, another 20 µl of 500 mM DTT were added, and the mixture was incubated for an additional 10 min at room temperature. After diluting the sample with 1,120 µl of 1x PBS to a concentration of less than 1 M GdnSCN, 80 µl of trypsin was then added. The digestion was then performed at 37 °C with shaking overnight. Then small aliquot was left as a control.

4.2.5 Chloroform-methanol extraction

Trypsin digestion of culture cell lysates with GdnSCN resulted in some protein precipitation in the samples. Chloroform-methanol extraction was performed to obtain the highest concentration of protein as possible. 400 µl of methanol were added and mixed with 100 µl of the digested protein sample. After that, 100 µl of chloroform were added and mixed. The sample then had 300 µl of ultrapure water added to it, and it became murky. The samples were mixed, and then centrifuged for 2 min at 14,000 rpm. A protein layer appeared between the water and chloroform phases after their separation. Next, the upper water phase was removed. 400 µl of methanol were added and the sample was mixed and centrifuged at 14,000 rpm for 3 min. A protein pellet formed at the microtube's bottom, and the methanol phase was entirely removed. After that, the pellet was set to air dry at room temperature. The pellet was then resuspended in 50 µl of ultrapure water and the concentration was measured using a NanoPhotometer (Implen) and ultrapure water as a blank. The samples were then stored at -20 °C for later purification.

4.2.6 Ultrafiltration optimalization

Following effective trypsin digestion and chloroform-methanol precipitation, 30K MWCO columns were used for ultrafiltration. The steps described in the manual were carried out as follows: First, peptide purification was performed out. The column was filled with 400 µl containing approximately 300 µg of digested sample (small aliquot was left as a control, marked as C2). According to the column manufacturer's guidelines, samples were centrifuged for 3 min at 5,000 rpm and then 100 µl of 1x PBS was added for rinsing.

The samples were once more centrifuged in the same manner as in the previous step. Upon completion, a small aliquot was obtained as an above-membrane control (marked as C3), and for the filtrate (marked C3a).

Lectins (SNA and MALII biotin – sialic acid purification) in ration 1:40 of each lectin were added to the filtrate to perform affinity purification. The samples were incubated at room temperature for 1 h. Following this step, a fresh 30K column was loaded with the samples and centrifuged as in the previous step. The samples were then washed once again with 100 µl of 1x PBS. After washing, a small aliquot was obtained as an above-column control (marked as C4), and a small aliquot was collected for the filtrate (marked as C4a). The next step was to cleave the lectins/antibodies from the proteins. Thus, 100 µl of 4 % SDS in 1x PBS solution was used, and the mixture was incubated for 30 min at 50 °C. The sample-containing column was moved to a clean microtube, where it was incubated. The samples were centrifuged once more following incubation, same as in the earlier procedures. A small aliquot was taken as an above-membrane control (marked as C5), and flow through was the sample with purified proteins (marked as E).

Still, there were issues with insufficient binding between proteins and lectins/antibody. Chlorides ($MnCl_2$, $MgCl_2$, and $CaCl_2$) were added to the incubation as a response. All solutions were made at a concentration of 100 mM and added to the mixture with lectins/antibodies at a final concentration of 1 mM after measuring the volume of the flow-through. In addition, the incubation time was increased from 1 h up to overnight at 4 °C.

After this procedure, protein denaturation of peptide-lectin/antibody complexes led to an additional problem. PNGase F or Endo H were added to release the bound peptides from the peptide-lectin complex. Two separate samples were done simultaneously. Before adding the PNGase F or Endo H, extra washing processes (three times by 50 µl of 1x PBS) were done. The sample-containing column was then transferred into clean microtube and 1 µl of PNGase F was added into one sample and 0.5 µl of Endo H was then added to another sample. Samples were then incubated for an hour at 37 °C. Centrifuged once more in the same manner as before.

Unfortunately, despite efforts to improve affinity purification conditions, it was not able to do so in a way that would produce satisfying results. Because of this, a new technique was used to purify glycosylated proteins from tick cell lines.

4.2.7 Purification of glycosylated proteins using magnetic/agarose beads from culture cell lines

A method using magnetic/agarose beads was selected for the purification of proteins that contain the desired immunogenic epitopes. A 200 µl volume containing 300 µg of protein was taken from the sample after chloroform-methanol extraction. An equivalent volume of incubation buffer with 1 mM chloride ($MnCl_2$, $MgCl_2$ and $CaCl_2$) concentration in 1x PBS was added to the mixture. Depending on whatever immunogenic epitope was isolated, lectins or antibodies were subsequently added to the sample. In case of purification of galactose-α-1,3-galactose, the samples were incubated for 30 min with the lectin and then incubated overnight at 4 °C with anti-biotin unconjugated. The lectins and antibodies and their ratio used for each epitope are listed in the Tab. II. A negative control was made by adding only anti-biotin unconjugated without adding lectins. The samples were then incubated at 4 °C overnight with constant mixing.

Following the completion of this incubation, the magnetic/agarose beads were added. Tab. II shows the amounts and types of beads that were used for each immunogenic epitope. The next step was a 4 h incubation at 4 °C with constant agitation. When the incubation period was over, the washing of the beads differed slightly depending on whether magnetic or agarose beads were used.

Table IV: A schematic table of used antibodies or lectins for purification of proteins that contains desired immunogenic epitopes. For a subsequent step in the purifying process, magnetic or agarose beads are also listed. The quantity of the compounds utilised in µl is indicated by the number included in square brackets.

	α1,3-core fucose	sialic acids	galactose-α-1,3-galactose
antibody [ratio]	anti-HRP made in rabbit [40:1]	X	antibiotin made in goat, unconjugated [40:1]
lectins [ratio]	X	SNA/MALII [80:1]	GSL-B4 [40:1]
magnetic beads [µl]	protein A-MB [150]	X	protein G-MB [150]
agarose beads [µl]	X	streptavidin agarose beads [30]	X

Magnetic/agarose bead washing

The volumes of the beads were calculated on the basis of the binding capacity according to manufacturer. The sample tubes were placed in a magnetic rack (purification using magnetic beads), or the sample tubes were centrifuged for 3 min at 1,000 rpm (purification using agarose beads). Then, the supernatant was collected (labelled as S/N). The beads were then washed three times with 1 ml of an incubation buffer (with chlorides) for 5 min. After washing, the first supernatant was collected and labelled as FT. The sample was divided into two parts for the third wash (500 µl each). After that, one half of the sample was analysed by MS, and the other half was denatured with 100 µl of 4 % SDS in 1x PBS and incubated for an hour at 50 °C. Once the incubation was complete, the obtained supernatant was given the sample F designation. The concentration of each sample (S/N, FT, and F) was then determined using a NanoPhotometer (Implen) (incubation buffer was used as a blank for S/N and FT, and 4 % SDS in 1xPBS was used for F).

This method was chosen because samples that had previously been denatured with 4 % SDS in 1x PBS and subsequently subjected to MS analysis were contaminated and unable to obtain data (PEG contamination occurred). These denatured samples were utilised for SDS-PAGE analysis.

Control

A control was performed for SDS-PAGE. Depending on whether immunogenic epitope was isolated, 10 µl of antibodies or 5 µl of lectins were added to 100 µl of 4 % SDS in 1x PBS. The samples were then incubated at 50 °C for an hour. After concentration measurement with the nanodrop, these samples were labelled as C.

All the samples were then loaded onto 15 % polyacrylamide gels with 15 wells and the proteins were transferred to membranes using Western blot techniques. The membranes were later incubated with appropriate primary and secondary antibodies and finally developed (see chapter 4.2.2).

4.2.8 Preparation of samples for MS analysis

All samples for MS analysis were prepared by Dr. Filip Dyčka by following protocols.

Protein Digestion

The gel was cut into the fractions and all fractions were subjected to in-gel digestion according to a previous protocol (Shevchenko et al., 2007). Proteins binded on microbeads were digested by trypsin as described previously (Chan et al., 2023). Briefly, a reduction and an alkylation of proteins were performed using 10 mM DTT and 55 mM IAA, respectively. Alkylation was stopped after 20 min by addition of 55 mM DTT. Proteins were digested directly on the microbeads by addition of 0.2 µg of trypsin and overnight incubation at 37 °C. The digestion was terminated by addition of formic acid to a final concentration of 2.5 %. The obtained peptide mixture was purified using C18 EmperoTM disks (3M, Saint Paul, MN, USA) (Rappsilber et al., 2007).

Protein Identification using MALDI TOF/TOF

A simple microgradient device was used for reversed-phase liquid chromatography of peptides (Franc et al., 2012; Moravcová et al., 2009). The system was first wetted with 80 % acetonitrile/0.1 % TFA and then equilibrated with 0.1 % TFA. The peptides were aspirated into a microsyringe and loaded on the column and separated by gradient elution in mobile phase with increasing acetonitrile content (v/v): 2 %, 8 %, 16 %, 24 %, 32 %, and 40 % acetonitrile in 0.1 % TFA, 2 µl per each fraction. The eluate was directly deposited onto an AnchorChipTM 800-3384 target plate (Bruker Daltonics) in 0.5 µl aliquots. Immediately after the deposition, 0.5 µl of CHCA was added for crystallization.

Mass spectrometric analyses were performed on an Autoflex Speed MALDI-TOF/TOF mass spectrometer (Bruker Daltonics). Mass spectra were obtained in the positive reflectron ion mode using pulsed extraction with an acceleration voltage of 19 kV, an extraction voltage 16.6 kV, a lens voltage 8.3 kV, a delayed extraction time of 120 ns and reflector voltages 21 kV and 9.7 kV. Resulting spectra were accumulated from up to 500 laser shots. Raw MALDI-MS data were subjected to protein identification to Mascot Server 2.5.0 (Matrix Science) against protein sequences of *I. ricinus* (UniProt, 8.6.2021) and *O. moubata* (UniProt 26.5.2021) supplemented with contaminant database (https://www.thegpm.org/crap/). Database search was performed using the same parameters as described previously (Loginov et al., 2019).

Protein Identification using nano LC-ESI MS/MS

Peptides were dissolved in 40 µl of 3 % acetonitril/0.1 % FA. The peptide analysis was carried out on an UltiMate 3000 RLSCnano system (Thermo Fisher Scientific,

Waltham, MA, USA) coupled on-line to mass spectrometer timsTOF Pro (Bruker Daltonics, Berlin, Germany) in same condition as described previously (Forinová et al., 2021).

Raw MS data were process in MaxQuant software (version 1.6.14.0 and 2.2.0.0) with integrated Andromeda search engine (Cox & Mann, 2008). Tick proteins were identified using databases for *I. ricinus* (UniProt, 8.6.2021), *O. moubata* (UniProt 26.5.2021), and Bos taurus (UniProt, 3.4.2023) supplemented with contaminant database included in MaxQuant software. The default parameters for TIMS-DDA search type and Bruker TIMS instrument were applied. Trypsin/P was set as enzyme allowing up to two missed cleavages in specific digestion mode. The search parameters were set: carbamidomethylation of Cys as fixed modification; Met oxidation was set as variable modifications; the minimum required peptide length was set to five amino acids; precursor ion tolerance was set at 20 ppm and 10 ppm in first and main peptide search, respectively; the mass tolerance for MS/MS fragments was set at 25 ppm. Peptide and protein identification were filtered using a target-decoy approach at a false discovery rate (FDR) of 1 %. Intensity based absolute quantification (iBAQ) was done using the algorithm integrated into MaxQuant software. Data tables obtained in MaxQuant were processed in Perseus software (version 2.0.9.0) (Tyanova & Cox, 2018). Protein contaminants, reverse proteins, protein identified only by sites and proteins identified by only 1 peptide and with score lower than 40 were further exclude from data set. Values of iBAQ were transformed using log2.

4.2.9 Sample preparation for fluorescence microscopy

Following passage, 50 µl of cells were mixed with the same amount of trypan blue, and the mixture was then used to count the cells using a Bürker chamber. The final number of cells in 1 ml of media was calculated by adding the three counts of 25 squares, averaging the result, and multiplying it by 20,000. A sterile pipette was used to measure the volume of medium containing cells, and the total number of cells in the volume was counted. One well should contain 70,000 cells. The well's working volume was 200 µl, and its surface area was 0.3 cm^2.

Three panels, each with 12 wells, were used, one for every cell line and immunogenic epitope. The distribution of the cell lines on a single panel is shown schematically in Table VI. The reusable flexiperm panels (flexiPERM® micro12 reusable, Sarstedt, Nümbrecht, Germany) were sterilised beforehand with 70 % ethanol in DEPC water for 20 min, then washed in deionized water for an additional 20 min, and then let dry in the flow box.

Table V: A schematic representation of one panel how the cell lines were divided. Every epitope was done in triplicate and one well from each cell line was prepared as a negative control (NC).

IRE20	IRE20	ISE18	ISE18	OME22	OME22
IRE20	IRE20 (NC)	ISE18	ISE18 (NC)	OME22	OME22 (NC)

After being divided into panels, the cells were placed in a box with moist paper towels to prevent the media from drying up. The panels were then placed in an incubator (28 °C) for three nights to allow the cells to settle down.

The medium was removed after three days, 150 µl of fixation solution was added, and panels were then incubated for 30 min. The wells were then washed three times for 5 min each with 1x PBS. 150 µl of the permeabilization solution was added after washing, incubated for 15 min, and then rinsed three times with 1x PBS for 5 min each time. Following that, 150 µl of the 3 % BSA blocking solution was added and the panels were incubated at room temperature for 1 h. Following the first incubation period, the cells were incubated twice for 10 min each with 50 mM NH_4Cl in 1 % BSA. The cells were once again rinsed with 1x PBS three times for 5 min.

Primary antibodies were added to 1 % BSA in a ratio of 1:500 and incubated over-night at 4 °C in a volume of 100 µl. One well from every cell line was intentionally left without primary antibodies as a negative control. In Tab. VII, the primary antibodies used with this incubation are listed. After incubation, cells were washed three times for 10 min in 1x PBS and secondary antibodies, listed in Tab. VII, were added to 1 % BSA in a 1:1,000 ratio (100 µl volume) and incubated for 1 h at room temperature. From this step the cells were protected from the light. After incubation with secondary antibodies, the cells were washed again with 1 % BSA three times for 10 min followed by washing with 1x PBS three times for 10 min.

Finally, the silicone wells were removed, and samples were allowed to dry for about 20 min at room temperature. After drying, a few drops of Vectashield mounting medium were added and covered by cover glass, and nail polish was used to seal the edges of the cover glass. The glasses were then stored in the dark at 4 °C until further use.

Table VI: A list of primary and secondary antibodies used in fluorescence microscopy.

	α1,3-core fucose	sialic acid	galactose-α-1,3-galactose	Negative control
primary antibody	anti-HRP made in rabbit	SNA/MALII	GSL-B4	-
secondary antibody	DyLight® 488 Anti-Rabbit IgG	DyLight® 549 Streptavidin	DyLight® 549 Streptavidin	DyLight® Streptavidin 549

The glasses were photographed with Olympus BX-60 fluorescence microscope equipped with Olympus DP71 digital camera with DPC Controller viewing software (Olympus, Tokyo, Japan). The captured images were assembled using the ImageJ program (Schneider et al., 2012) .

5 Results

5.1 Detection of glycosylated proteins, SDS-PAGE and Western blotting

The detection of α1,3-core fucose, Sialic acid, and galactose-α-1,3-galactose was performed by SDS-PAGE with CBB staining and Western blotting with detection using a specific antibody or lectin and secondary labelling followed by alkaline phosphatase signal developing. Each detection in every cell line was performed in triplicates, and below are shown only representative gels or membranes.

5.1.1 α1,3-core fucose

Gels for every cell lines from which protein bands were cut out are shown in Fig. 8. (IRE20) and Fig. 9 (ISE18, OME22), along with the developed membrane with labelled proteins that contain α1,3-core fucose.

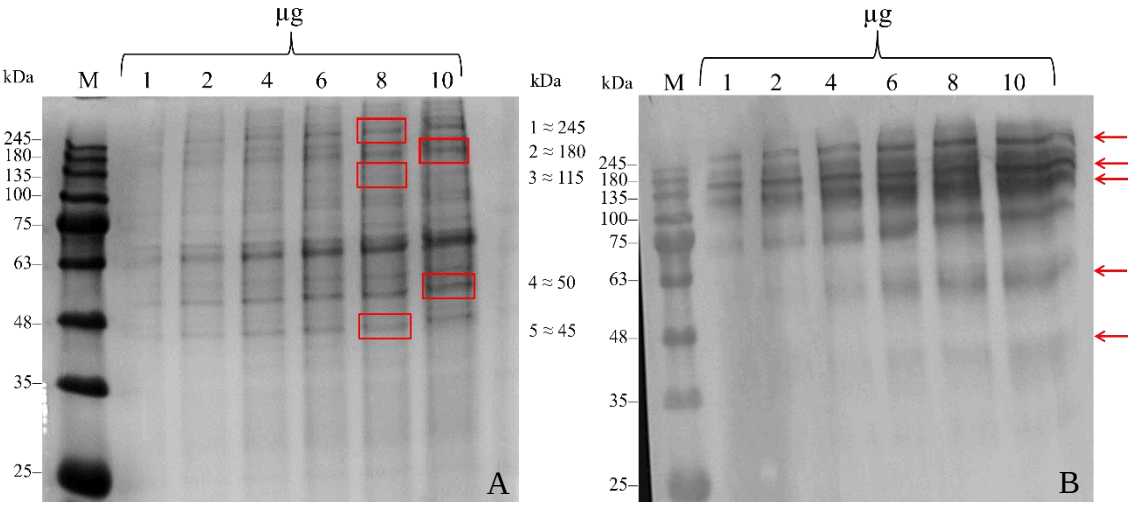


Figure 8: Detection of α1,3-core fucose in the IRE20 cell line lysates. CBB-stained gel (A), and a membrane with labelled proteins that contain α1,3 core fucose (B) are shown. Red squares in the CBB-stained gel show the protein bands that were identified as containing α1,3 core fucose and were cut out for MS analysis; the arrows show the corresponding protein bands on the membrane. The number above shows the amount of protein loaded onto gel. The number on the right side of the picture (A) shows how each band was marked and its approximate mass in kDa. M – protein marker. Full gels are shown in appendix 1.

