# Bio-lubrication

## Structural Investigation of Lubricin and its Glycosylation

Liaqat Ali

Department of Medical Biochemistry

Institute of Biomedicine

Sahlgrenska Academy at University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2014

Cover illustration: Antibody immunofluorescence on arthritis cartilage biopsy cryosection with anti-lubricin antibody. Gratefully provided by Dr Sarah Flowers.

# ABSTRACT

The sliding articular cartilage surfaces of the human diarthrodial joints are surrounded by biolubricating synovial fluid (SF), creating a perfect low friction biological biobearing structure with excellent lubrication and wear resistance, even during motion. Lubrication is predominately provided by surface adhered biomolecules including phospholipids, hyaluronic acid and synovial lubricin. Inflammation, such as arthritis and injury, changes the joint assembly resulting in detachment of essential surface molecules, increasing friction and wear of the sliding articular cartilage. Changes in the composition and distribution of these surface molecules are suggested to aggravate the disease. The lubricative, heavily glycosylated mucin-like synovial glycoprotein, lubricin, has previously been observed to contain glycosylation changes related to rheumatoid arthritis and osteoarthritis. Therefore, a structural investigation of lubricin and its glycosylation was initiated in order to better understand the biolubricating ability of lubricin and its pathological involvement in arthritic diseases.

An investigation was undertaken to better understand the nature of the dominant glycan structure, sialic acid. Sialidase specific for α2-3 linked sialic acid and subsequent UniCarb-DB fragment spectra comparison of the resultant structure indicated the exclusive 3-linked nature of core 2 sialylation. However, core 1 structures had both 3 and 6 linked sialylation. In arthritis, lubricin has been shown to degrade as identified by its fragments in the synovial fluid. This may open up a new possibility for identification of disease specific biomarker. The mass spectrometric glycopeptide analysis showed that lubricin contains an extended serine, threonine and proline (STP) rich domain composed of imperfect tandem repeats (EPAPTTPK), the target for *O*-glycosylation. The *N*-acetylgalactosaminyltransferase (*GALNTs*) expression analysis of the fibroblast-like synoviocytes showed high expression of the less understood *GALNT5* and *15* in addition to the ubiquitously expressed *GALNT1* and *2*. This indicated that lubricin required a unique combination of transferase genes for its glycosylation.

Overall, this study revealed that negatively charged sialic acid in the mucin-like domain makes the lubricin molecule amphoteric in nature, as the arginine and lysine rich protein core is positively charged. The number of glycosylation sites and sialylation were shown to be essential for this amphoteric nature and may be important for its function as an amphoteric biolubricator.

**Keywords**: Lubricin, mass spectrometry, EPAPTTPK, GALNTs, biolubricator

**ISBN:** 978-91-628-9206-7

# SAMMANFATTNING PÅ SVENSKA

Tillsammans utgör ledbrosk och ledvätska skavskyddet för att inte underliggande ben ska förslitas när vi rör oss. Detta åstadkoms genom att minska friktionen i leden. Ledvätskans smörjande effekt beror framförallt på biomolekyler förankrade till broskytan och inkluderar fosfolipider, hyaluronsyra och glykoproteinet lubricin. Vid ledinflammation och ledskada kan deras interaktion med broskytan förstöras så friktionen och slitaget av leden ökar och förvärrar sjukdomsbilden. Förändring har tidigare uppmärksammats för det mucin-liknande högglykosylerade lubricinet i samband med ledgångsreumatism och artros. Denna förändring förmodades härröra till lubricinets glykosylering, men kan även bero på molekylär nedbrytning av lubricinet. Detta, då det kan påvisas att lubricinfragment finns i ledvätskan vid ledinflammation. Dessa upptäckter tillsammans indikerar att lubricin kan användas som biomarkör för att identifiera ledsjukdomar. För att bättre förstå lubricinets smörjande förmåga och hur lubricin är inblandat i ledinflammatoriska sjukdomar inleddes en undersökning av lubricins kemiska struktur inkluderande dess glykosylering. Som ett led i undersökningen karakteriserades sialinsyre-innehållande lubricin oligosackarider då dessa är de huvudsakliga glykanerna på lubricin. Frisatta lubricin oligosackarider behandlades med $\alpha$2-3 specifikt sialidas enzym och produkternas MS/MS spektra jämfördes med referens spektra från den mass spektrometriska databases UniCarb-DB. Resultaten visade att core 2 strukturer på lubricin hade enbart $\alpha$2-3 bunden sialinsyra, medan core 1 strukturer även hade $\alpha$2-6 bunden. I ett annat försök visade glykopeptider analyserade med masspektrometri att mucindomänen innehållande serin, treonin och prolin (STP) och dess tandem repetera aminosyrasekvens EPAPTTPK var substrat för O-glykosylering. Vidare visade gentranskription av *N*-acetylgalactosaminyltransferaser *(GALNTs)* i fibroblast-liknande synoviocyter på vilka enzymer som troligtvis var ansvariga för att initiera glykosyleringen av lubricin. Analysen visade att inte bara de vanligast förkommande *GALNT1* och *GALNT2* var transkriberade, men även att de mer vävnadspecifika *GALNT5* och *GALNT15*, var högt utryckta. Resultaten från gentranskriptionsanalysen indikerade att glykosylering av lubricin kräver en unik kombination av transferas gener. Sammanfattningsvis visar den här studien att negativit laddad sialinsyra i mucin-liknande domäner gör lubricin amfotär pga dessa negativa kolhydratsidokedjor och dess positiva proteinkärnan som innehåller mycket arginin och lysin. Antalet glykosyleringsställen och antalet bundna sialinsyra visades vara av yttersta vikt för den amfotära zwitterjon egenskapen hos lubricin och kan vara av betydelse för dess smörjande funktion.

# LIST OF PAPERS

This thesis is based on the following published studies, referred to in the text by their Roman numerals.

I.    Ali, L., Kenny, D.T., Hayes, C.A., and Karlsson, N.G. (2012) **Structural Identification of *O*-Linked Oligosaccharides Using Exoglycosidases and MSn Together with UniCarb-DB fragment Spectra Comparison**. *Metabolites*. 2(4): 648-666.

II.   Ali, L., Jin, C., and Karlsson, N.G. (2012) **Glycoproteomics of Lubricin-Implication of Important Biological Glyco- and Peptide-Epitopes in Synovial Fluid**, In Rheumatoid Arthritis –Etiology Consequences and Co-Morbidities. *Intech*. 131-150

III.  Flowers, S.A., Ali, L., Lane, C.S., Olin, M., and Karlsson, N.G. (2013) **Selected reaction monitoring to differentiate and relatively quantitate isomers of sulfated and unsulfated core 1 *O*-glycans from salivary MUC7 protein in rheumatoid arthritis**. *Mol. Cell. Proteomics*. 12: 921-931

IV.   Ali, L., Flowers, S.A., Jin, C., Bennet, E.P., Ekwall, A.K., and Karlsson, N.G. (2014) **The *O*-glycomap of Lubricin, a Novel Mucin Responsible for Joint Lubrication, Identified by Site-Specific Glycopeptide Analysis**. *Mol. Cell. Proteomics*. Accepted

Reprints were made with the permission from publisher.

# CONTENTS

ii

iii

# ABBREVIATIONS

| | |
|---|---|
| Cosmc | ClGalT1-specific chaperone 1 |
| CXP | Collision cell exit potential |
| DP | De-clustering potential |
| EPI | Enhanced product ion |
| ETD | Electron transfer dissociation |
| FLS | Fibroblast-like synoviocytes |
| GalNAc | *N*-acetylgalactosamine |
| GlcNAc | *N*-acetylglucosamine |
| GalNAcol | *N*-acetylgalactosaminitol |
| *GALNT* | Polypeptide *N*-acetylgalactosaminyltransferase gene |
| LC-MS$^2$ | Liquid chromatography tandem mass spectrometry |
| Le$^{a/x}$ | Lewis a/x |
| *m/z* | Mass to charge ratio |
| $NaBH_4$ | Sodium borohydride |
| NaOH | Sodium hydroxide |
| OA | Osteoarthritis |
| PGM | Porcine gastric mucin |
| PVDF | Polyvinylidine fluoride |
| PGC | Porous graphitized carbon |
| PNGase F | Peptide: N-glycosidase F |

| | |
|---|---|
| PNA | Peanut agglutinin |
| pp-GalNAc-T | Polypeptide GalNAc-transferases |
| RA | Rheumatoid arthritis |
| ReA | Reactive arthritis |
| SDS-AgPAGE | Sodium dodecyl sulfate-agarose/polyacrylamide composite gel electrophoresis |
| SDS-PAGE | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| Ser | Serine |
| SF | Synovial fluid |
| SRM | Selected reaction monitoring |
| Thr | Threonine |
| WGA | Wheat germ agglutinin |