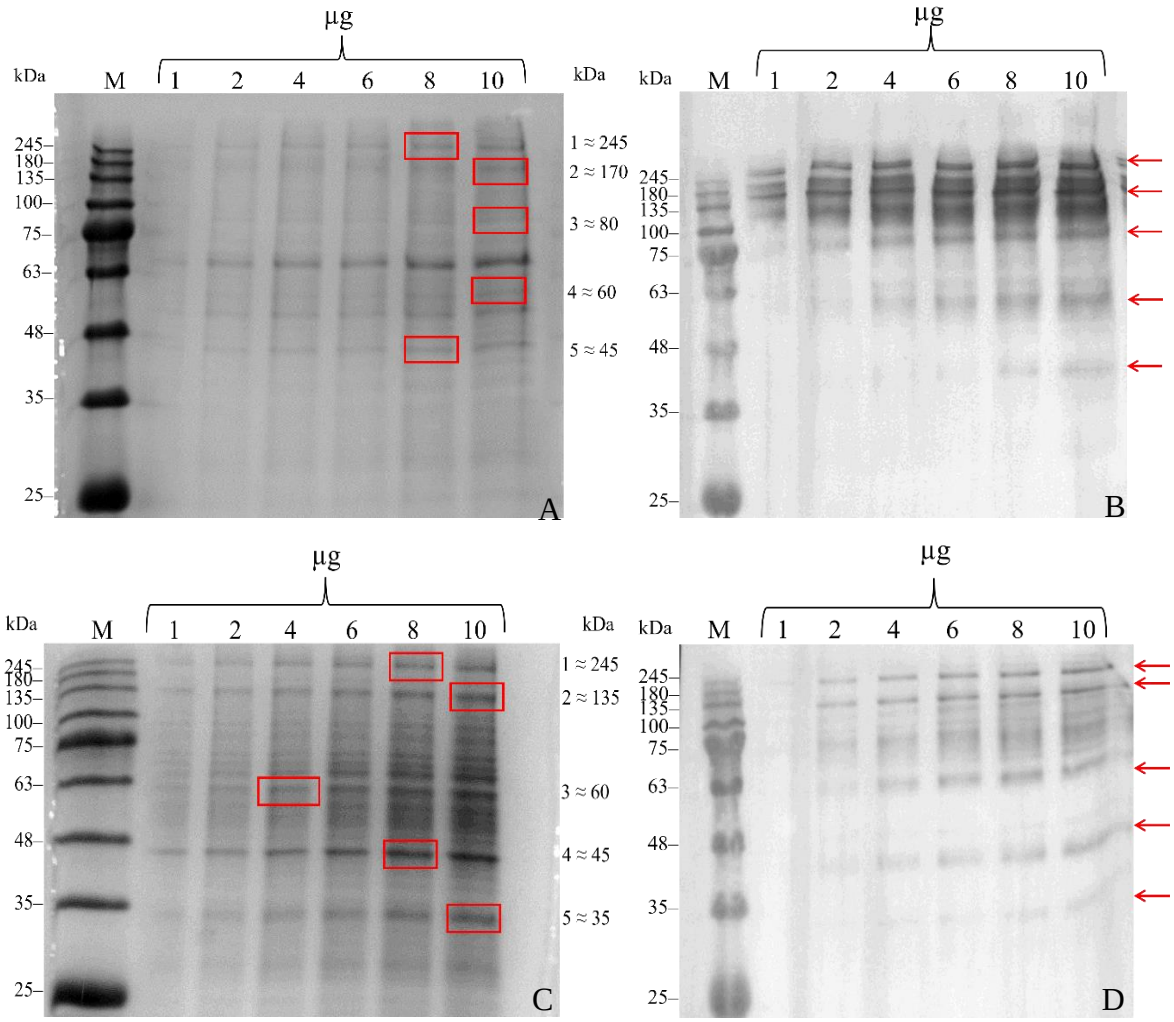


Figure 9: Detection of α1,3-core fucose-containing proteins in the ISE18 (A, B) and OME22 (C, D) cell line lysates. CBB-stained gels (A, C), and membranes with labelled proteins that contain α1,3-core fucose (B, D) are shown. Red squares in the CBB-stained gels show the protein bands that were identified as those containing α1,3-core fucose and were cut out for MS analysis; the arrows show the corresponding bands on the membrane. The numbers above show the amount of protein loaded onto gel. The number on the right side of the picture (A, C) shows how each band was marked and its approximate mass in kDa. M – protein marker. Full gels are shown in appendix 2 and 3.

It can be assumed from Figs. 8 and 9 that every tick cell line contains several fucosylated proteins. Since IRE20 and ISE18 are from the same family *Ixodidae,* the labelled proteins are similar. Also, the cut-out bands have a similar molecular weight (except the band marked as 3). On the other hand, when comparing the hard tick cell lines and the soft tick cell line, differences in the profiles of fucosylated proteins can be seen. Even though the cut-out bands have similar molecular weights, the signals strength differs in between the cell lines. The fucosylated proteins are observed even in samples with only 1 µg of proteins loaded.

5.1.2 Sialic acid

Fig. 10 displays the developed membranes with labelled proteins that contain sialic acid along with the gels for every studied cell line and protein bands cut out for MS identification.

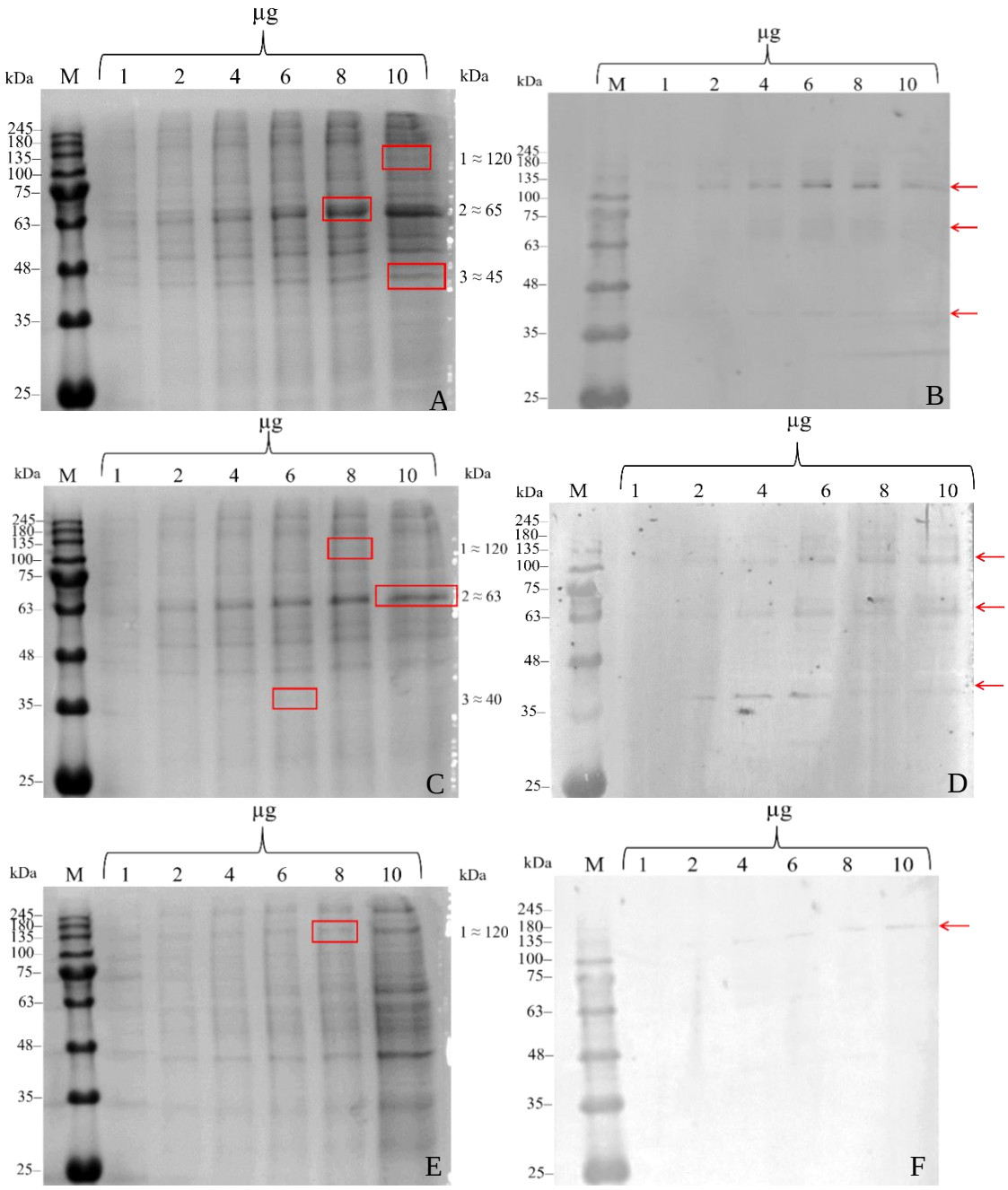


Figure 10: Detection of sialylated proteins in the IRE20 (A, B), ISE18 (C, D), and OME22 (E, F) cell lines lysates. CBB-stained gels (A, C, E), and membranes with labelled proteins that contain Sia (B, D, F) are shown. Red squares in the CBB-stained gels show the protein bands that were identified as containing Sia and were cut out for MS analysis; the arrows show corresponding bands on the membrane. The numbers above show the amount of protein loaded onto gel. The number on the right side of the picture (A, C, E) shows how each band was marked and their approximate mass in kDa. M – protein marker. Full gels are shown in appendix 4, 5, and 6.

From Fig. 10, it can be assumed that every tick cell line contains fewer sialylated proteins than fucosylated proteins. Three bands with a strong signal on the membranes in IRE20 and ISE18 are seen, with almost the same molecular weight in both hard tic cell lines. When comparing the hard tick cell lines and the soft tick cell line, there is a major difference. In OME22, only one strong band of labelled sialylated protein on the membrane can be seen. However, this one band has a similar molecular weight as the bands in hard tick cell lines marked as 1. For the hard tick cell lines, sialylated proteins can be observed even in samples with only 2 μg of proteins loaded, faintly also in 1 μg of proteins. On the other hand, for soft tick cell line sialylated proteins can be seen only in a sample lane with 4 μg of proteins loaded.

5.1.3 Galactose-α-1,3-galactose

Membranes with labelled proteins that contain the αGal epitope and the corresponding final gels for each cell line from which the bands were cut out are shown in Fig. 11.

From the results shown in Fig. 11, it can be seen that almost all proteins labelled for αGal in every tick cell line have a molecular weight that is almost similar. On the contrary to α1,3-core fucose and Sia detection, where differences are seen between hard and soft tick cell line, in the detection of αGal these differences cannot be seen. The labelled protein bands on the membrane can be seen from a concentration of 2 μg of proteins.

In comparison of all the data shown above, it can be found that some proteins with similar molecular weights can contain more than one immunogenic epitope. A protein of approximately 45 kDa is present in the detection of α1,3-core fucose and Sia in the IRE20 (Figs. 8A and 10A) cell line. Also, in the lines ISE18 (Figs. 10C and 11C) and OME22 (Figs. 10E and 11E) there is a protein with a molecular weight of about 120 kDa for the detection of Sia and also of the αGal. Thus, there is some possibility that these proteins could contain more than one immunogenic epitope.

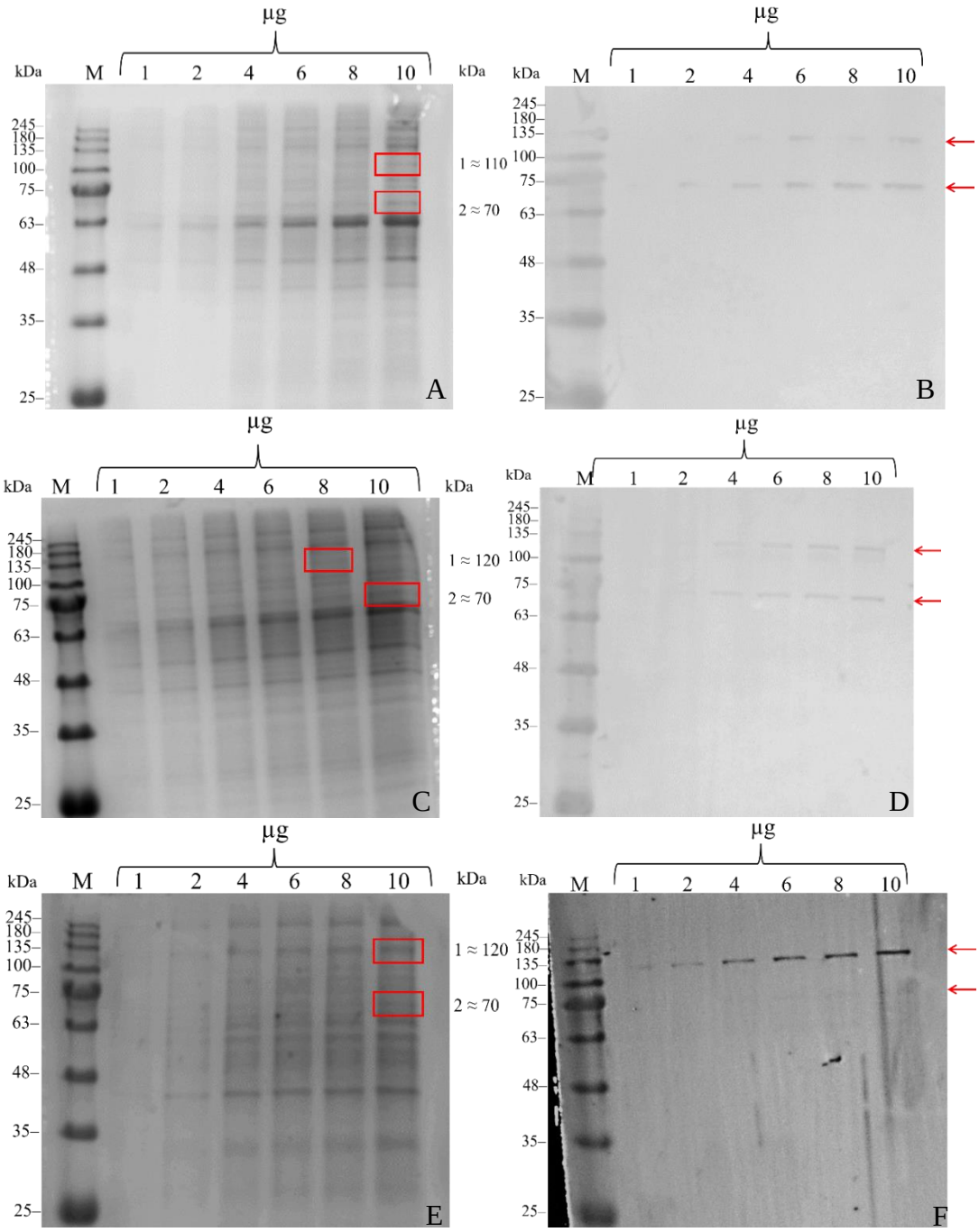


Figure 11: Detection of αGal epitope in the IRE20 (A, B), ISE18 (C, D), and OME22 (E, F) cell lines. CBB-stained gels (A, C, E), and membranes with labelled proteins that contain αGal (B, D, F) are shown. Red squares in the CBB-stained gels show the protein bands that were identified as containing αGal and were cut out for MS analysis; the arrows show corresponding bands on the membrane. The numbers above show the amount of protein loaded onto gel. The number on the right side of the picture (A, C, E) shows how each band was marked and their approximate mass in kDa. M – protein marker. Full gels are shown in appendix 7, 8, and 9.

5.2 Identification of glycosylated proteins in protein bands, MS data

Proteins detected by imuno/lectinoblotting were cut out from CBB stained gels and analysed by MALDI-TOF/TOF MS. Each analysis was done in triplicates. The MS data from each repeat were compared and only the proteins that occurred in all three replicates (in some cases only two) were chosen. For each band, also the molecular weight was considered. Proteins with the highest scores were chosen from the three repeats.

Subsequently, because the MS data of each sample comprised hundreds of proteins, they were categorised as follows: From the highest to the lowest intensity, the top 10 % of the highest protein intensities were chosen. Moreover, contaminants like actin and tubulin that could have come from the media during cell development were taken out. The first two proteins that were still present after this sorting were also checked to determine if they come from *Ixodida* and if they contained putative N-glycosylation sites. An online program NetNGlyc – 1.0 was used to find the N-glycosylation sites (Gupta & Brunak, 2002).

α1,3-core fucose

Proteins that were identified by MS analysis in IRE20, ISE18, and OME22 cell lines probably containing α1,3-core fucose based on immunoblotting are listed in Table VII, VIII, and IX, respectively.

Sialic acid

Proteins that were identified by MS analysis in IRE20, ISE18, and OME22 cell lines probably containing Sia based on lectinoblotting are listed in Table X, XI, and XII, respectively.

Galactose-α-1,3-galactose

Proteins that were identified by MS analysis for IRE20, ISE18, and OME22 cell lines probably containing αGal based on lectinoblotting are listed in Table XIII, XIV, and XV, respectively.

In the following tables, lists of proteins identified by MS in protein bands cut out from SDS-PAGE gels are shown. The information about Uniprot number and protein molecular weight is shown as well. The column with N-glycosylation sites is showing the sequons identified in the sequences of these proteins predicted by NetNGlyc – 1.0, the number represents the number of the first amino acid in the amino acid sequence. The star in the names of some proteins means that these proteins were identified only in two replicates.

Table VII: List of proteins that were identified by MS in protein bands cut out from SDS-PAGE gels and corresponded to bands shown on membranes as those containing α1,3-core fucose in the IRE20 cell line. The protein band numbers correspond with the bands cut out of the gel represented in Fig. 8A.

Protein band	Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
1	V5H3D0	Putative cation-independent mannose-6-phosphate receptor*	7.5	224.0	98 (NTT), 347 (NIT), 507 (NRT), 663 (NLT), 697 (NLT), 706 (NLT), 835 (NGT), 877 (NYT), 1163 (NDT), 1736 (NGS), 1963 (NVT)
	A0A147BHZ6	Putative mannose-6-phosphate/insulin receptor type ii*	323.3	255.3	347 (NIT), 507 (NRT), 663 (NLT), 697 (NLT), 706 (NLT), 835 (NGT), 877 (NYT), 1163 (NDT), 1736 (NGS), 1963 (NVT)
2	A0A0K8RE48	Putative prosaposin	323.3	148.8	87 (NET), 300 (NAT), 467 (NMT), 554 (NST), 657 (NAT), 860 (NGT), 1147 (NVT)
	A0A6B0VEP9	Putative saposin	39.9	126.6	87 (NET), 300 (NAT), 467 (NMT), 944 (NVT)
3	A0A0K8R8N9	Heat shock protein 83	323.3	84.2	48 (NSS), 286 (NKT), 392 (NIS)
	A0A131XQ15	Putative translation elongation factor 2	323.3	94.3	3 (NFT), 257 (NPT)
4	A0A7D5BM75	Elongation factor 1-alpha	241.0	50.7	284 (NIT)
	A0A6B0VEL5	Putative formyltetrahydrofolate dehydrogenase (Fragment)	323.3	61.1	488 (NVT)
5	A0A131Y342	Adenosylhomocysteinase*	270.5	47.8	183 (NDS)
	A0A131Y2X7	Putative metallopeptidase*	233.4	41.5	182 (NPT)

Table VIII: List of proteins that were identified by MS in protein bands cut out from SDS-PAGE gels and corresponded to bands shown on membranes as those containing α1,3-core fucose in the ISE18 cell line. The protein band numbers correspond with the bands cut out of the gel represented in Fig. 9A.