# 1 BACKGROUND

The diarthrodial joint is an excellent biological biobearing creation that facilitates motion of the bones. It has high wear resistance and a low coefficient of friction, primarily due to articular cartilage and synovial fluid (SF), which provide lubrication during movement. The articular cartilage, an elastic and porous material, reduces stress during joint loading and can also contribute to joint lubrication mechanisms [1]. SF, an ultrafilterate of blood plasma, serves as the primary biological lubricant in the joint. The joint diseases are primarily characterized by changes in SF and articular cartilage that result in the deformation of the joint assembly. This deformation causes increased friction and wear of articular cartilage leading to degradation of the cartilage and severe pain during joint movement. SF is composed of proteins, glycoproteins, proteoglycans and phospholipids. These molecules serve as lubricating molecule that, in combination, provide the essential boundary lubrication in the synovial joint (Fig. 1). The aim of this thesis was to investigate a mucin-like synovial glycoprotein, lubricin, essential for lubricating the synovial joint. The structural investigation of lubricin and its glycosylation may provide an insight into how the molecule provides lubrication in the joint.
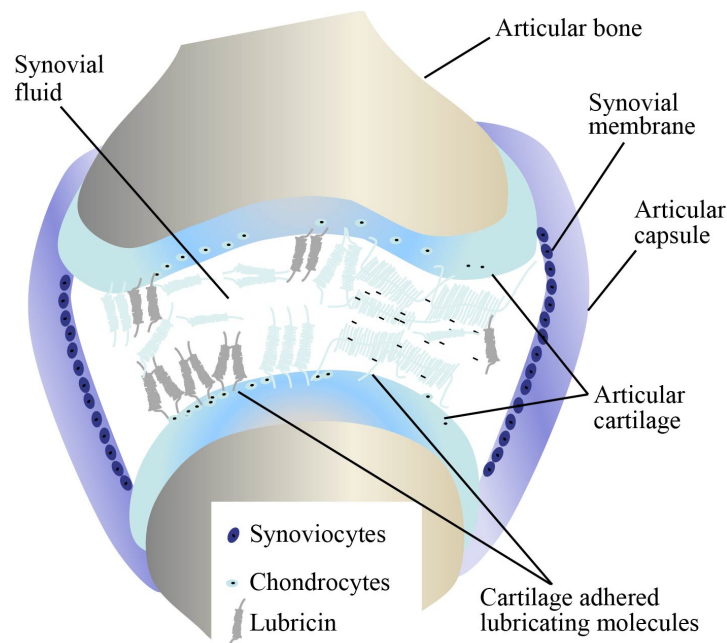
Figure 1. A schematic overview of the human synovial joint showing the cartilage adhered lubricating molecules, including lubricin, that provide lubrication during joint movement.

1

## 1.1 Bio-Lubrication

In nature, biological surfaces in motion are common. These surfaces are often surrounded by fluid-film, composed of biological molecules such as proteins, glycoproteins, lipids, polysaccharides, proteoglycans and phospholipids, and have developed efficient lubrication mechanisms. Under high pressure, these biomolecules in the fluid-film significantly increase in fluid viscosity thereby reducing friction and wear of the surface during motion [2]. This mechanism is different from water-based lubrication [3], as the viscosity of water does not rise significantly under pressure making it a poor lubricant.

Saliva has also been investigated in this thesis, mostly for method development as well as for its connection to joint degrading diseases. Saliva is composed of water and macromolecules (glycoproteins, electrolytes, lipids) and serves as a biological lubricant in the oral cavity [4]. It has been shown that the rheological (viscous) properties of saliva are strongly influenced by these macromolecules, particularly the high molecular weight mucin MUC5B [5]. However, compared to synovial fluid (SF), it is relatively less viscous [6]. Similarly to other biological lubricants, the macromolecules of saliva provide lubrication between tongue and tooth surfaces to facilitate speech [5].

The presence of biomolecules in the fluid-film can also enable the fluid to interact with the opposed fluid forming surfaces (in particular sliding surfaces) resulting in boundary lubrication [7]. In boundary lubrication, the biomolecules in the fluid often establish contact among themselves and with surface molecules or the underlying surface layer. This results in the adherence of fluid-molecules to the sliding surfaces forcing opposing surface apart and preventing surface contact during motion. In the synovial joint, the two opposing sliding surfaces and the attached biological molecules are in close contact almost constantly during motion. It is known that when these sliding surfaces are subjected to a high load, wear of soft cartilage surfaces can easily and quickly propagate [8]. During this condition, lubrication is provided by the synovial fluid macromolecules such as hyaluronic acid (HA), phospholipids and lubricin present in the interstitial gap separating the surfaces. It has also been documented that the rheological properties of the SF in the joint do not contribute significantly to the lubrication of the surfaces. [8]. This justifies the investigation and characterization of the SF macromolecules, in particular lubricin investigated in this thesis, in order to better understand joint lubrication.

2

## 1.2   Synovial Fluid (SF)

Synovial fluid (SF) is an ultrafiltrate of blood with additives that are produced by the cells located in the synovial lining (synoviocytes and chondrocytes). It surrounds the synovial joints and functions as a biological lubricant, responsible for the reduced friction and wear of articulating cartilage in synovial joints [9]. Due to the lack of blood supply to cartilage, synovial fluid is also responsible for cartilage nourishment by providing nutrients and removing catabolic products. Small molecules such as urea, amino acids and uric acid diffuse, relatively unrestricted, into the synovial fluid and thus exist in steady state equilibrium with plasma. However, macromolecules such as glycoproteins and globulins enter the synovial fluid from the plasma by restricted osmotic diffusion, or are produced by the synovial cells. Hence, the make up of the SF is not a pure reflection of the plasma. SF has a low coefficient of friction ($\mu$) and is specifically designed for protection of the cartilage [10].

SF is composed of macromolecules that, in combination, play a key role in lubricating the joint. These lubricating molecules include, lubricin (also known as proteoglycan 4, PRG4) [11] at a concentration of 0.05-0.35 mg/ml, hyaluronan (HA) at 1-4 mg/ml [12] and surface active phospholipids (SAPL) at 0.1 mg/ml [13]. Synoviocytes are the major source of surface active phospholipids and hyaluronan [14] secreted into the synovial fluid while lubricin is secreted by the synoviocytes as well as chondrocytes and, to a lesser extent, by the cells in the meniscus. The synovium is a thin layer surrounding the joints (~50 μm in humans), which is backed by a thick layer (~100 μm) of subsynovium and is responsible for the clearance of transported molecules. The cells of the synovium form a discontinuous matrix layer that contains collagen types I, III, V, chondroitin sulfate [15], biglycan, decorin proteoglycan [16] and fibronectin.

In healthy individuals, the joint and SF constitutes a system of reduced friction and wear which is in homeostasis. However, increased friction between the opposing articular cartilage surfaces has been suggested during disease and inflammation such as in arthritis and injury, which may be due to an altered SF composition. It has been documented that during acute arthritis, lubricin concentration decreases which results in increased friction [17]. A similar trend is seen in osteoarthritis where the concentration of SAPL, as well as hyaluronan, decreases [18]. However, the exact mechanisms behind SF composition changes remain unknown.

3

# 1.3 Osteoarthritis (OA) and rheumatoid arthritis (RA)

Osteoarthritis (OA) [19] and rheumatoid arthritis (RA) [20] are the two main arthritic diseases of the joint. Although both diseases result in the degeneration of the articular cartilage, the mechanisms of degradation are different with mechanical degradation in OA and chemical degradation in RA. Cartilage degeneration can be detected by glycoprotein fragments in the SF [21] (Paper II). Due to the limited efficacy of the available treatments, particularly for OA, understanding the biological factors related to arthritis is essential. OA is defined by the formation of osteophytes and the gradual loss of articular cartilage. The exact mechanisms of cartilage degradation is unclear; however, enhanced synthesis of matrix metalloproteinases (MMPs) has been suggested due to the resulting loss of collagen and proteoglycan fragments from the cartilage [19]. In addition to this, inflammatory cytokines such as interleukin (IL-1$\beta$, IL-17 and IL-18) and tumor necrosis factor alpha (TNF$\alpha$) also decrease the collagen synthesis and in turn increase the activity of MMPs that degrade collagen [22]. RA, on the other hand, is a chronic autoimmune inflammatory disorder that targets synovial tissues resulting in tissue remodeling and destruction [23]. Studies using tissue from RA patients and animal models of inflammatory arthritis indicated that inflammatory cytokines such as IL-1 and TNF$\alpha$ are responsible for the pathogenesis of abnormal tissue remodeling. It has been shown that these cytokines play a major role in enhancing the destructive capability of the inflamed synovial tissues [24, 25].

The joints of arthritis patients, both RA and OA, have indicated a change in the expression, degradation and glycosylation of lubricin [26] (Paper II). Patients with advanced RA show low levels of lubricin expression and studies conducted on OA animal models suggest that there is a relationship between pathogenesis and the downregulation of lubricin [27, 28]. This variation in lubricin expression exacerbates the disease by accelerating joint destruction. This suggests that characteristics of lubricin may be an effective indicator of disease progression in both RA and OA.