Protein band	Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
1	A0A147BHZ6	Putative mannose-6-phosphate/insulin receptor type ii*	323.3	255.3	347 (NIT), 507 (NRT), 663 (NLT), 697 (NLT), 706 (NLT), 835 (NGT), 877 (NYT), 984 (NVT), 1163 (NDT), 1736 (NGS), 1963 (NVT)
	V5HC54	Putative beta-spectrin*	323.3	277.9	626 (NRS), 2376 (NLT)
2	A0A0K8RE48	Putative prosaposin	323.3	148.8	87 (NET), 300 (NAT), 467 (NMT), 544 (NST), 657 (NAT), 860 (NGT), 1147 (NVT)
	V5IFM5	Putative clip-associating protein 1 (Fragment)	49.8	161.2	153 (NLS), 458 (NLS)
3	A0A131XQ15	Putative translation elongation factor 2	323.3	94.3	3 (NFT), 257 (NPT)
	A0A0K8RGG2	Heat shock protein 83	166.3	91.2	454 (NVS)
4	A0A6B0VEL5	Putative formyltetrahydrofolate dehy-drogenase (Fragment)	323.3	61.1	488 (NVT)
	A0A7D5BM75	Elongation factor 1-alpha	241.0	50.7	284 (NIT)
5	A0A131Y0R0	Citrate synthase (Fragment)	61.7	50.2	17 (NIS), 261 (NVS)
	A0A131Y342	Adenosylhomocysteinase	270.5	47.8	183 (NDS)

Table IX: List of proteins that were identified by MS in protein bands cut out from SDS-PAGE gels and corresponded to bands shown on membranes as those containing α1,3-core fucose in the OME22 cell line. The protein band numbers correspond with the bands cut out of the gel represented in Fig. 9C.

Protein band	Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
1	A0A1Z5L4Q7	Alpha-mannosidase (Fragment)	251.3	90.1	376 (NST), 442 (NGT), 616 (NGT)
	A0A1Z5L522	E3 ubiquitin-protein ligase TRIP12 (Fragment)	131.7	106.7	176 (NVT), 227 (NYT), 378 (NVT), 431 (NGT), 458 (NLT), 573 (NCS), 612 (NIT), 674 (NIT), 698 (NLT), 774 (NVT)
2	A0A1Z5L522	E3 ubiquitin-protein ligase TRIP12 (Fragment)	131.7	106.7	176 (NVT), 227 (NYT), 378 (NVT), 431 (NGT), 458 (NLT), 573 (NCS), 612 (NIT), 674 (NIT), 698 (NLT), 774 (NVT)
3	A0A1Z5L916	Beta-galactosidase (Fragment)	31.4	72.9	24 (NFT), 208 (NVT), 243 (NAT), 491 (NVT)
4	A0A1Z5L6C8	Peptidase A1 domain-containing protein (Fragment)	69.1	42.5	125 (NGT), 238 (NIT)
	A0A1Z5KXU7	Uncharacterized protein (Fragment)*	65.5	41.9	123 (NMS), 366 (NLT)
5	Q6U8A8	Serine protease-like protein	90.0	33.2	121 (NYT)
	A0A1Z5L7X9	Pept_C1 domain-containing protein (Fragment)*	43.2	26.0	40 (NTT), 137 (NIS)

From the results in Tab. VII, VIII, and IX can be seen that not all molecular masses that are assumed from gels corresponds with the molecular masses of proteins that are listed in these tables. On the other hand, similar proteins are identified between IRE20 and ISE18. However, when comparing hard tick cell lines and soft tick cell line not a single one protein matches.

Table X: List of proteins that were identified by MS in protein bands cut out from SDS-PAGE gels and corresponded to bands shown on membranes as those containing Sia in the IRE20 cell line. The protein band numbers correspond with the bands cut out of the gel represented in Fig. 10A.

Protein band	Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
1	A0A147BBM6	Protein-tyrosine-phosphatase (Fragment)*	29.7	139.5	381 (NVS), 501 (NFS)
2	A0A0K8R9H1	Putative heat shock cognate 70 isoform 1	323.3	71.1	64 (NNT), 417 (NTT), 487 (NVS)
	V5I2P5	Putative ca2+-binding actin-bundling protein (Fragment)	56.6	66.8	170 (NLT), 266 (NFT), 409 (NWT)
3	A0A0K8RID7	Elongation factor 1-alpha*	23.5	49.2	NONE
	A0A6B0VEL5	Putative formyltetrahydrofolate dehydrogenase (Fragment)*	45.5	61.1	488 (NVT)

Table XI: List of proteins that were identified by MS in protein bands cut out from SDS-PAGE gels and corresponded to bands shown on membranes as those containing Sia in the ISE18 cell line. The protein band numbers correspond with the bands cut out of the gel represented in Fig. 10C.

Protein band	Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
1	A0A131XQ15	Putative translation elongation factor 2	109.1	94.3	3 (NFT)
	V5HJ32	Putative scinderin like b (Fragment)*	16.3	74.3	406 (NES)
	A0A0K8RGG2	Heat shock protein 83	86.2	91.2	454 (NVS)
2	A0A0K8R9H1	Putative heat shock cognate 70 isoform 1	276.9	71.1	64 (NNT), 417 (NTT), 487 (NVS)
	A0A131XRZ5	Putative polyubiquitin (Fragment)	16.8	64.2	NONE
	V5I2P5	Putative ca2+-binding actin-bundling protein (Fragment)	45.1	66.8	170 (NLT), 266 (NFT), 409 (NWT)
3	A0A0K8R4R6	Isocitrate dehydrogenase*	14.6	46.1	95 (NTS)

Table XII: List of proteins that were identified by MS in protein bands cut out from SDS-PAGE gels and corresponded to bands shown on membranes as those containing Sia in the OME22 cell line. The protein band numbers correspond with the bands cut out of the gel represented in Fig. 10E.

Protein band	Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
1	A0A1Z5L7U0	Uncharacterized protein (Fragment)	162.6	84.9	131 (NGT), 252 (NVT), 425 (NIS)
	A0A1Z5KVK9	Heat shock protein 83 (Fragment)	323.3	90.4	450 (NVS)

Comparing Tab. X and XI where both are the result for hard tick cell lines only in one band (labelled as 2) the two protein matches. Comparing Tab. XI and XI even though one is the results for hard tick cell line and the other one is for soft tick cell line one protein almost matches (Heat shock protein 83). For ISE18 cell line it is a whole protein but for OME22 is a fragment of the protein.

Table XIII: List of proteins that were identified by MS in protein bands cut out from SDS-PAGE gels and corresponded to bands shown on membranes as those containing αGal in the IRE20 cell line. The protein band numbers correspond with the bands cut out of the gel represented in Fig. 11A.

Protein band	Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
1	A0A131XSJ1	Ubiquitin-activating enzyme E1 (Fragment)*	123.8	112.6	6 (NGS)
2	A0A131XTW4	Putative thimet oligopeptidase*	72.0	84.3	67 (NKT)
	A0A0K8R9H1	Putative heat shock cognate 70 isoform 1	276.9	71.1	64 (NNT), 417 (NTT), 487 (NVS)

Table XIV: List of proteins that were identified by MS in protein bands cut out from SDS-PAGE gels and corresponded to bands shown on membranes as those containing αGal in the ISE18 cell line. The protein band numbers correspond with the bands cut out of the gel represented in Fig. 10A.

Protein band	Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
1	V5HVY0	Putative transcription factor nfat subunit nf90	218.8	98.8	NONE
	A0A131XSJ1	Ubiquitin-activating enzyme E1 (Fragment)*	123.8	112.6	6 (NGS)
2	A0A0K8RIK6	Putative heat shock 70 kDa protein	238.7	74.7	49 (NKS)
	A0A147BUS7	H (+)-transporting two-sector ATPase*	144.1	68.1	180 (NYT), 298 (NVS)
	A0A0K8R9H1	Putative heat shock cognate 70 isoform 1*	312.0	71.1	64 (NNT), 417 (NTT), 487 (NVS)

Table XV: List of proteins that were identified by MS in protein bands cut out from SDS-PAGE gels and corresponded to bands shown on membranes as those containing αGal in the OME22 cell line. The protein band numbers correspond with the bands cut out of the gel represented in Fig. 10A.

Protein band	Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
1	A0A1Z5KVK9	Heat shock protein 83 (Fragment)*	87.1	90.4	107 (NAS)
	A0A1Z5L7U0	Uncharacterized protein (Fragment)	51.4	84.9	131 (NGT), 252 (NVT), 425 (NIS)
	A0A1Z5L522	E3 ubiquitin-protein ligase TRIP12 (Fragment)*	44.7	106.7	176 (NVT), 227 (NYT), 378 (NVT), 431 (NGT), 458 (NLT), 573 (NCS), 612 (NIT), 674 (NIT), 698 (NLT), 774 (NVT)
2	A0A1Z5KW47	Uncharacterized protein (Fragment)	323.3	72.3	NONE
	A0A1Z5L577	Cathepsin L (Fragment)	119.4	83.7	254 (NVT), 303 (NKS), 449 (NSS)
	A0A1Z5L916	Beta-galactosidase (Fragment)	78.6	72.9	24 (NFT), 208 (NVT), 243 (NAT), 491 (NVT)

Comparing the Tabs. XIII and XIV from the individual bands, there is at least one protein matching. On the other hand, comparing tables for hard tick (Tabs. XII, XIV) and soft (Tab. XV) tick cell lines does not show even one protein match even though on the gels, the protein bands have similar molecular weights.

When comparing proteins identified for each epitope in IRE20, one protein occurs in the identification of α1,3-core fucose and Sia and another protein in identification of Sia and αGal. Comparing ISE18 cell line for every epitope, three protein matches occur, two in the identification of α1,3-core fucose and Sia and one in the identification of Sia and αGal. Finally, in OME22 cell line three uncharacterised proteins were identified (only cell line when this happens) and four matches. In the identification of α1,3-core fucose and αGal two proteins match and in the identification of Sia and αGal also two proteins match.

5.3 Ultrafiltration optimalization

Ultrafiltration affinity purification was the first method used to purify proteins containing α1,3-core fucose, Sia or αGal for better MS identification. In this section, only some results of optimization are shown. Fig. 12 represents the results from ultrafiltration before and after adding divalent cations into the incubation mixtures with lectins. Both results were performed with 30K MWCO columns, incubation overnight with at 4 °C, eluted with 4 % SDS.

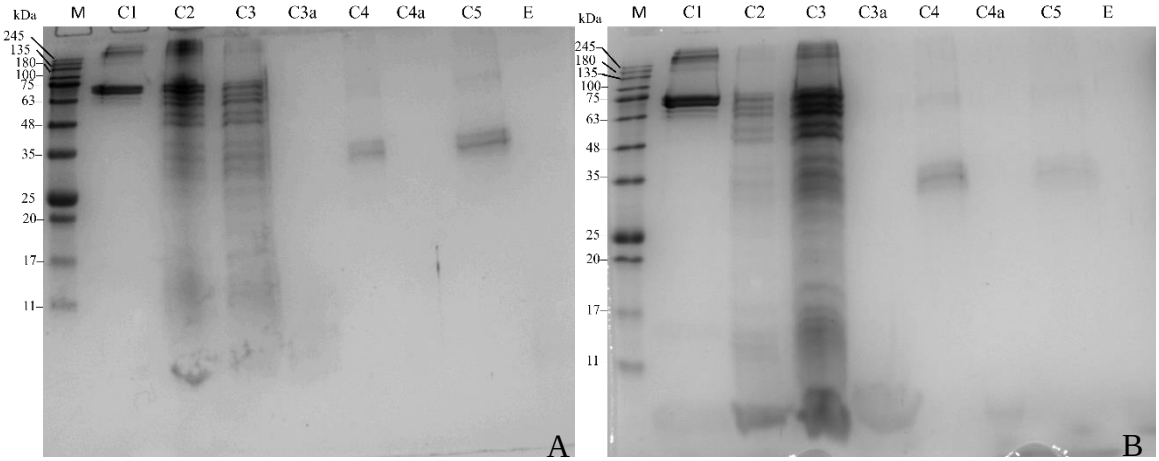


Figure 12: Sia affinity ultrafiltration of BT on CBB stained gel. Result before (A) and after (B) addition of divalent cations into the incubation mixtures with lectins. M – marker; C1 – BT (4 ug); C2 – BT after digestion with trypsin; C3 – above-membrane control after peptides purification; C3a – ultrafiltrate after peptides purification; C4 – above-membrane control after affinity purification; C4a – ultrafiltrate control after affinity ultrafiltration; C5 – above membrane control after denaturation with 4 % SDS in 1x PBS; E – ultrafiltrate after denaturation (sample with purified proteins).

As seen in Fig. 12, incubation with divalent cations did not enhance the binding of lectins to glycopeptides, since the wells labelled as E are both empty. However, divalent cations were used in further experiments as they should enhance the lectin binding to glycans, and they did not affect the purification in unwanted way.

Fig. 13 represent the result of adding Endo H and PNGase F to cleave peptides from the peptide-lectin complex, which should "elute" the peptide part of the bound glycopeptides.

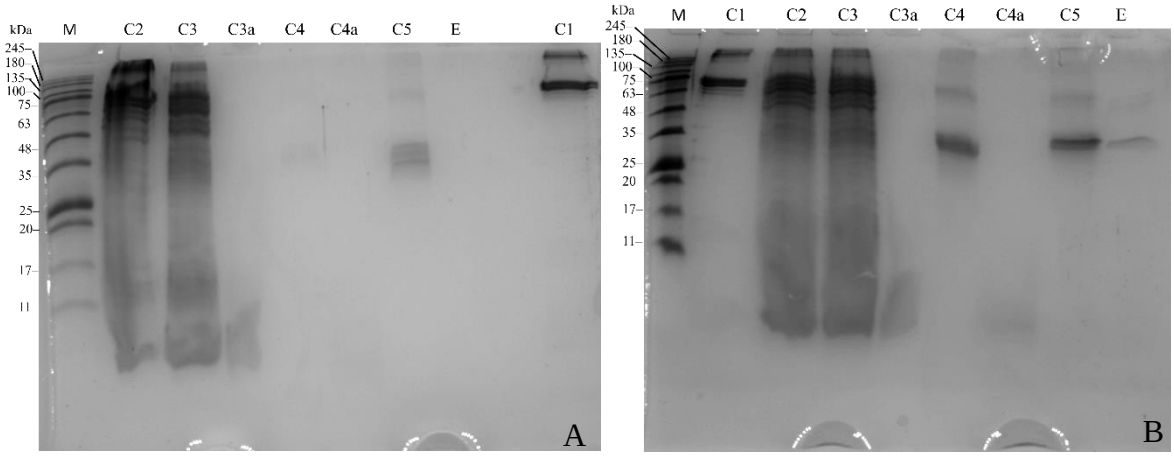


Figure 13: Result of Sia affinity ultrafiltration for BT on CBB stained gel. Result after adding Endo H (A) or PNGase F (B). Both experiments were performed using 30K MWCO columns, incubation overnight with at 4 °C. M – marker; C1 – BT (4 ug); C2 – BT after digestion with trypsin; C3 – above-membrane control after peptides purification; C3a – ultrafiltrate control after peptides purification; C4 – above-membrane control after affinity purification; C4a – ultrafiltrate control after affinity ultrafiltration; C5 – above membrane control after Endo H (A) or PNGase F (B) treatment; E – ultrafiltrate after denaturation (sample with purified proteins)

A protein band occurred in Fig. 13B in the well labelled E. As PNGase F has a molecular weight of 36 kDa, since it was assumed, it is the enzyme. However, when the protein band was sent for MS analysis, no proteins were identified. As optimization of affinity ultrafiltration did not provide any results, the purification using magnetic/agarose beads was chosen as a new approach.

5.4 Purification of glycosylated proteins using magnetic/agarose beads, SDS-PAGE, and Western blotting

After unsuccessful affinity ultrafiltration a new method of purifying proteins by using magnetic/agarose beads was performed. Every purification was done in triplicates. Also, for galactose-α-1,3-galactosea negative control was prepared alongside the control. Only representative membranes or gels for each cell line and immunogenic epitope are shown below.

5.4.1 α1,3-core fucose

Results obtained from purification of α1,3-core fucose for IRE20 (Fig. 14), ISE18 and OME22 (Fig. 15) are shown in the following figures on SDS-PAGE gels together with developed membranes.

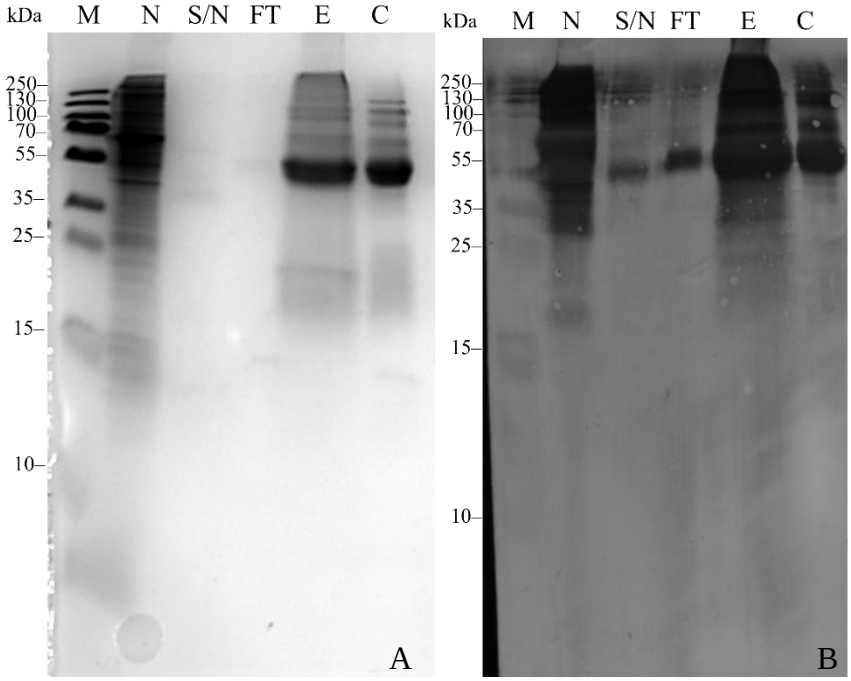


Figure 14: Purification of α1,3-core fucose from IRE20 cell line. CBB stained gel (A), membrane (B) with labelled proteins that contain α1,3-core fucose are shown. M – marker; N – cell lysate of IRE20 cell line (10 μg); S/N – supernatant after incubation with magnetic beads; FT – flow through; E – elution sample after denaturation of magnetic beads with 4 % SDS; C – control.