# 1.4   Synovial lubricin

Lubrication of the synovial joints under load requires the presence of molecules in the SF that lower the adhesion energy between opposing articulating cartilage surfaces. Lubricin is an abundant mucin-like glycoprotein (~227-345 kDa) considered responsible for the boundary lubrication of the synovial joints [11]. The molecule is encoded by the proteoglycan 4 gene and is synthesized in synovial fibroblast (synoviocytes) and superficial zone chondrocytes. The different translation products of the PRG4 gene are also referred to as superficial zone protein (SZP), megakaryocyte stimulating factor (MSF) precursor, camptodactyly arthopathy coxa vara pericarditis (CACP) protein, and hemangiopoietin (HAPO) [29]. The name lubricin was first applied to a lubricating glycoprotein (LGP-I) found in synovial fluid which represents the function of the protein with regards to lubrication and chondroprotection, as described by the PRG4 knockout mice model [30]. However, the protein is also synthesized in the superficial layer of articular cartilage [31], synovial lining, tendon and menisci [31-33] and is referred to as SZP. This makes lubricin a potential biomarker during inflammation.

Human synovial lubricin is composed of a 1404 amino acid apoprotein with a central mucin like domain (Fig. 2). This mucin domain is extensively *O*-linked glycosylated and contains 59 imperfectly repeated sequences of EPAPTTPK. The *O*-glycosylated mucin domain of lubricin is predominantly occupied by mono and di-sialylated core 1 and 2 structures [34]. The presence of a small amount of sulfated *O*-glycans in the mucin domain has also been confirmed [34]. It has also been shown that a strong repulsive hydration force, due to the presence of these negatively charged glycans, is responsible for the boundary lubrication of the protein to the cartilage surface [35]. The mucin domain is surrounded by a hemopexin (PEX) like domain at the *C*-terminus and two somatomedin B (SMB) like domains at the *N*-terminus. Lubricin lacking the terminal end domains has been shown to bind weakly to the cartilage surface resulting in in-efficient lubrication [36]. In addition to boundary lubrication, lubricin is also considered responsible for protecting the cartilage surface from protein accumulation and cell adhesion [30].
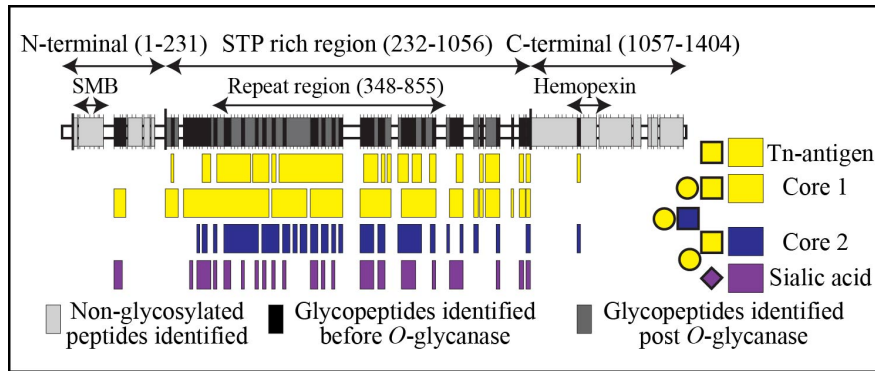
*Figure 2. Lubricin protein sequence with the glycosylated peptides identified in the STP rich region and non-glycosylated peptides (light grey) in the terminal domains before (black) and after (dark grey) partial de-glycosylation. All the identified glycans, including GalNAcα1- (yellow), core 1(yellow), core 2 (blue) and sialylated core 1 and 2 (pink), and their position in the protein sequence are also presented. The molecule is also composed of the somatomedin B (SMB) and hemopexin-like domains in the N- and C-terminal respectively.*

A down regulation of expression and glycosylation changes of lubricin have been shown in the joints of both RA and OA patients [26]. An association between this down regulation and pathogenesis has been suggested in OA animal models [28]. RA and OA are very different in disease etiology. However, in both, the cartilage degeneration detected by the deposition of proteoglycan fragments into the SF has been shown. Although several synovial joint specific markers have been identified in OA and RA patients such as calgranulin A, B and C [37], fructose bisphosphonate aldose A, fibrinogen β-chain, alpha-enolase, tenascin-C [38], serum amyloid A (SAA), C-reactive protein [39] and haptoglobin [28], a specific biomarker for early disease diagnosis is still required. The effectiveness of lubricin as a monitor of the state of the joint during disease has been little investigated despite the importance of lubricin as a synovial glycoprotein important to biolubrication. Therefore, the investigation of lubricin and its glycosylation, which is the main focus of this thesis, is essential in order to better understand its boundary lubricating properties.

## 1.4.1 The identified *O*-glycans on synovial lubricin

The *O*-glycans on synovial lubricin have been investigated previously in the group and the identified *O*-glycan structures are presented in figure 3 [34].

Core 1 *O*-linked oligosaccharides (Galβ1-3GalNAcα1-) and sialylated core 1 (NeuAcα2-3Galβ1-3GalNAcα1- and NeuAcα2-6)GalNAcα1-) were identified to be the predominant structures on lubricin. In addition to core 1, a low proportion of core 2 structures such as NeuAcα2-3Galβ1-3(NeuAcα2-3Galβ1-3/4GlcNAcβ1-6)GalNAcα1- were also found. The data also indicated that a small proportion of the core 2 structures were sulfated. The presence of sulfated core 2 structures, as sulfate was identified to be on the 6-position of the GlcNAc, makes lubricin a potential candidate for L-selectin binding, as previously reported [40]. The present investigation of *O*-glycans on synovial lubricin from arthritis patients identified a novel isomeric core 1 sulfation in addition to the previously reported structures (Paper III). Core 1 from arthritis patient was identified to have both Gal and GalNAc sulfation indicating a possible inflammatory type of role for this type of sulfation in the joint.
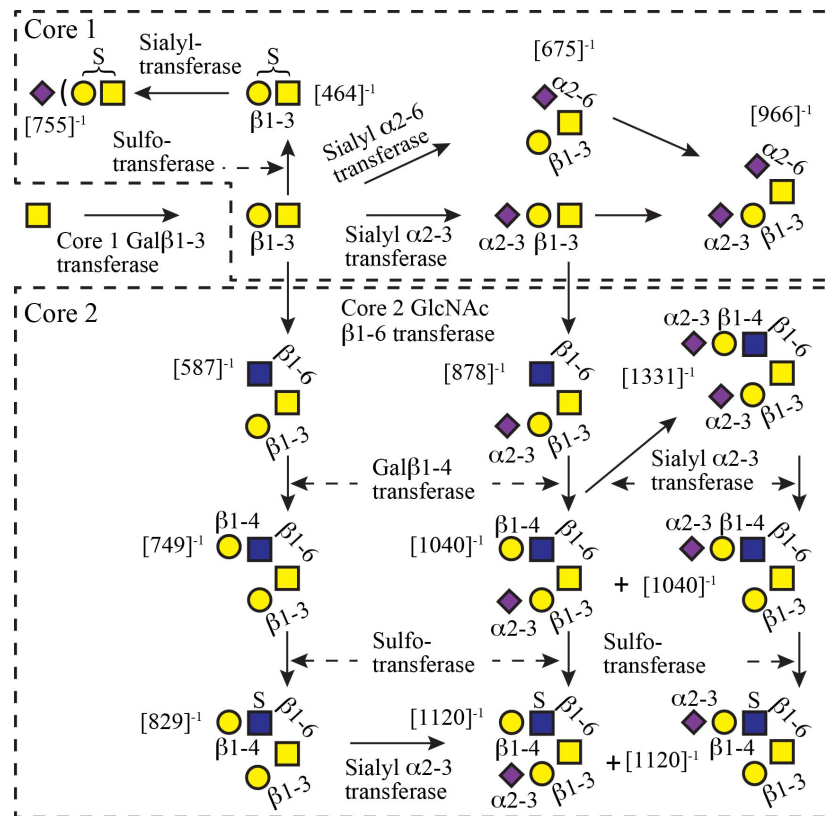


*Figure 3. O-glycan structures identified on synovial lubricin, including the novel core 1 sulfation.*

7

# 1.5   Glycosyltransferases for *O*-linked glycosylation relevant for synovial tissue

*O*-glycosylation is an important post-translational modification of secretory and membrane-bound glycoproteins. Mucin-type (or GalNAc) *O*-glycosylation of proteins, compared to other types (*N*-linked), is unique, as the prediction of the site of *O*-glycosylation is difficult. The main reason for this is the large, redundant UDP-GalNAc:polypeptide α-*N*-acetylgalactosaminyltransferase (ppGalNAc T's) gene family containing 20 gene encoded isoenzymes, all possessing unique and/or overlapping substrate specificities [41, 42]. Once the protein is fully folded, these ppGalNAc T's initiate *O*-glycan biosynthesis. The main area for initiation is in the Golgi apparatus and is started by transferring a GalNAc from the sugar nucleotide donor UDP-GalNAc to the hydroxyl group of a serine (Ser) or threonine (Thr) forming GalNAcα1-*O*-Ser/Thr (Tn antigen). Except proteoglycan (*O*-xylose) [43], all other types of *O*-glycosylation [including *O*-mannose, *O*-fucose, *O*-glucose and *O*-galactose (galactose added to hydroxylysine, Hyl)] as well as *N*-glycosylation of secretory proteins are initiated in the endoplasmic reticulum (ER). In contrast to mucin-type glycosylation, the biosynthesis of most other types of protein glycosylation is initiated by 1 or 2 oligo/glycosyl-transferase genes. The abundant *O*-GlcNAc (*N*-acetylglucosamine) glycosylation, catalyzed by a single enzyme in animals, occurs in the cytosol and nucleus. Even though this has recently been identified on extracellular proteins [44], this type of glycosylation is not usually found on proteins synthesized in the secretory pathway [45].