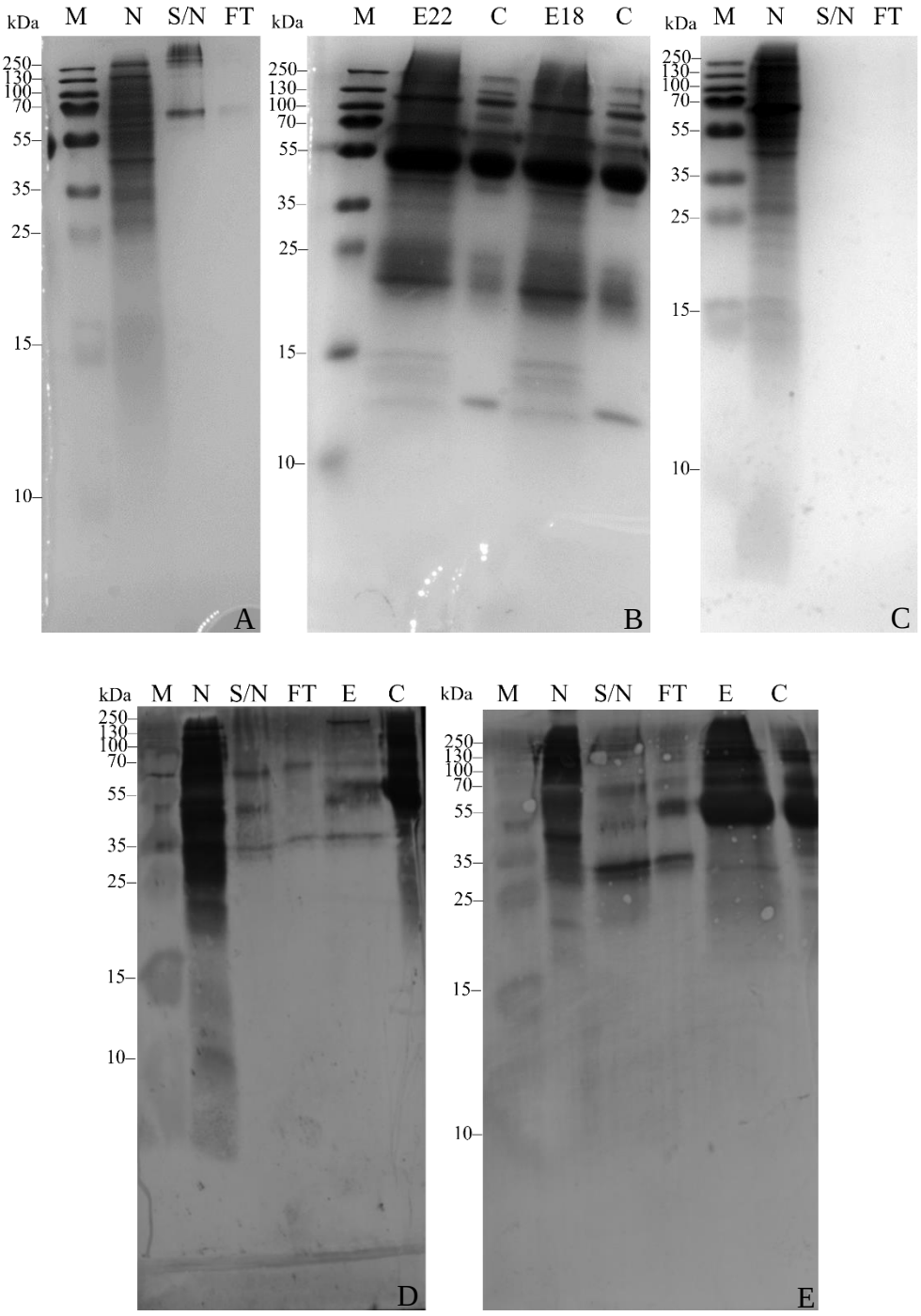


Figure 15: Result of the purification of α1,3-core fucose from ISE18 (B, C, E) and OME22 (A, D) cell line. CBB stained gels are shown (A, B, C), membranes with labelled proteins that contain α1,3-core fucose are shown (D, E). M – marker; N – cell lysate of OME22 or ISE18 cell line (10 μg); S/N – supernatant after incubation with magnetic beads; FT – flow through; E – elution sample after denaturation of magnetic beads with 4 % SDS; E22 (B) – elution for OME22; E18 (B) – elution for ISE18; C – control.

The results shown in Figs. 14 and 15 show that for IRE20 and ISE18 cell lines the purified samples look similar. Similar protein bands on gels and membranes are visible. On the contrary, protein profiles from hard tick cell lines and soft tick cell line look similar

on the gel; however, signal strength varies after immunoblotting on the membranes between the hard and soft tick cell lines. Also, some proteins with similar molecular weights appeared occurred in the FT samples.

5.4.2 Sialic acid

Fig. 16 shows the representative CBB-stained gels and the developed membranes for every cell line, after the purification of sialylated glycoproteins.

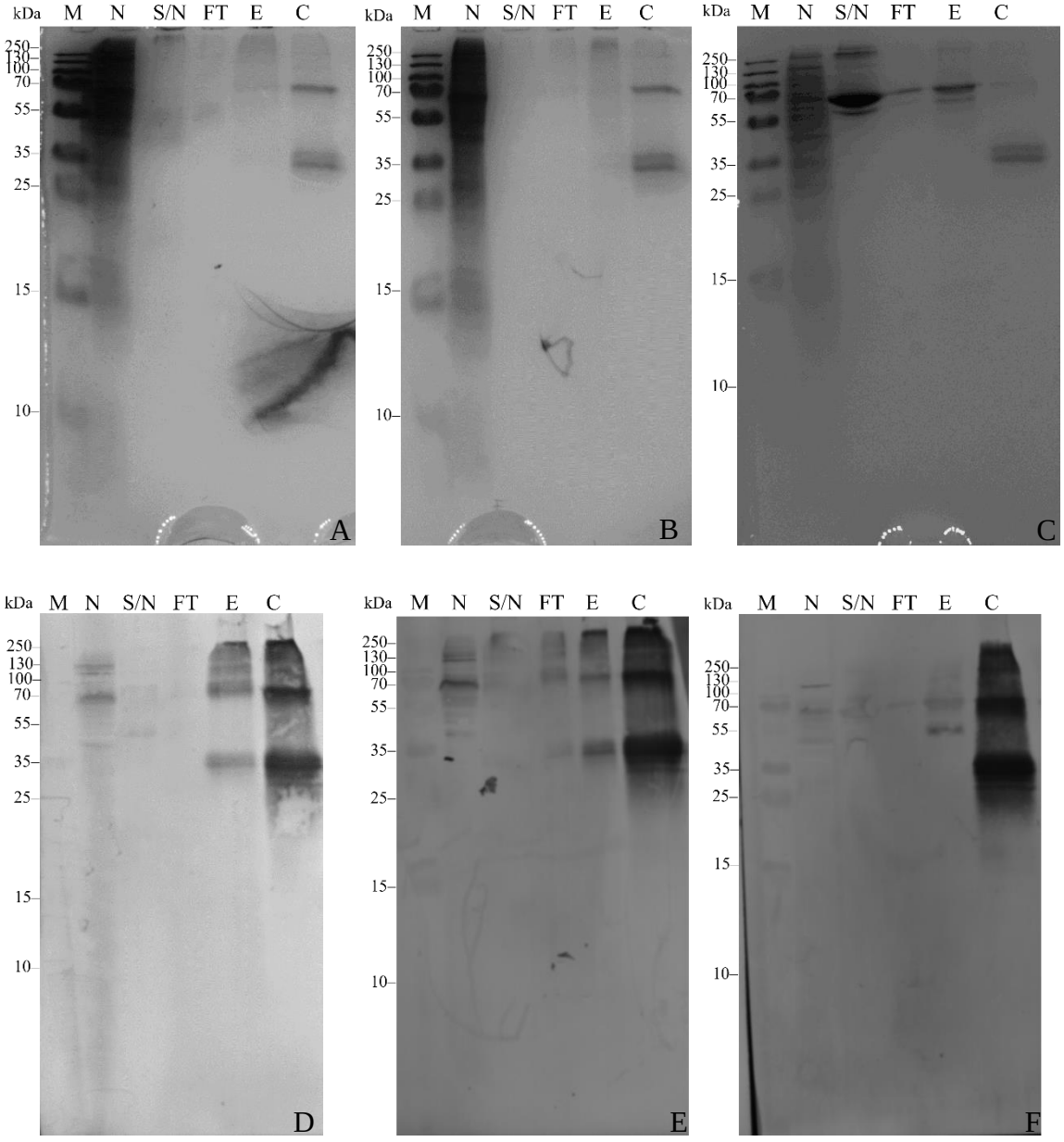


Figure 16: Purification of sialylated proteins from IRE20 (A, D), ISE18 (B, E), and OME22 (C, F) cell line. CBB stained gels (A, B, C), and membranes (D, E, F) with labelled proteins that contain Sia are shown. M – marker; N – cell lysate of corresponding cell line (10 µg); S/N – supernatant after incubation with magnetic beads; FT – flow through; E – elution sample after denaturation of magnetic beads with 4 % SDS; C – control.

It can be seen from the result shown in Fig. 16 that the protein profiles as well as profiles of labelled proteins on membranes for the hard tick cell lines and their purified samples, look almost the same. On the contrary, the protein profile of eluate and the corresponding immunoblotting results for the OME22 cell line is different compared to the hard tick cell lines. Only one major signal on developed membranes (molecular weight of ca 70 kDa) is shared between all three tick cell lines.

5.4.3 Galactose-α-1,3-galactose

The final gels from the purification of αGal for every cell lines are shown in Fig. 17, together with the developed membranes.

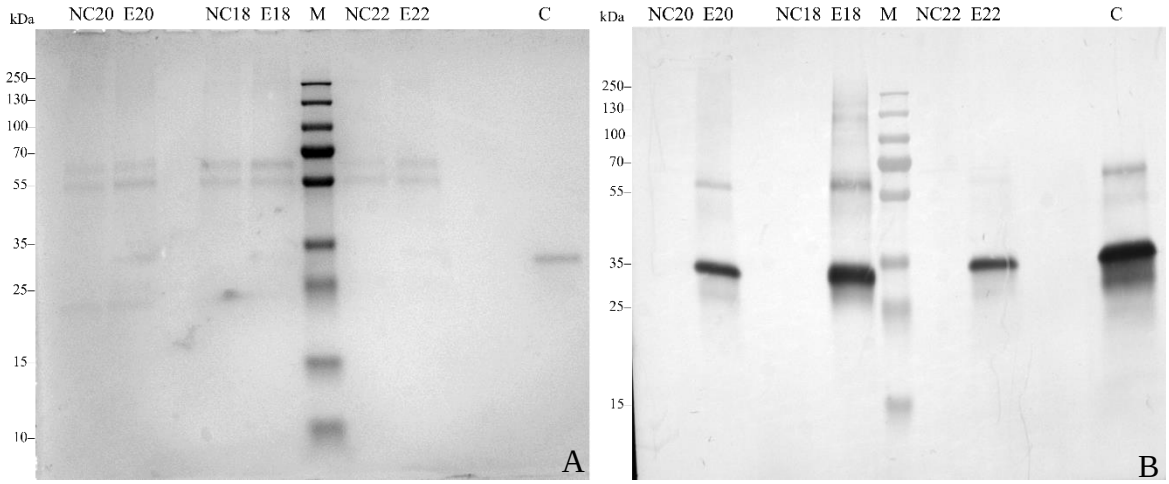


Figure 17: Purification of αGal from the IRE20, ISE18, and OME22 cell lines. CBB stained gel (A), and membrane (B) with labelled proteins that contain αGal are shown. M – marker; NC20 – negative control for IRE20; E20 –elution sample for IRE20; NC18 – negative control for ISE18; E18 – elution sample for ISE18; NC22 – negative control for OME22; E22 – elution sample for OME22; C – control.

The major protein bands detected after lectinoblotting with GSL-B4 (Fig. 17B) show the lectin itself and only a few bands corresponding to the purified αGal-containing proteins are visible in eluates for the ISE18 and IRE20 cell lines; these proteins are of low abundance, as they were not stained by CBB (Fig. 17A). Size of the detected bands correspond to the detection of αGal-containing proteins (Fig. 11).

5.5 Purification of glycosylated proteins, using magnetic/agarose beads, MS data

After purification, MS analysis was used for the identification of proteins present in eluates. All purifications were done in triplicates and every MS analysis was done in

triplicates as well. The MS data from each replicate were compared and only the proteins that occurred in all three repeats (in some cases only two) were chosen. Also, only the proteins with the highest intensity and in all repeats were chosen as a representative for each immunogenic epitope and cell line. The chosen proteins were checked to determine if they come from *Ixodida* and if they had N-glycosylation sites. An online program NetNGlyc – 1.0 was used to find the N-glycosylation sites (Gupta & Brunak, 2002).

The data for the purification of αGal in IRE18 were not conclusive. None of the proteins there were identified in all three replicates and at the same time did not occur in negative control. This is the reason why these results are not shown here.

α1,3-core fucose

The following tables provide a list of the proteins that were detected using MS analysis for the cell lines IRE20 (Tab. XVI), ISE18 (Tab. XVII), and OME22 (Tab. XVII) that most likely contain 1,3-corefucose after purification.

Sialic acid

When processing data from MS identification of Sia-containing proteins, tick and bovine databases were used to identify sialylated proteins of tick-origin vs. bovine-origin, which come from the culture medium. Tick and bovine proteins which presumably contain sialic acid from the IRE20 (Tab. XIX), ISE18 (Tab. XX), and OME22 (Tab. XXI) cell lines, are shown.

Galactose-α-1,3-galactose

Proteins presumably containing the αGal epitope that were identified by MS analysis in the cell lines IRE20 (Tab. XXII), and OME22 (Tab. XXIII), are listed in the following tables.

In the following tables, lists of proteins identified by MS in eluates after affinity purification are shown. The information about Uniprot number and protein molecular weight is shown as well. The column with N-glycosylation sites is showing the sequons identified in the sequences of these proteins predicted by NetNGlyc – 1.0, the number represents the number of the first amino acid in the amino acid sequence. The star in the names of some proteins means that these proteins were identified only in two replicates.

Table XVI: The list of representative proteins with the highest intensity that could contain α1,3-core fucose after purification selected from MS data analysis in the IRE20 cell line. MS analysis data are provided in Supplementary S1.

Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
A0A0K8R9H1	Putative heat shock cognate 70 isoform 1	323.3	71.1	35 (NRT), 64(NNT), 151 (NDS), 360 (NKS), 417 (NTT), 487 (NVS)
A0A7D5BM75	Elongation factor 1-alpha	323.3	50.7	284 (NIT)
A0A6B0VEP9	Putative saposin	323.3	126.4	87 (NET), 300 (NAT), 467 (NMT), 944 (NVT)
A0A131XZ04	Putative prohibitins and stomatins of the pid superfamily protein	323.3	34.7	123 (NYT)
A0A0K8R8I3	Putative low-density lipoprotein receptor domain class a	323.3	75.9	35 (NCT), 42 (NAT), 159 (NKT), 165 (NFT), 534 (NGS), 562 (NGT), 604 (NRT)
A0A0K8R8N9	Heat shock protein 83	323.3	84.2	48 (NSS), 286 (NKT), 392 (NIS)
A0A0K8RIU3	Putative heat shock 70 kDa protein 5	323.3	72.6	NONE
A0A131Y234	Putative transketolase	323.3	67.3	31 (NAS), 335 (NSS)
V5I1A9	Rab GDP dissociation inhibitor (Fragment)	323.3	50.3	167 (NVT)
V5GSI9	Putative glyoxylate/hydroxypyruvate reduct	323.3	35.9	26 (NVT)

Table XVII: The list of representative proteins with the highest intensity that could contain α1,3-core fucose after purification selected from MS data analysis in the ISE18 cell line. MS analysis data are provided in Supplementary S2.

Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
A0A0K8R9H1	Putative heat shock cognate 70 isoform 1	323.3	71.1	64 (NNT), 417 (NTT), 487 (NVS)
A0A131XZ04	Putative prohibitins and stomatins of the pid superfamily protein	323.3	34.7	123 (NYT)
A0A6B0VEP9	Putative saposin	323.3	126.4	87 (NET), 300 (NAT), 467 (NMT), 944 (NVT)
A0A7D5BM75	Elongation factor 1-alpha	323.3	50.7	284 (NIT)
A0A0K8RNI3	Voltage-dependent anion-selective channel protein 3	323.4	30.3	54 (NAS)
A0A6B0VEL5	Putative formyltetrahydrofolate dehydrogenase (Fragment)	323.4	61.1	448 (NVT)
A0A0K8RNA0	Malate dehydrogenase	323.4	35.8	149 (NST)
A0A131Y234	Putative transketolase	323.4	67.3	31 (NAS), 335 (NSS)
A0A0K8R8I3	Putative low-density lipoprotein receptor domain class a	323.4	75.9	35 (NCT), 42 (NAT), 159 (NKT), 165 (NFT), 534 (NGS), 562 (NGT), 604 (NRT)
A0A131XQ15	Putative translation elongation factor 2	323.4	94.3	3 (NFT)

Table XVIII: The list of representative proteins with the highest intensity that could contain α1,3-core fucose after purification selected from MS data analysis in the OME22 cell line. MS analysis data are provided in Supplementary S3.

Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
A0A1Z5L522	E3 ubiquitin-protein ligase TRIP12 (Fragment)	323.3	106.7	176 (NVT), 227 (NYT), 378 (NVT), 431 (NGT), 458 (NLT), 573 (NCS), 612 (NIT), 612 (NIT), 698 (NLT), 774 (NVT)
A0A1Z5KX19	H15 domain-containing protein (Fragment)	323.3	20.3	NONE
A0A1Z5KX72	GTP-binding nuclear protein (Fragment)	92.3	25.1	210 (NTS)
A0A1Z5KXA0	Cofilin (Fragment)	120.8	17.0	46 (NAS)
A0A1Z5L658	GI18784 (Fragment)	258.1	49.6	27 (NGS)
A0A1Z5KYG4	Calponin-homology (CH) domain-containing protein (Fragment)	49.1	26.5	30 (NSS)
A0A1Z5KY24	Sodium/potassium-transporting ATPase subunit alpha (Fragment)*	323.3	114.9	230 (NSS)
A0A1Z5L7U0	Uncharacterized protein (Fragment)*	301.6	84.9	131 (NGT), 252 (NVT), 425 (NIS)
A0A1Z5KWX1	SCP domain-containing protein (Fragment)*	102.9	31.4	NONE
A0A1Z5KXF4	Lamin (Fragment)*	294.1	70.0	339 (NAS)

As it can be seen from the Tab. XVI and XVII for hard tick cell lines a several proteins are similar. On the other hand, when comparing of hard tick cell lines and soft tick cell line (Tab. XVIII) there is not a single protein match.