Mucin-type *O*-glycosylation has the potential to be differentially regulated in specific cell and tissue types because of this large family of up to 20 known homologous genes [42]. Genetic deficiencies in the GalNAc-T gene family may also contribute to differential regulation, which will ultimately lead to a change in the function of *O*-glycans. It is now well understood that deficiencies in genes that synthesize or initiate other types of protein glycosylation affects the overall glycosylation profile resulting in severe diseases. The main reason for this is the fewer glycosyltransferase genes (only 1 or 2) that initiate biosynthesis in most types of protein glycosylation, therefore alterations in these non-redundant pathways leads to devastating results. For example, in humans, mutations in protein *O*-mannosyltransferase T1 and T2 cause severe congenital muscular dystrophies [46]. In mice, mutations in protein *O*-xylosyltransferase 2 (xylosyltransferase T1 and T2

synthesize proteoglycans) results in a substantial decrease in proteoglycans that ultimately leads to polycystic liver and kidney disease [47].

As mucin-type *O*-glycosylation is initiated by a large family of *GALNT* genes, loss of a single gene may not affect the overall glycosylation profile or produce distinct phenotypes. This redundancy among isoforms was confirmed in earlier studies when several *GALNTs* were targeted ''knock out'' in mice but did not produce any visible phenotypes [48]. However, later studies demonstrated that individual GalNAc-T isoforms are important and can be essential for specific functions [49, 50]. It has been shown that *GALNT1* deficient mice exhibit defects in blood coagulation and lymphocyte homing [50]. In addition, *GALNT1* deficient mice resulted in reduce secretion of the major basement membrane proteins laminin and collagen IV and induce endoplasmic reticulum stress. This resulted in changes to epithelial cell proliferation, fibroblast growth factor signaling and organ growth [51]. *GALNT13* deficient mice showed a reduced GalNAc$\alpha$1-*O*-Ser/Thr (Tn antigen) expression in brain tissues but with no visible disease [52]. These studies indicate that some *GALNT* genes in mucin-type *O*-glycosylation are indeed essential for specialized functions and are not redundant.

## 1.5.1  Core 1 and 2 specific glycosyltranferases

GalNAc$\alpha$1-*O*-Ser/Thr (Tn antigen) serves as a precursor for the extension of *O*-glycan structures in mucin-type *O*-glycoproteins. The multiple glycosyltransferases transfer sugar in a stepwise manner from a sugar nucleotide donor to the acceptor substrate. It results in the formation of more complex *O*-glycan structures, which are further classified according to their core structures. In mucin glycosylation, eight different *O*-linked core structures have been described so far (Table 1) where core 1-4 are the most common in mucin glycoproteins and core 1 and 2 in other non-mucin glycoproteins [53].

*Table 1. Table of mucin O-linked core structures*

| Core | Structures |
|------|-----------|
| Core 1 | Galβ1-3GalNAcα1-*O*-Ser/Thr |
| Core 2 | Galβ1-3(GlcNAcβ1-6)GalNAcα1-*O*-Ser/Thr |
| Core 3 | GlcNAcβ1-3GalNAcα1-*O*-Ser/Thr |
| Core 4 | GlcNAcβ1-3(GlcNAcβ1-6)GalNAcα1-*O*-Ser/Thr |
| Core 5 | GalNAcα1-3GalNAcα1-*O*-Ser/Thr |
| Core 6 | GlcNAcβ1-6GalNAcα1-*O*-Ser/Thr |
| Core 7 | GalNAcα1-6GalNAcα1-*O*-Ser/Thr |
| Core 8 | GalNAcα1-3GalNAcα1-*O*-Ser/Thr |

Core 1 β3-galactosyltransferase (core 1β3-GalT or T-synthase) is a mammalian core 1 specific transferase. It transfers galactose from a sugar nucleotide donor UDP-Gal to Tn antigen on glycoproteins via a 1,3 linkage that results in the synthesis of the core 1 Galβ1-3GalNAcα1-*O*-Ser/Thr structure (Thomson-Freidenreich antigen) [54]. The T-synthase activity, identified in most cell types and mammalian tissues, depends on the expression of molecular chaperone Cosmc that resides in the endoplasmic reticulum (ER) [55]. Cosmc, essential for folding and stability of T-synthase [55], prevents proteosomal degradation of nascent T-synthase [56, 57] and can directly interact with denatured T-synthase *in vitro* to partially restore its activity [57]. Altered T-synthase activity has been suggested to play key roles in many diseases such as Tn syndrome and IgA nephropathy (IgAN).