Table XIX: The list of representative proteins with the highest intensity that could contain Sia after purification selected from MS data analysis in the IRE20 cell line. The information about origin of the proteins is shown as well. MS analysis data are provided in Supplementary S4.

Uniprot number	Protein name	Protein score	Protein mass (kDa)	Organism		N-glycosylation sites (bovine)	N-glycosylation sites (tick)
P00760	Serine protease 1	323.3	25.8	Bos taurus		NONE	
A0A140T897	Albumin	323.3	69.3	Bos taurus		NONE	
B0JYQ0	ALB protein	323.3	69.3	Bos taurus		NONE	
G3X6N3	Serotransferrin	323.3	77.7	Bos taurus		NONE	
P12763	Alpha-2-HS-glycoprotein	323.3	38.4	Bos taurus		99 (NCS), 156 (NDS)	
A0A0K8R9H1	Putative heat shock cognate 70 isoform 1	323.3	71.1		Ixodes ricinus		64 (NNT), 417 (NTT), 487 (NVS)
A0A6B0VEL5	Putative formyltetrahydrofolate dehydrogenase (Fragment)	323.3	61.1		Ixodes ricinus		448 (NVT)
V5GXF3	Putative conserved secreted protein	323.3	44.9		Ixodes ricinus		67 (NST)
A0A0K8R8I3	Putative low-density lipoprotein receptor domain class a	308.6	75.9		Ixodes ricinus		35 (NCT), 42 (NAT), 159 (NKT), 165 (NFT), 534 (NGS), 562 (NGT),604 (NRT)
A0A131Y234	Putative transketolase	323.3	67.3		Ixodes ricinus		31 (NAS), 335 (NSS)

Table XX: The list of representative proteins with the highest intensity that could contain Sia after purification selected from MS data analysis in the ISE18 cell line. The information about origin of the proteins is shown as well. MS analysis data are provided in Supplementary S5.

Uniprot number	Protein name	Protein score	Protein mass (kDa)	Organism		N-glycosylation sites (bovine)	N-glycosylation sites (tick)
A0A140T897	Albumin	323.3	69.3	Bos taurus		NONE	
P00760	Serine protease 1	323.3	25.8	Bos taurus		NONE	
G3X6N3	Serotransferrin	323.3	77.7	Bos taurus		NONE	
B0JYQ0	ALB protein	323.3	69.3	Bos taurus		NONE	
A0A6B0VEL5	Putative formyltetrahydrofolate dehydrogenase (Fragment)	323.3	61.1		Ixodes ricinus		448 (NVT)
A0A0K8R9H1	Putative heat shock cognate 70 isoform 1	323.3	71.1		Ixodes ricinus		64 (NNT), 417 (NTT), 487 (NVS)
P12763	Alpha-2-HS-glycoprotein	323.3	38.4	Bos taurus		99 (NCS), 156 (NDS)	
A0A0K8R8I3	Putative low-density lipoprotein receptor domain class a	308.6	75.9		Ixodes ricinus		35 (NCT), 42 (NAT), 159 (NKT), 165 (NFT), 534 (NGS), 562 (NGT), 604 (NRT)
A0A131XZ04	Putative prohibitins and stomatins of the pid superfamily protein	278.5	34.7		Ixodes ricinus		123 (NYT)
A0A0K8RNA0	Malate dehydrogenase	317.0	35.8		Ixodes ricinus		149 (NST)

Table XXI: The list of representative proteins with the highest intensity that could contain Sia after purification selected from MS data analysis in the OME22 cell line. The information about origin of the proteins is shown as well. MS analysis data are provided in Supplementary S6.

Uniprot number	Protein name	Protein score	Protein mass (kDa)	Organism		N-glycosylation sites (bovine)	N-glycosylation sites (tick)
A0A140T897	Albumin	323.3	69.3	Bos taurus		NONE	
P00760	Serine protease 1	323.3	25.8	Bos taurus		NONE	
Q1RMN8	Immunoglobulin light chain	88.0	24.5	Bos taurus		NONE	
A0A1Z5KX19	H15 domain-containing protein (Fragment)	127.8	20.3		Ornithodoros moubata		NONE
A0A1Z5KYD3	Histone 1 (Fragment)	92.4	10.4		Ornithodoros moubata		NONE
P12763	Alpha-2-HS-glycoprotein	40.0	38.4	Bos taurus		99 (NCS), 156 (NDS)	
E1B8G8	Sushi, nidogen and EGF like domains 1	13.4	145.9	Bos taurus		338 (NFT), 642 (NGT, 907 (NVS)	
A0A1Z5L522	E3 ubiquitin-protein ligase TRIP12 (Fragment)	28.1	106.7		Ornithodoros moubata		176 (NVT), 227 (NYT), 378 (NVT), 431 (NGT), 458 (NLT), 573 (NCS), 612 (NIT), 674 (NIT), 698 (NLT), 774 (NVT)
A0A1Z5L8V7	40S ribosomal protein S9 (Fragment)	91.8	11.9		Ornithodoros moubata		NONE
A0A1Z5KWX1	SCP domain-containing protein (Fragment)	13.5	31.4		Ornithodoros moubata		NONE

In all tables shown above (Tabs. XIX, XX, XXI)the bovine taurus database show at least three proteins match in all tick cell lines. In the IRE20 and ISE18 cell lines even all proteins form bovine taurus are matching. If comparing the hard tick cell lines and soft tick cell lines only two protein matches.

Table XXII: The list of representative proteins with the highest intensity that could contain αGal after purification selected from MS data analysis in the IRE20 cell line. MS analysis data are provided in Supplementary S7.

Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glycosylation sites
V5H1E7	Putative sequestosome 1	267.4	46.1	388 (NGS)
B6V3B5	Glutathione peroxidase	173.6	24.6	12 (NFT)
V5H3S3	Putative 60s ribosomal protein l23 (Fragment)	44.4	21.7	130 (NNT)
A0A131Y447	Profilin	109.3	13.9	38 (NVT)
A0A0K8RE48	Putative prosaposin*	307.9	148.8	87 (NET), 300 (NAT), 467 (NMT), 554 (NST), 657 (NAT), 860 (NGT), 1147 (NVT)
A0A131XUI5	Putative rna-binding protein lark (Fragment)*	39.4	25.8	14 (NAS)
V5I5E9	Putative vesicle-associated membrane protein/syn-aptobrevin-binding protein*	74.9	29.2	203 (NIT)
V5HVJ7	Putative small heat shock protein ii (Fragment)*	246.3	19.7	NONE
A0A131Y2B4	Prohibitin*	323.3	30.0	89 (NIT)
V5H9L8	AP-2 complex subunit alpha*	29.8	104.7	589 (NLS), 748 (NKT), 760 (NVT)
A0A131Y950	Putative low-density lipoprotein receptor domain class a*	167.6	75.9	35 (NCT), 42 (NAT), 159 (NKT), 165 (NFT), 534 (NGS), 562 (NGT), 562 (NGT)

Table XXIII: The list of representative proteins with the highest intensity that could contain αGal after purification selected from MS data analysis in the OME22 cell line. MS analysis data are provided in Supplementary S8.

Uniprot number	Protein name	Protein score	Protein mass (kDa)	N-glykosylation sites
A0A1Z5KY37	Calreticulin (Fragment)	91.4	48.2	NONE
A0A1Z5KWK6	Uncharacterized protein (Fragment)*	12.1	54.7	72 (NST)
A0A1Z5LG77	Elongation factor 1-alpha (Fragment)*	84.1	50.7	284 (NLT)
A0A1Z5KWW2	IDP (Fragment)*	150.2	50.6	131 (NGT)
A0A1Z5KUQ5	Galectin (Fragment)*	79.6	54.6	NONE (only with Pro)
A0A1Z5KV65	Myosin heavy chain non-muscle (Fragment)*	312.5	122.8	NONE
A0A1Z5L6V6	Catalase (Fragment)*	49.4	56.8	435 (NFT), 476 (NFT)
A0A1Z5L5M7	Methionine aminopeptidase 1 (Fragment)*	64.9	37.8	331 (NLT)
A0A1Z5LER6	60S ribosomal protein L40 (Fragment)*	78.0	14.7	NONE
A0A1Z5L2U2	Peptidylprolyl isomerase (Fragment)*	166.7	28.4	NONE

Since for the purification of αGal, only data from IRE20 and OME22 are reliable, we cannot compare the hard tick cell lines together. When comparing Tab. XXII and XXIII, there is no single protein match between the hard and soft tick cell lines.

When comparing the identified proteins from the purification of different epitopes in one tick cell line, three protein match in purification of α1,3-core fucose and Sia in the IRE20 cell line, but not a single protein matches between the proteins possibly bearing the αGal epitope and any of the other two epitopes. In ISE18 there are four matching proteins in purification of proteins containing α1,3-core fucose and Sia. In OME22 cell line, two proteins matched in purification of proteins containing α1,3-core fucose and Sia.

5.6 Fluorescence microscopy

During fluorescence microscopy experiments, major issues emerged in the process. In the first experiment, 70,000 cells per well were used, which proved to be a too-high number with overlapping cellular bodies. Thus, 30,000 to 40,000 cells per well were used in further experiments. Another issue was the non-adherent OME22 cell line, which prevented effective fixation since all the cells were washed away. In the case of αGal detection, the results are not reliable since the cells were not labelled and there was a significant background (the pictures are not shown).

Because of these issues, only a couple of pictures are shown below. These experiments, however, require additional repetition and optimalization. A representative image of α1,3-core fucose detection in IRE20 cell line is shown in Fig. 18.

5.6.1 α1,3-core fucose

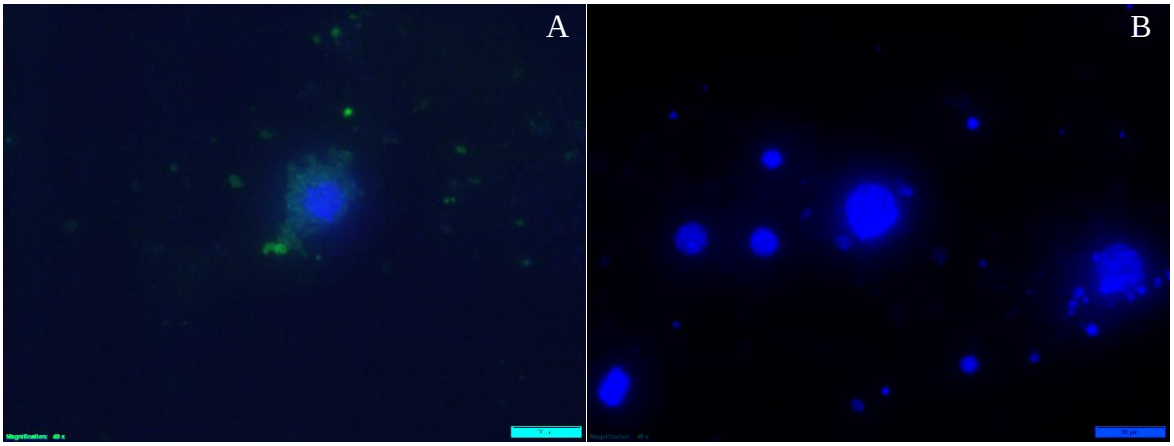


Figure 18: Detection of α1,3-core fucose in IRE20 cell line using fluorescence microscopy. A – detection of α1,3-core fucose using rabbit anti-HRP antibodies; B – negative control. Green – DyLight® 488 Anti-Rabbit IgG, blue – DAPI. The magnification is 40x, and the scale is 20 μm.

The labelling of α1,3-core fucose was very weak. Also, cells formed big clusters where it was hard to distinguish the individual cells and their cellular structures. In some cases, cells were damaged during the process of sample preparation. For the ISE18 cell line, big clusters of cells with strong background were detected without a specific signal; these results were determined as inconclusive.

5.6.2 Sialic acid

Cell fixation during Sia-staining proceeded more successfully, and results for Sia staining in IRE20 and ISE18 are shown in Fig. 19. Still, damaged cells were present in high numbers. Sia was detected in various cellular structures throughout the tick cells.

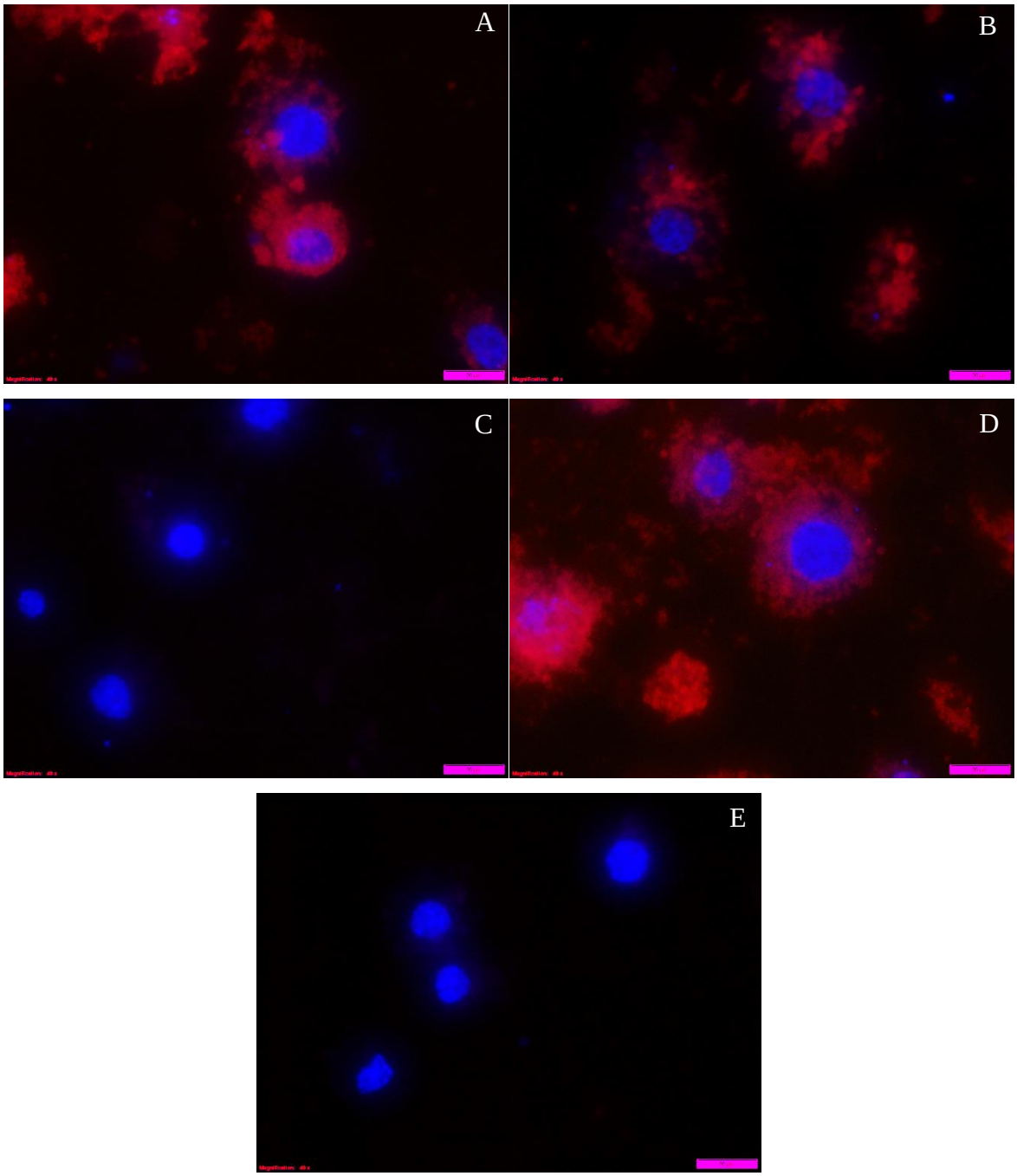


Figure 19: Detection of Sia using fluorescence microscopy in IRE20 (A, B, C) and ISE18 (D, E) cell line. Sia detection using SNA-biotin (A, B, D), negative control (C, E). Red – DyLight®549 Streptavidin; blue – DAPI. The magnification is 40x, and the scale is 20 µm.

6 Discussion

Ticks are blood-sucking parasites that are distributed worldwide. To this date, more than 900 species of ticks have been identified. Their only food source is blood, which they consume in disproportionally vast amounts to their size. In addition to people, they also feed on birds, reptiles, and mammals. Yet, the most relevant and onerous aspect is their potential to transfer pathogens. Several different pathogens that cause diseases in humans are spread by ticks. Lyme disease, anaplasmosis, babesiosis, and ehrlichiosis are the most common infections transmitted by ticks. These illnesses impact not only humans but also livestock, since almost 80 % of the global livestock population is infested with ticks (Bowman & Nuttal, 2008).

Although the immunogenic glycopeptides under investigation in this work have already been examined in other species, there is very little in-depth information on ticks connected to these epitopes (Altmann, 2007; Van Die et al., 1999). There are several potential reasons why the tick does not immediately trigger an immunological response in the host (one epitope weakens another, or multiple epitopes are required). It is also important to find out the purpose and importance of the proteins to which these glycan units are linked.

Proteins bands were cut out from the gels that corresponded with bands on labelled membranes for each epitope in all three cell lines. The most widely used method for protein identification is in-gel digestion of proteins separated by SDS-PAGE in one or two dimensions with subsequent MS identification. However, a low peptide yield is a drawback of in-gel digestion (Yeung et al., 2008), especially while extracting, as some peptides can still be trapped in the gel matrix. The time it takes to complete most protocols is another limitation of in-gel digestion. On the other hand, substances that might possibly affect MS (like detergents and salts) are eliminated by this type of protein analysis (Rosenfeld et al., 1992). To lower the risk of losing glycoproteins, they were purified by two biochemical methods for better results.

The first approach was to purify glycoproteins containing one of the three immunogenic epitopes by affinity ultrafiltration. Different approaches, e.g., columns with different cut-offs, higher concentrations of protein, and adding PNGase F or Endo H to cleave the glycan part from proteins, were tried. However, this method was still insufficient in terms of obtaining purified glycoproteins or glycopeptides. One possible reason was that the binding between glycoprotein and the antibody/lectin interaction was not strong enough to withstand centrifugation. Another reason could be a wrongly chosen column where the glycoproteins

could be trapped in the membrane. Another reason can be that this purification method is usually used for even smaller molecules (less than 30 kDa) (Wei et al., 2016). Finding a suitable macromolecule affinity ligand, however, is typically the main challenge. In these situations, the internal macromolecule can be improved by the addition of a low-molecular-weight affinity ligand (Glatz, 2000).

Because of the unsuccessful result, purification by magnetic/agarose beads was tried as a new approach and it was successful. Elution of the bound glycoproteins from the magnetic beads with 4 % SDS proved to hinder the MS analysis, and thus, glycoproteins bound to the beads were analysed by MS directly, without elution.

Lectinoblotting showed that the galactose-α-1,3-galactose epitope is the least abundant in tick cell lines from the three studied glycoimmunogens. Therefore, a combination of a biotinylated GSL-B4 isoform with goat anti-biotin antibody and Protein G-magnetic beads was used for purification. Purification without lectins but only with anti-biotin as a negative control was also performed. For ISE18, the results from MS analysis show that probably during the preparation of the purified samples or during preparation for MS analysis some unwanted contamination occurred. For this reason, these data are not shown, and the experiment for this cell line will have to be repeated.

With the FM experiment, there were some major issues with cell fixation to the glass, especially with the OME22 cell line, which is non-adherent. The common approach with this type of cells to attach to the glass is performed by cytocentrifugation, where cells are flattened against the glass by subsequent centrifugation. However, cytocentrifugation has a number of disadvantages, slowness when processing numerous samples, high equipment costs, and slightly altered morphology (Fischer et al., 2008; Stokes, 2004). The coating of plastic or glass surfaces with charged substrates like poly-L-lysine to stimulate cellular attachment is another method for the adhesion of suspension cells. However, this calls for extensive planning and overnight treatment, which takes time (Mazia et al., 1975). Tsang et al. (2017) tried different attempt by suspending the cells in PBS rather than growth media, and after 30 minutes of gravity sedimentation, they were able to adhere the cells to plastic or glass (Tsang et al., 2017).

6.1 α1,3-core fucose

α1,3-core fucose is an immunogenic epitope regularly found in insects, especially in insect venoms (e.g., honeybee venom). Also, it has been found in insect neuronal cells (Kurosaka et al., 1991). In this thesis an anti-HRP antibody was used to detect this epitope.

From the result of the Western blot (Figs. 8, 9), it is clear that all three tick cell lines contain a lot of core – fucosylated glycoproteins. These fucosylated glycoproteins play an important role in the colonisation of ticks by pathogens that can be later transferred to humans. In their study, Pedra et al. (2010) discovered that *Anaplasma phagocytophilum*, which is spread by the genus *Ixodes* and causes human granulocytic anaplasmosis, employs α1,3-core fucosylated glycoproteins to colonise the ticks (Pedra et al., 2010).

From the purification results of α1,3-core fucosylated glycoproteins, hundreds of possible fucosylated glycoproteins were identified by MS. When comparing the ten proteins with the highest intensities, all the proteins from the hard tick cell lines (IRE20 and ISE18) have generally more potential glycosylation sites in their sequences than the proteins from the soft tick cell line (OME22). That can be caused by the fact that hard ticks are more likely to spread diseases than soft ticks and are in contact with the host for a longer time, where hiding of the protein surface by a glycan can lower the recognition by the immune system (Shan, 2012). Also, not a single protein matches between the hard and soft ticks in the selected proteins with the highest intensities. When comparing if any protein matches between hard tick cell lines from Tabs. XVI and XVII, it is obvious that six proteins match (Putative heat shock cognate 70 isoform 1; Elongation factor 1-alpha; Putative saposin; Putative prohibitins and stomatins of the pid superfamily protein; Putative low-density lipoprotein receptor domain class a; Putative transketolase). There is no published information yet that these identified proteins were fucosylated in tick. However, Pedersen et al. (2014) found out that low density lipoprotein receptor domain class a is O- and N-glycosylated in humans (Pedersen et al., 2014).

In the future, it will still be necessary to focus more on the results of the MS analysis and to investigate in more detail the functions and determination of the role of individual proteins to understand the transmission of pathogens by ticks.

6.2 Sialic acid

This thesis builds on earlier research oriented on sialic acid in ticks, where Sia was detected in tick guts and salivary glands by Western blot and MS (Vancova et al., 2012). Later, the host blood vs. tick origin of Sia was studied. The result of this study points out that the majority of Sia comes from the host in tick salivary glands and gut (Sterba et al., 2014).

The combination of SDS-PAPGE followed by Western blot with the used lectins cannot determine if the Sia detected is host-origin or tick-origin. To determine the origin of

Sia in all tick cell lines. data from the MS were compared both with the tick database for individual lines and with the bovine database. since the medium in which the cells were grown contains BOFES, and they could thus receive sialylated glycoproteins from the culture medium.

Sia was detected and purified by two types of lectins – SNA, which preferably bind to Sia in α2,6- and less likely in α2,3-linkage, and MAL II, which preferably binds to Sia in α2,3-linkage. With this combination, it was possible to purify glycoproteins containing Sia in both linkages. Even thought, all the sialylated glycoproteins were purified, purification was not done specifically for the immunogenic Neu5Gc. Some of the identified proteins occurred in both MS analyses performed in this study. Also, the first three to four most abundant proteins identified in all three studied cell lines comes from a bovine database. However, these proteins do not include any glycosylation sites and thus are a result of insufficient washing during the purification process. Karamessinis et al., (2008) found out that human Alpha-2-HS-glycoprotein, identified as a sialylated glycoprotein in thick cell lines, contains two N-linked and three O-linked glycosylation sites (Karamessinis et al., 2008).

At the same time, using fluorescence microscopy, it was partially (since the data were limited) determined in which parts of cells the Sia is located. The results from FM experiment are surprising because Sia is usually expressed on cell surfaces (Varki & Gagneux, 2012); thus, it was expected that membranes will be mostly labelled. However, as seen in Fig.19, cellular structures in almost the whole cell (except the nucleus) are labelled for Sia, this corresponds to cellular localization of host-derived sialylated glycoproteins in tick cell lines (J. Štěrbová, personal communication). Still, most of the cells in these samples were damaged, and thus, these results cannot be considered reliable.

When comparing lectinoblot results from hard and soft ticks (Fig. 10), it is clear that the soft tick cell line contains less sialylated glycoproteins and overall Sia than hard tick cell lines. Also, similarly to glycoproteins containing α1,3- core fucose, neither of the ten most abundant identified proteins matched between the hard tick cell lines and the soft tick cell line.

6.3 Galactose-α-1,3-galactose

Galactose-α-1,3-galactose epitope has been thoroughly studied lately because of its connection to red meat allergy. It was first identified during clinical trials where patients were treated with cetuximab, which is a monoclonal antibody used as a treatment for cancer patients. Cetuximab carries the galactose-α-1,3-galactose epitope, which binds to anti-αGal

IgE antibodies, causing anaphylactic and allergic reactions. Since this epitope is not expressed in humans due to mutation, it is a strong immunogen (Commins et al., 2009; Galili, 2018). To this date, a connection has been made between red meat allergy and tick bites. To be more accurate, bites of *Amblyomma americanum* (lone star tick) in the USA, *I. holocyclus* (Australian paralysis tick) in Australia, and *Haemaphysalis longicornis* (Asian long-horned tick) in Asia have been connected to this allergy (Chinuki et al., 2016; Commins et al., 2011; Van Nunen et al., 2009).

The galactose-α-1,3-galactose epitope was detected in all three tick cell lines by Western blot (Fig. 11). Also, the proteins detected by lectinoblot have similar molecular weights. For the hard tick cell lines, the same proteins (Ubiquitin-activating enzyme E1 in band No. 1; Putative heat shock cognate 70 isoform 1 in band No. 2) were identified in protein bands cut out from the gel by MS analysis. On the other hand, no protein was identified in both the hard and soft tick cell lines from the bands cut out of the gels. For MS identification of proteins from purification, the data from purification for ISE18 were not reliable; thus, we can only compare the results from the purification of IRE20 and OME22 tick cell lines. Also in this case, none of the proteins matched between the top ten most abundant proteins. A lot of these proteins do not, to this day, have a clearly stated function. So, in the future, it will be needed to look further into these proteins. To this date there are no data about these proteins in ticks and about their glycosylation. Of the identified putative core-fucosylated proteins, low density lipoprotein receptor domain class a was reported as glycosylated in humans. (Pedersen et al., 2014).

7 Conclusion

The first aim of this work was to detect α1,3-core fucose, Sialic acid, and galactose-α-1,3-galactose in three cell lines – IRE/CTVM20, ISE18, and OME/CTVM22. Detection was carried out using SDS-PAGE and Western blot methods, followed by labelling of membranes using specific antibodies/lectins. The membrane bands were compared to the gel bands, the bands that were then cut out, and were identified later using mass spectrometry. The proteins with highest intensity that appeared in at least two replicates were selected, and their amino acid sequences were examined for N-glycosylation sites.

The second goal of the work was to purify proteins carrying α1,3-core fucose, Sialic acid, and galactose-α-1,3-galactose from IRE/CTVM20, ISE18, and OME/CTVM22. First, the samples were digested with trypsin. Then, we attempted purification using affinity ultrafiltration, which was unsuccessful. The second method was purification using magnetic/agarose beads, which already gave satisfactory results. The eluates were analysed by mass spectrometry. Furthermore, ten proteins with the highest intensity that occurred in at least two repetitions were selected. Their amino acid sequences were then examined to see if they contained N-glycosylation sites. For the MS analysis of Sia, a comparison was made not only with the tick database but also with the bovine database in order to determine whether the Sia originates from the medium or originates from the tick.

The third and final goal was to visualize the IRE/CTVM20, ISE18, and OME/CTVM22 tick cell lines bearing α1,3 – core fucose, Sialic acid, and galactose-α-1,3-galactose using fluorescence microscopy. The Sia visualisation brought the most results here. However, there were some issues with the other epitopes, so the experiment would need further optimisation and repetition.

8 Bibliography

Altmann, F. (2007). The role of protein glycosylation in allergy. *International Archives of Allergy and Immunology*, *142*(2), 99–115. https://doi.org/10.1159/000096114

Anderson, J. F., & Magnarelli, L. A. (1980). Vertebrate host relationships and distribution of *Ixodid* ticks (Acari: *Ixodidae*) in Connecticut, USA. *Journal of Medical Entomology*, *17*(4), 314–323. https://doi.org/10.1093/JMEDENT/17.4.314

Anwar, A., Khan, N. A., & Siddiqui, R. (2019). Galactose as novel target against Acanthamoeba cysts. *PLOS Neglected Tropical Diseases*, *13*(7), e0007385. https://doi.org/10.1371/JOURNAL.PNTD.0007385

Barker, S. C., & Walker, A. R. (2014). Ticks of Australia. The species that infest domestic animals and humans. *Zootaxa*, *3816*(1), 1–144. https://doi.org/10.11646/ZOOTAXA.3816.1.1

Basu, A. K., & Charles, R. A. (2017). *Ticks of Trinidad and Tobago -- an overview*. Academic Press. ISBN 978-0-12-809744-1.

Bell-Sakyi, L., Růžek, D., & Gould, E. A. (2009). Cell lines from the soft tick *Ornithodoros moubata*. *Experimental and Applied Acarology*, *49*(3), 209–219. https://doi.org/10.1007/S10493-009-9258-Y/TABLES/2

Bell-Sakyi, L., Zweygarth, E., Blouin, E. F., Gould, E. A., & Jongejan, F. (2007). Tick cell lines: tools for tick and tick-borne disease research. *Trends in Parasitology*, *23*(9), 450–457. https://doi.org/10.1016/J.PT.2007.07.009

Bowman, A. S., & Nuttal, P. (Eds.). (2008). *Ticks: biology, disease, and control*. Cambridge University Press. ISBN 978-0-521-86761-0.

Breton, C., Šnajdrová, L., Jeanneau, C., Koča, J., & Imberty, A. (2006). Structures and mechanisms of glycosyltransferases. *Glycobiology*, *16*(2), 29R-37R. https://doi.org/10.1093/GLYCOB/CWJ016

Cao, J., Guo, S., Arai, K., Lo, E. H., & Ning, M. M. (2013). Studying extracellular signaling utilizing a glycoproteomic approach: lectin blot surveys, a first and important step. *Methods in Molecular Biology (Clifton, N.J.)*, *1013*, 227–233. https://doi.org/10.1007/978-1-62703-426-5_15

Capinera, J. L. (Ed.). (2008). *Encyclopedia of Entomology* (Vol. 2). Springer Science and Business Media.ISBN 978-1-4020-6359-6.

Carroll, K. K., Guthrie, N., & Ravi, K. (1992). Dolichol: function, metabolism, and accumulation in human tissues. *Biochemistry and Cell Biology, 70*(6), 382–384. https://doi.org/10.1139/O92-059

Chan, J. T. H., Picard-Sánchez, A., Majstorović, J., Rebl, A., Koczan, D., Dyčka, F., Holzer, A. S., & Korytář, T. (2023). Red blood cells in proliferative kidney disease—rainbow trout (*Oncorhynchus mykiss*) infected by *Tetracapsuloides bryosalmonae harbor* IgM+ red blood cells. *Frontiers in Immunology, 14*, 642. https://doi.org/10.3389/FIMMU.2023.1041325

Chinuki, Y., Ishiwata, K., Yamaji, K., Takahashi, H., & Morita, E. (2016). *Haemaphysalis longicornis* tick bites are a possible cause of red meat allergy in Japan. *Allergy*, *71*(3), 421–425. https://doi.org/10.1111/ALL.12804

Cime-Castillo, J., Delannoy, P., Mendoza-Hernández, G., Monroy-Martínez, V., Harduin-Lepers, A., Lanz-Mendoza, H., Hernández-Hernández, F. D. L. C., Zenteno, E., Cabello-Gutiérrez, C., & Ruiz-Ordaz, B. H. (2014). Sialic acid expression in the mosquito aedes aegypti and its possible role in dengue virus-vector interactions. *BioMed Research International*, *2015*. https://doi.org/10.1155/2015/504187

Commins, S. P., James, H. R., Kelly, L. A., Pochan, S. L., Workman, L. J., Perzanowski, M. S., Kocan, K. M., Fahy, J. V., Nganga, L. W., Ronmark, E., Cooper, P. J., & Platts-Mills, T. A. E. (2011). The relevance of tick bites to the production of IgE antibodies to the mammalian oligosaccharide galactose-α-1,3-galactose. *Journal of Allergy and Clinical Immunology*, *127*(5), 1286-1293.e6. https://doi.org/10.1016/j.jaci.2011.02.019

Commins, S. P., & Platts-Mills, T. A. E. (2009). Anaphylaxis syndromes related to a new mammalian cross-reactive carbohydrate determinant. *Journal of Allergy and Clinical Immunology, 124*(4), 652–657. https://doi.org/10.1016/j.jaci.2009.08.026

Commins, S. P., Satinover, S. M., Hosen, J., Mozena, J., Borish, L., Lewis, B. D., Woodfolk, J. A., & Platts-Mills, T. A. E. (2009a). Delayed anaphylaxis, angioedema, or urticaria after consumption of red meat in patients with IgE antibodies specific for galactose-α-1,3-galactose. *Journal of Allergy and Clinical Immunology, 123*(2), 426–433. https://doi.org/10.1016/j.jaci.2008.10.052

Commins, S. P., Satinover, S. M., Hosen, J., Mozena, J., Borish, L., Lewis, B. D., Woodfolk, J. A., & Platts-Mills, T. A. E. (2009b). Delayed anaphylaxis, angioedema, or urticaria after consumption of red meat in patients with IgE antibodies specific for galactose-α-1,3-galactose. *Journal of Allergy and Clinical Immunology*, *123*(2), 426-433.e2. https://doi.org/10.1016/j.jaci.2008.10.052

Connie, R., Robert, W., Vladimir, J., Jean, D., Jung, C., & Yael, A. (2013). 3.4 Proteins. In *Biology*. Open Stax College. https://openstax.org/books/biology/pages/3-4-proteins

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology 2008 26:12*, *26*(12), 1367–1372. https://doi.org/10.1038/nbt.1511

De Schutter, K., & Van Damme, E. J. M. (2015). Protein-carbohydrate interactions as part of plant defense and animal immunity. *Molecules 2015*, *20*(5), 9029–9053. https://doi.org/10.3390/MOLECULES20059029

Domon, B., & Aebersold, R. (2006). Mass spectrometry and protein analysis. *Science*, *312*(5771), 212–217. https://doi.org/10.1126/SCIENCE.1124619

Drummond, R. (2013). *Ticks and what you can do about them*. Wilderness Press. ISBN 978-0-89997-353-1.

Endo, T. (1999). O-Mannosyl glycans in mammals. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1473*(1), 237–246. https://doi.org/10.1016/S0304-4165(99)00182-8

Esmail, S., & Manolson, M. F. (2021). Advances in understanding N-glycosylation structure, function, and regulation in health and disease. *European Journal of Cell Biology*, *100*(7–8), 151186. https://doi.org/10.1016/J.EJCB.2021.151186

Fischer, A. H., Jacobson, K. A., Rose, J., & Zeller, R. (2008). Preparation of slides and coverslips for microscopy. *Cold Spring Harbor Laboratory Press*, *3*(5). https://doi.org/10.1101/pdb.prot4988

Forinová, M., Pilipenco, A., Víšová, I., Lynn, N. S., Dostálek, J., Mašková, H., Hönig, V., Palus, M., Selinger, M., Kočová, P., Dyčka, F., Štěrba, J., Houska, M., Vrabcová, M., Horák, P., Anthi, J., Tung, C. P., Yu, C. M., Chen, C. Y., … Vaisocherová-Lísalová, H. (2021). Functionalized terpolymer-brush-based biointerface with improved antifouling properties for ultra-sensitive direct detection of virus in crude clinical samples. *ACS Applied Materials and Interfaces*, *13*(50), 60612–60624. https://doi.org/10.1021/ACSAMI.1C16930/SUPPL_FILE/AM1C16930_SI_003.XLS

Franc, V., Šebela, M., Řehulka, P., Končitíková, R., Lenobel, R., Madzak, C., & Kopečný, D. (2012). Analysis of N-glycosylation in maize cytokinin oxidase/dehydrogenase 1 using a manual microgradient chromatographic separation coupled offline to MALDI-TOF/TOF mass spectrometry. *Journal of Proteomics*, *75*(13), 4027–4037. https://doi.org/10.1016/J.JPROT.2012.05.013

Frant, M., Woźniakowski, G., & Pejsak, Z. (2017). African swine fever (ASF) and ticks. No risk of tick-mediated ASF spread in Poland and Baltic States. *Journal of Veterinary Research (Poland)*, *61*(4), 375–380. https://doi.org/10.1515/JVETRES-2017-0055

Gagneux, P., & Varki, A. (1999). Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology*, *9*(8), 747–755. https://doi.org/10.1093/GLYCOB/9.8.747

Galili, U. (2018). Anti-Gal IgE mediates allergies to red meat. *The Natural Anti-Gal Antibody As Foe Turned Friend In Medicine*, 117–128. https://doi.org/10.1016/B978-0-12-813362-0.00007-5

Gallagher, S. R. (2006). One-dimensional SDS gel electrophoresis of proteins. *Current Protocols in Immunology*, *75*(1), 8.4.1-8.4.37. https://doi.org/10.1002/0471142735.IM0804S75

Gentzsch, M., & Tanner, W. (1996). The PMT gene family: protein O-glycosylation in Saccharomyces cerevisiae is vital. *The EMBO Journal*, *15*(21), 5752–5759. https://doi.org/10.1002/J.1460-2075.1996.TB00961.X

Glatz, Z. (2000). Affinity ultrafiltration of proteins. *Chemické Listy*, *94*(8).

Gupta, R., & Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 310–322. https://doi.org/10.1142/9789812799623_0029

Harris, R. J., & Spellman, M. W. (1993). O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology*, *3*(3), 219–224. https://doi.org/10.1093/GLYCOB/3.3.219

Helenius, A., & Zurich, E. (2001). Intracellular functions of N-linked glycans. *Science*, *291*(5512), 2364–2369. https://doi.org/10.1126/science.291.5512.2364

Ho, C. S., Lam, C. W. K., Chan, M. H. M., Cheung, R. C. K., Law, L. K., Lit, L. C. W., Ng, K. F., Suen, M. W. M., & Tai, H. L. (2003). Electrospray ionisation mass spectrometry: principles and clinical applications. *The Clinical Biochemist. Reviews*, *24*(1), 3–12.

Ishihara, H., Takahashi, N., Oguri, S., & Tejima, S. (1979). Complete structure of the carbohydrate moiety of stem bromelain. An application of the almond glycopeptidase for structural studies of glycopeptides. *Journal of Biological Chemistry*, *254*(21), 10715–10719. https://doi.org/10.1016/S0021-9258(19)86580-8

Kamerling, J. P., & Gerwig, G. J. (2006). Structural analysis of naturally occurring sialic acids. *Methods in Molecular Biology (Clifton, N.J.)*, *347*, 69–91. https://doi.org/10.1385/1-59745-167-3:69

Karamessinis, P. M., Malamitsi-Puchner, A., Boutsikou, T., Makridakis, M., Vougas, K., Fountoulakis, M., Vlahou, A., & Chrousos, G. (2008). Marked defects in the expression and glycosylation of 2-HS glycoprotein/fetuin-A in plasma from neonates with intrauterine growth restriction. *Molecular and Cellular Proteomics*, *7*, 591–599. https://doi.org/10.1074/mcp.M700422-MCP200

Kiszewski, A. E., & Spielman, A. (2002). Preprandial inhibition of re-mating in *Ixodes* ticks (Acari: *Ixodidae*). *Journal of Medical Entomology*, *39*(6), 847–853. https://doi.org/https://doi.org/10.1603/0022-2585-39.6.847

Kobata, A. (1992). Structures and functions of the sugar chains of glycoproteins. *European Journal of Biochemistry*, *209*, 992. https://doi.org/10.1111/j.1432-1033.1992.tb17313.x

Kreppel, L. K., Blomberg, M. A., & Hart, G. W. (1997). Dynamic glycosylation of nuclear and cytosolic proteins. *Journal of Biological Chemistry*, *272*(14), 9308–9315. https://doi.org/10.1074/jbc.272.14.9308

Kurosaka, A., Yano, A., Itoh, N., Kuroda, Y., Nakagawa, T., & Kawasaki, T. (1991). The structure of a neural specific carbohydrate epitope of horseradish peroxidase recognized by anti-horseradish peroxidase antiserum. *Journal of Biological Chemistry*, *266*(7), 4168–4172. https://doi.org/10.1016/S0021-9258(20)64302-2

Loginov, D. S., Loginova, Y. F., Dycka, F., Böttinger, K., Vechtova, P., & Sterba, J. (2019). Tissue-specific signatures in tick cell line MS profiles. *Parasites and Vectors*, *12*(1), 1–12. https://doi.org/10.1186/S13071-019-3460-5/FIGURES/6

Lubas, W. A., Frank, D. W., Krause, M., & Hanover, J. A. (1997). O-linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. *Journal of Biological Chemistry*, *272*(14), 9316–9324. https://doi.org/10.1074/jbc.272.14.9316

Macher, B. A., & Galili, U. (2008). The Galα1,3Galβ1,4GlcNAc-R (α-Gal) epitope: a carbohydrate of unique evolution and clinical relevance. *Biochimica et Biophysica Acta*, *1780*(2), 75. https://doi.org/10.1016/J.BBAGEN.2007.11.003

Mahmood, T., & Yang, P. C. (2012). Western blot: technique, theory, and trouble shooting. *North American Journal of Medical Sciences*, *4*(9), 429. https://doi.org/10.4103/1947-2714.100998

Mazia, D., Schatten, G., & Sale, W. (1975). Adhesion of cells to surfaces coated with polylysine. Applications to electron microscopy. *The Journal of Cell Biology*, *66*(1), 198–200.

Moravcová, D., Kahle, V., Řehulková, H., Chmelík, J., & Řehulka, P. (2009). Short monolithic columns for purification and fractionation of peptide samples for matrix-assisted laser desorption/ionization time-of-flight/time-of-flight mass spectrometry analysis in proteomics. *Journal of Chromatography A*, *1216*(17), 3629–3636. https://doi.org/10.1016/J.CHROMA.2009.01.075

Mullins, R. J., James, H., Platts-Mills, T. A. E., & Commins, S. (2012). Relationship between red meat allergy and sensitization to gelatin and galactose-α-1,3-galactose. *Journal of Allergy and Clinical Immunology*, *129*(5), 1334–1342. https://doi.org/10.1016/J.JACI.2012.02.038

Munderloh, U. G., Liu, Y., Wang, M., Chen, C., & Kurtti, T. J. (1994). Establishment, maintenance and description of cell lines from the tick Ixodes scapularis. *Journal of Parasitology*, *80*(4), 533–543. https://doi.org/10.2307/3283188

Nagamune, K., Acosta-Serrano, A., Uemura, H., Brun, R., Kunz-Renggli, C., Maeda, Y., Ferguson, M. A. J., & Kinoshita, T. (2004). Brief definitive report surface sialic acids taken from the host allow Trypanosome survival in Tsetse fly vectors. *The Journal of Experimental Medicine*, *199*(10), 1445–1450. https://doi.org/10.1084/jem.20030635

Nunen, S. A. van, O'connor, K. S., Clarke, L. R., Boyle, R. X., & Fernando, S. L. (2009). An association between tick bite reaction and red meat allergy in humans. *The Medical Journal of Australia*, *190*(9), 510–511. https://doi.org/https://doi.org/10.5694/j.1326-5377.2009.tb02533.x

Pedersen, N. B., Wang, S., Narimatsu, Y., Yang, Z., Halim, A., Schjoldager, K. T. B. G., Madsen, T. D., Seidah, N. G., Bennett, E. P., Levery, S. B., & Clausen, H. (2014). Low density lipoprotein receptor class A repeats are O-glycosylated in linker regions. *Journal of Biological Chemistry*, *289*(25), 17312–17324. https://doi.org/10.1074/jbc.M113.545053

Pedra, J. H. F., Narasimhan, S., Rendić, D., DePonte, K., Bell-Sakyi, L., Wilson, I. B. H., & Fikrig, E. (2010). Fucosylation enhances colonization of ticks by *Anaplasma phagocytophilum*. *Cellular Microbiology*, *12*(9), 1222–1234. https://doi.org/10.1111/J.1462-5822.2010.01464.X

Rand, P. W., Lubelczyk, C., Lavigne, G. R., Elias, S., Holman, M. S., Lacombe, E. H., & Smith, R. P. (2003). Deer density and the abundance of *Ixodes scapularis* (Acari: *Ixodidae*). *Journal of Medical Entomology*, *40*(2), 179–184. https://doi.org/https://doi.org/10.1603/0022-2585-40.2.179

Rappsilber, J., Mann, M., & Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols 2007 2:8*, *2*(8), 1896–1906. https://doi.org/10.1038/nprot.2007.261

Raymond, S., & Weintraub, L. (1959). Acrylamide gel as a supporting medium for zone electrophoresis. *Science*, *130*(3377), 711–711. https://doi.org/10.1126/SCIENCE.130.3377.711.A

Rosenfeld, J., Capdevielle, J., Guillemot, J. C., & Ferrara, P. (1992). In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Analytical Biochemistry*, *203*(1), 173–179. https://doi.org/10.1016/0003-2697(92)90061-B

Roth, J., Kempf, A., Reuter, G., Schauer, R., & Gehring, W. J. (1992). Occurrence of sialic acids in *Drosophila melanogaster*. *Science*, *256*(5057), 673–675. https://doi.org/10.1126/science.1585182

Rovid Spickler, A. (2022). *Ixodes ricinus*. https://www.cfsph.iastate.edu/Factsheets/pdfs/ixodes_ricinus.pdf

Rudd, P. M., Dwek, R., Opdenakker, G., van den Steen, P., & Dwek, R. A. (1998). Concepts and principles of O-linked glycosylation. *Critical Reviews in Biochemistry and Molecular Biology*, *33*(3), 151–208. https://doi.org/10.1080/10409239891204198

Saari, S., Näreaho, A., & Nikander, S. (2019). Arachnida. In *Canine Parasites and Parasitic Diseases* (pp. 187–228). Elsevier. https://doi.org/10.1016/B978-0-12-814112-0.00009-X

Sadler, E., J. (1984). Biosynthesis of glycoproteins: formation of O-linked oligosaccharides. *Biology of Carbohydrates*, *2*, 199–213.

Saldova, R., & Wilkinson, H. (2020). Current methods for the characterization of O-glycans. *Journal of Proteome Research*, *19*(10), 3890–3905. https://doi.org/10.1021/ACS.JPROTEOME.0C00435/ASSET/IMAGES/LARGE/PR0C00435_0005.JPEG

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

Schotthoefer, A. M., & Frost, H. M. (2015). Ecology and epidemiology of lyme borreliosis. *Clinics in Laboratory Medicine*, *35*(4), 723–743. https://doi.org/10.1016/J.CLL.2015.08.003

Seeberger, P. H. (2022). Essentials of glycobiology. In *Essentials of Glycobiology* (Vol. 4). Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/GLYCOBIOLOGY.4E.2. ISBN 978-1-621824-22-0

Shan, M. M. (Ed.). (2012). *Parasitology*. InTech. ISBN 978-953-51-0149-9

Shevchenko, A., Tomas, H., Havliš, J., Olsen, J. V., & Mann, M. (2007). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols 2007 1:6*, *1*(6), 2856–2860. https://doi.org/10.1038/nprot.2006.468

Sonenshine, D. E., & Roe, M. R. (Eds.). (2013). *Biology of ticks volume 2*. Oxford University Press. ISBN 978-0-19-974405-3.

Spiro, R. G. (1973). Glycoproteins. *Advances in Protein Chemistry*, *27*(C), 349–467. https://doi.org/10.1016/S0065-3233(08)60451-9

Spiro, R. G. (2002). Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, *12*(4), 43R-56R. https://doi.org/10.1093/GLYCOB/12.4.43R

Sterba, J., Vancova, M., Sterbova, J., Bell-Sakyi, L., & Grubhoffer, L. (2014). The majority of sialylated glycoproteins in adult *Ixodes ricinus* ticks originate in the host, not the tick. *Carbohydrate Research*, *389*(1), 93–99. https://doi.org/10.1016/J.CARRES.2014.02.017

Stokes, B. O. (2004). Principles of cytocentrifugation. *Laboratory Medicine*, *35*(7), 434–437. https://doi.org/10.1309/FTT59GWKDWH69FB0

Taylor, M., & Drickamer, K. (2011). Introduction to glycobiology. In *Introduction to glycobiology* (2nd ed., Vol. 40). Oxford University Press.

Troughton, D. R., & Levin, M. L. (2007). Life cycles of seven *Ixodid* tick species (Acari: *Ixodidae*) under standardized laboratory conditions. *Journal of Medical Entomology*, *44*(5), 732–740. https://doi.org/https://doi.org/10.1093/jmedent/44.5.732

Tsang, M., Gantchev, J., Ghazawi, F. M., & Litvinov, I. V. (2017). Protocol for adhesion and immunostaining of lymphocytes and other non-adherent cells in culture. *BioTechniques*, *63*(5), 230–233. https://doi.org/10.2144/000114610/

Tyanova, S., & Cox, J. (2018). Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. *Methods in Molecular Biology*, *1711*, 133–148. https://doi.org/10.1007/978-1-4939-7493-1_7/FIGURES/5

van der Veen, M. J., van Ree, R., Aalberse, R. C., Akkerdaas, J., Koppelman, S. J., Jansen, H. M., & van der Zee, J. S. (1997). Poor biologic activity of cross-reactive IgE directed to carbohydrate determinants of glycoproteins. *Journal of Allergy and Clinical Immunology*, *100*(3), 327–334. https://doi.org/10.1016/S0091-6749(97)70245-8

Van Die, I., Gomord, V., Kooyman, F. N. J., Van Den Berg, T. K., Cummings, R. D., & Vervelde, L. (1999). Core α1 → 3-fucose is a common modification of N-glycans in parasitic helminths and constitutes an important epitope for IgE from Haemonchus contortus infected sheep. *FEBS Letters*, *463*(1–2), 189–193. https://doi.org/10.1016/S0014-5793(99)01508-2

Van Nunen, S. A., O'Connor, K. S., Clarke, L. R., Boyle, R. X., & Fernando, S. L. (2009). An association between tick bite reactions and red meat allergy in humans. *The Medical Journal of Australia*, *190*(9), 510–511. https://doi.org/10.5694/J.1326-5377.2009.TB02533.X

Vancová, M., & Nebesářová, J. (2015). Correlative fluorescence and scanning electron microscopy of labelled core fucosylated glycans using cryosections mounted on carbon-patterned glass slides. *PLoS One*, *10*(12). https://doi.org/10.1371/journal.pone.0145034

Vancova, M., Sterba, J., Dupejova, J., Simonova, Z., Nebesarova, J., Novotny, M. V., & Grubhoffer, L. (2012a). Uptake and incorporation of sialic acid by the tick *Ixodes ricinus*. *Journal of Insect Physiology*, *58*(9), 1277–1287. https://doi.org/10.1016/J.JINSPHYS.2012.06.016

Vancova, M., Sterba, J., Dupejova, J., Simonova, Z., Nebesarova, J., Novotny, M. V., & Grubhoffer, L. (2012b). Uptake and incorporation of sialic acid by the tick *Ixodes ricinus*. *Journal of Insect Physiology*, *58*(9), 1277–1287. https://doi.org/10.1016/J.JINSPHYS.2012.06.016

Varki, A. (2010). Colloquium Paper: Uniquely human evolution of sialic acid genetics and biology. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(Suppl 2), 8939–8946. https://doi.org/10.1073/PNAS.0914634107

Varki, A., & Gagneux, P. (2012). Multifarious roles of sialic acids in immunity. *Annals of the New York Academy of Science*, *1253*(1), 16–36. https://doi.org/10.1111/j.1749-6632.2012.06517.x

Věchtová, P., Štěrbová, J., Štěrba, J., Vancová, M., Rego, R. O. M., Selinger, M., Strnad, M., Golovchenko, M., Rudenko, N., & Grubhoffer, L. (2018). A bite so sweet: the glycobiology interface of tick-host-pathogen interactions. *Parasites & Vectors 2018*, *11*(1), 1–27. https://doi.org/10.1186/S13071-018-3062-7

Wardyn, J. D., & Jeyasekharan, A. D. (2018). Immunofluorescence. *ELS*, 1–9. https://doi.org/10.1002/9780470015902.a0001174.pub2

Weerapana, E., & Imperiali, B. (2006). Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology*, *16*(6), 91–101. https://doi.org/10.1093/glycob/cwj099

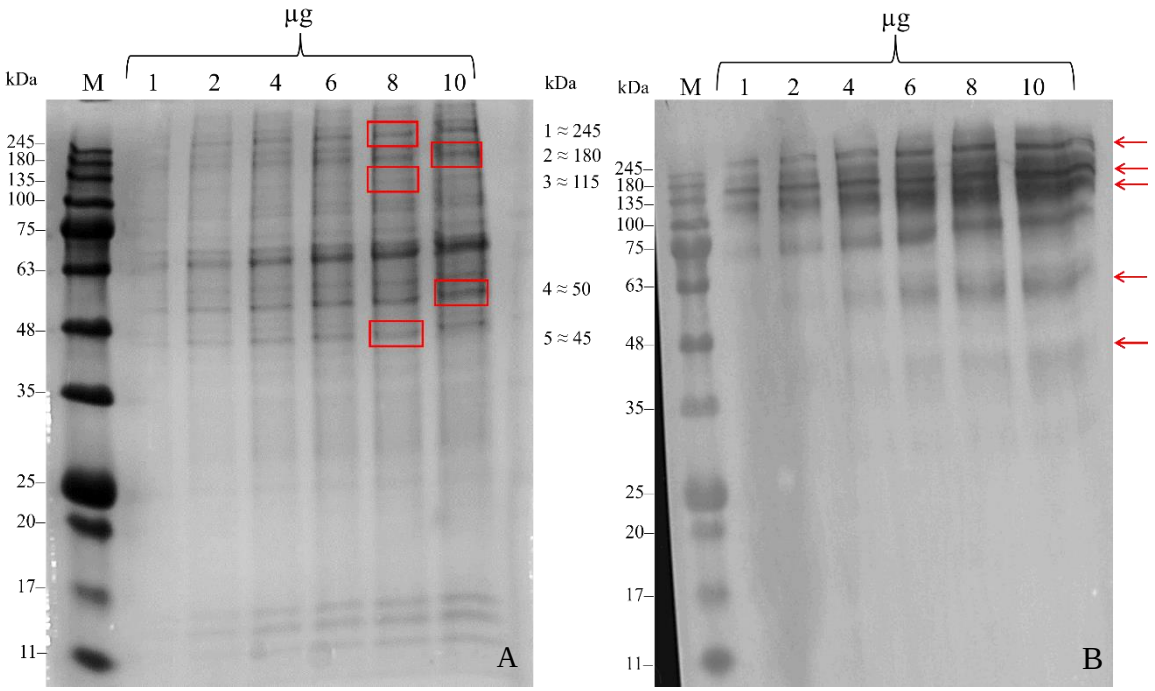
Wei, H., Zhang, X., Tian, X., & Wu, G. (2016). Pharmaceutical applications of affinity-ultrafiltration mass spectrometry: Recent advances and future prospects. In *Journal of Pharmaceutical and Biomedical Analysis* (Vol. 131, pp. 444–453). Elsevier B.V. https://doi.org/10.1016/j.jpba.2016.09.021

Yehuda, S., & Padler-Karavani, V. (2020). Glycosylated biotherapeutics: immunological effects of N-glycolylneuraminic acid. *Frontiers in Immunology*, *11*. https://doi.org/10.3389/FIMMU.2020.00021

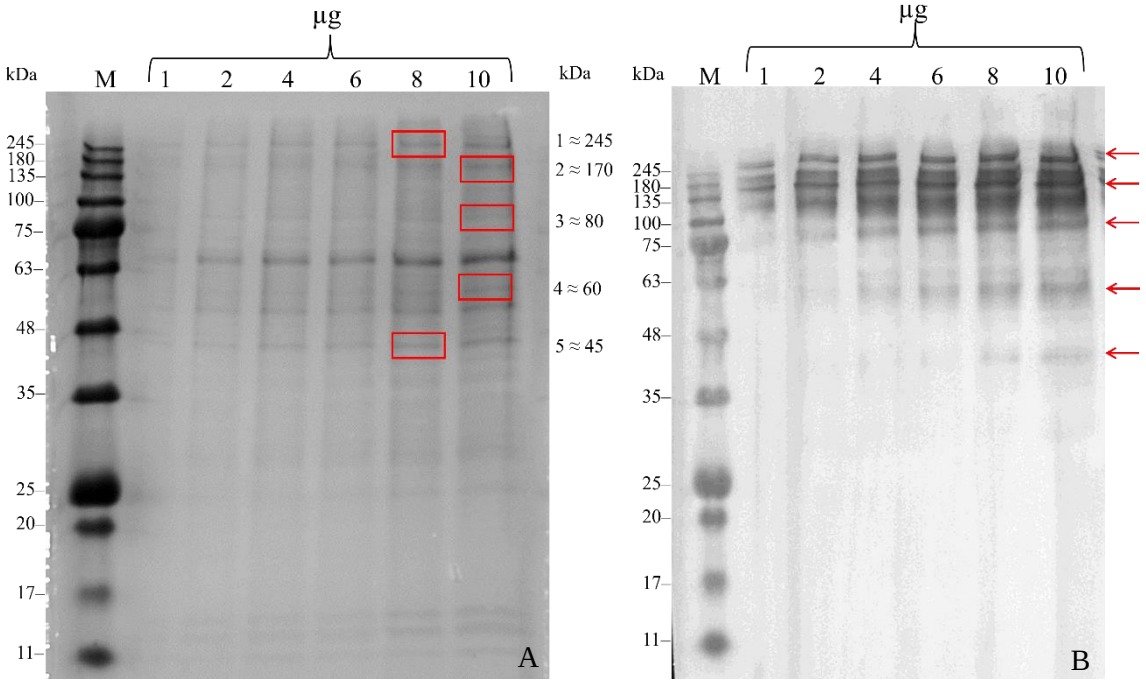
Yeung, Y. G., Nieves, E., Angeletti, R. H., & Stanley, E. R. (2008). Removal of detergents from protein digests for mass spectrometry analysis. *Analytical Biochemistry*, *382*(2), 135. https://doi.org/10.1016/J.AB.2008.07.034

9 Appendix

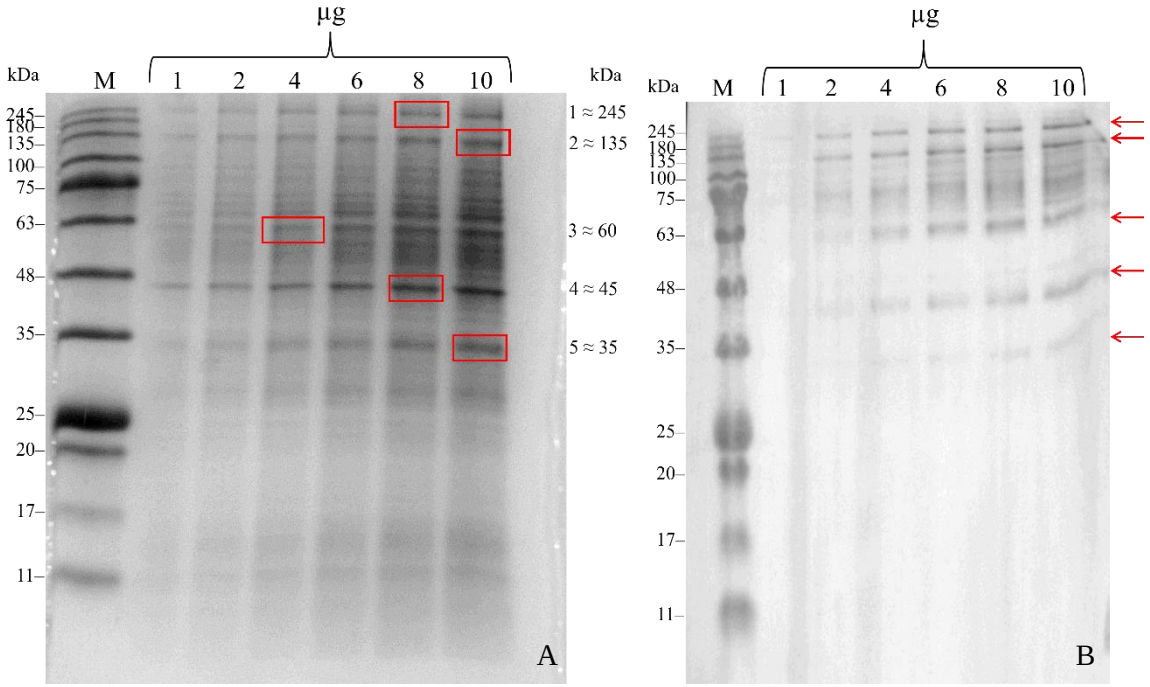
Appendix 1:Full CBB-stained gels (A) and membranes (B) for detection of α1,3-core fucose in IRE20.



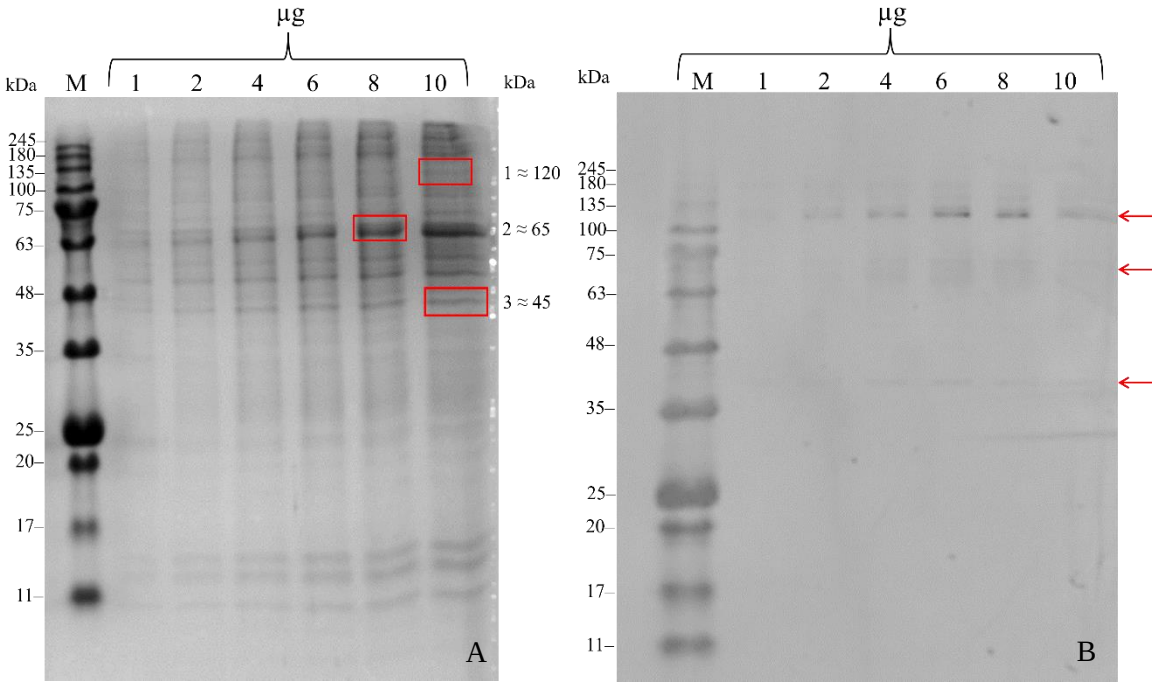
Appendix 2: Full CBB-stained gels (A) and membranes (B) for detection of α1,3-core fucose in ISE18.



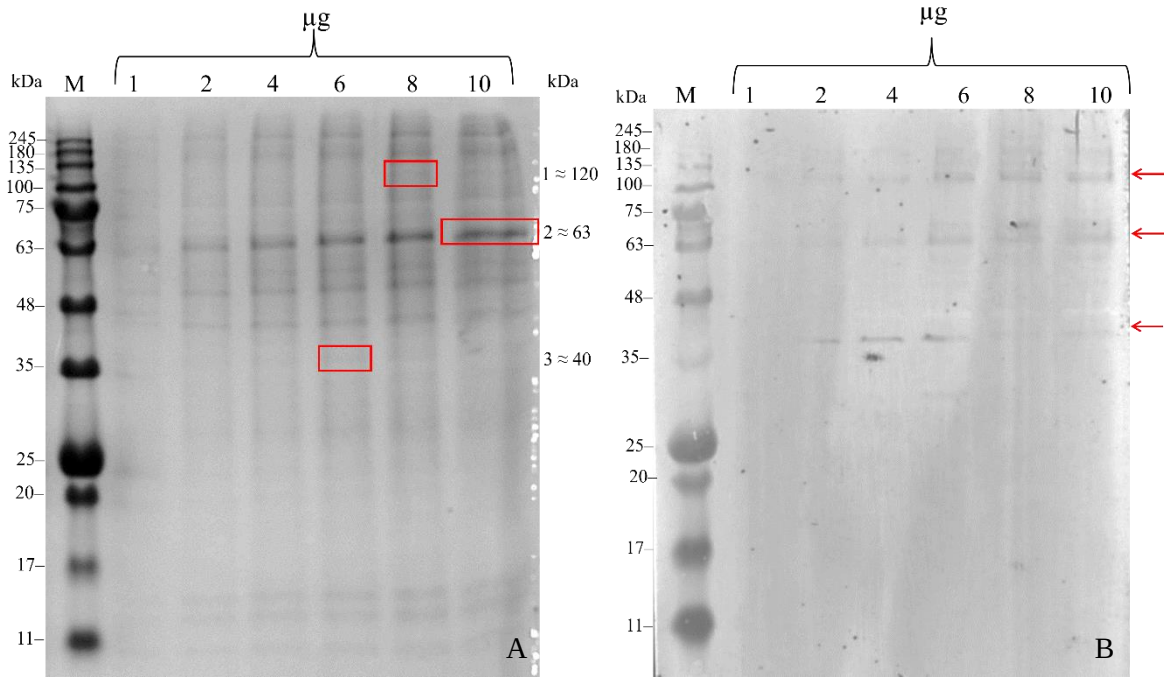
Appendix 3: Full CBB-stained gels (A) and membranes (B) for detection of α1,3-core fucose in OME22.



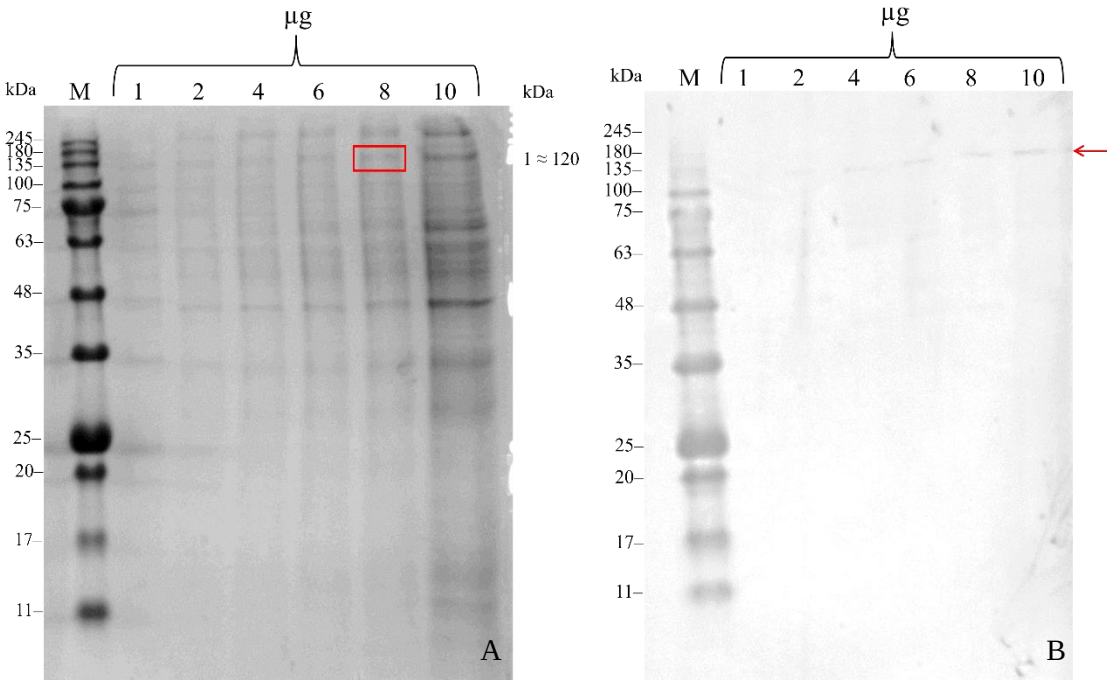
Appendix 4: Full CBB-stained gels (A) and membranes (B) for detection of Sialic acid in IRE20.



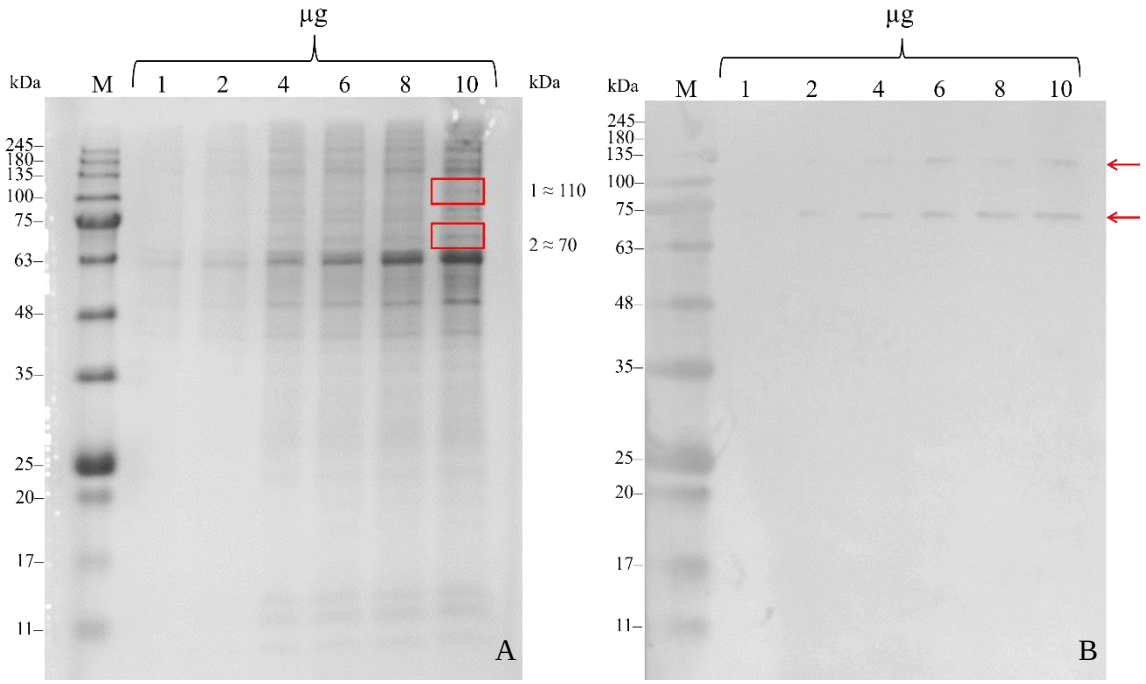
Appendix 5: Full CBB-stained gels (A) and membranes (B) for detection of Sialic acid in ISE18.



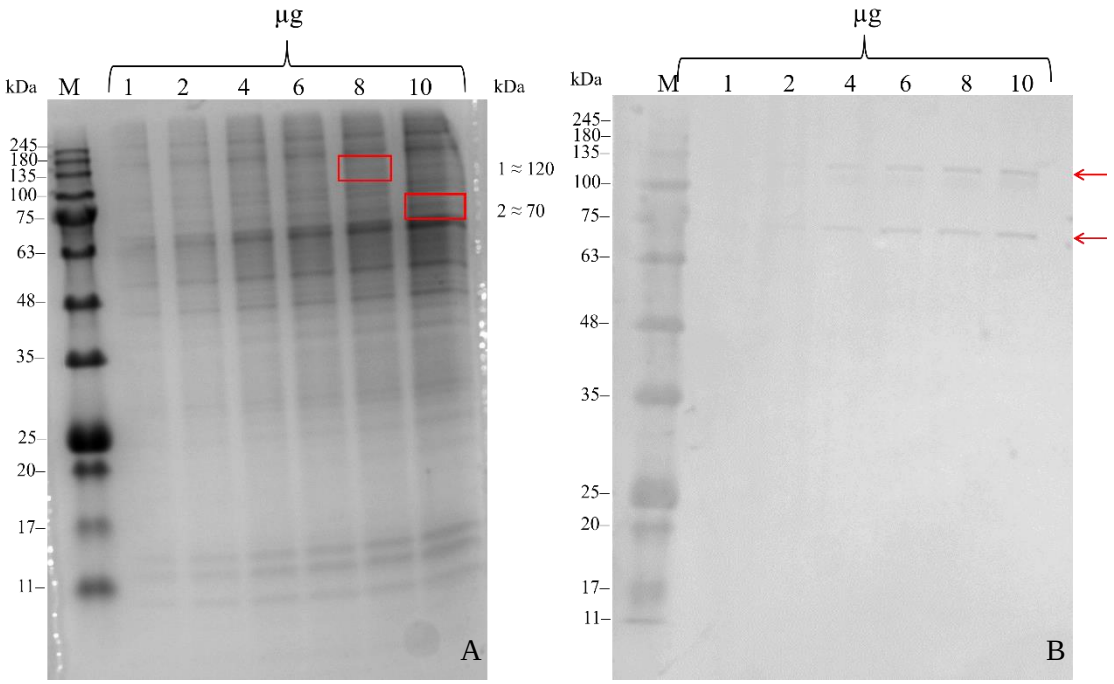
Appendix 6: Full CBB-stained gels (A) and membranes (B) for detection of Sialic acid in OME22.



Appendix 7: Full CBB-stained gels (A) and membranes (B) for detection of galactose-α-1,3-galactose in IRE20.



Appendix 8: Full CBB-stained gels (A) and membranes (B) for detection of galactose-α-1,3-galactose in ISE18.



Appendix 9: Full CBB-stained gels (A) and membranes (B) for detection of galactose-α-1,3-galactose in OME22.