Core 2 *O*-glycans are synthesized by the core 2 β-1,6-*N*-acetyglucosaminyltransferase (C2GnTs) family, which include C2GnT-1, C2GnT-2 or C2GnT-M, and C2GnT-3 [58, 59]. Once the core 1 structure is fully catalyzed, either of these C2GnTs transfers *N*-acetylglucosamine (GlcNAc) to a GalNAc residue on the core 1 structure via a 1,6 linkage that results in the synthesis of a core 2 branch [59]. In the case of the core 3 *O*-glycan (GlcNAcβ1-3GalNAcα1-), C2GnT-2 has also been shown to transfer GlcNAc to a GalNAc residue of the core 3 disaccharide via a 1,6 linkage, resulting in the core 4 structure [59]. Deficiency of C2GnTs is usually associated with distinct phenotypes. C2GnT-1 deficient mice have been shown to exhibit phenotypes with neutrophilia and a decreased biosynthesis of selectin ligands among myeloid cells [60]. This leads to a decreased recruitment of neutrophils to the site of inflammation and vascular disease pathogenesis. In addition, a partial reduction in the biosynthesis of L-selectin on high endothelial venules has also been shown in C2GnT-1 deficient mice,

10

resulting in reduced B-cell homing [61]. C2GnT-2 deficiency in mice can damage the mucosal barrier, which can lead to colitis and colon cancer [62]. In contrast, C2GnT-3 deficiency in mice is usually associated with reduced thyroxine levels in circulation [62].

## 1.5.2   Sialyltransferases

Sialyltransferases, unlike, glycosyltransferases direct the glycosylation pathway from glycan extension towards termination resulting in simplified sialylated *O*-glycan structures, a hallmark of various types of cancer. In lubricin, $\alpha$2-3 and $\alpha$2-6 sialyltransferases may be active, since only $\alpha$2-3 and $\alpha$2-6 sialylated structures have been identified. The $\alpha$2-3 sialyltransferase (ST3Gal) transfers *N*-acetylneuraminic acid (NeuAc) to the 3 position of galactose in core 1 (Gal$\beta$1,3GalNAc$\alpha$1-) and core 2 [Gal$\beta$1,3(GlcNAc$\beta$1,6)GalNAc$\alpha$1-] *O*-glycans respectively [63]. The sialylation of galactose on core 1 results in the synthesis of sialyl-T antigen (NeuAc$\alpha$2,3Gal$\beta$1,3GalNAc$\alpha$1-). In contrast, the $\alpha$2-6 sialyltransferases (ST6GalNAcI and II) compete with core 2 glycosyltransferases (C2GnTs) since both use the same core 1 acceptor substrate. These enzymes transfer NeuAc to the 6 position of GalNAc$\alpha$1- resulting in structures that lack a core 2 branch such as the structure Gal$\beta$1,3(NeuAc$\alpha$2,6)GalNAc$\alpha$1- and the sialyl-Tn antigen (NeuAc$\alpha$2,6GalNAc$\alpha$1-) [64]. The ST6GalNAcI has been suggested to be the predominant sialyltransferase in the synthesis of sialyl-Tn [64]. Increased expression of sialyl-T and sialyl-Tn antigen has been shown in gastric [65], breast [66], ovarian and pancreatic human carcinoma [67]. In addition, altered glycosylation of the MUC1 mucin in carcinoma has been shown to be associated with the expression of sialyl-Tn antigen. Both MUC1 and sialyl-Tn are used as targets for cancer immunotherapy [68, 69].

## 1.5.3   Sulfotransferases

Sulfotransferases are enzymes that can modify oligosaccharides with sulfate groups providing functionally specific negative charges to the oligosaccharides. They reside in the lumen of the Golgi apparatus and transfers sulfate from a coenzyme-3'-phosphoadenosine-5'–phosphosulfate (PAPS) to the particular monosaccharides residues that is to be sulfated in the structure [70]. Similar to oligosaccharides, in which a particular glycosyltransferase is required for the synthesis of a specific monosaccharide to a linkage, the addition of a sulfate to a particular monosaccharide and site

11

requires a specific sulfotransferase. Sulfotransferases recognize a specific monosaccharide residue, anomeric configuration or linkage position in the structure.

In synovial tissues and fluid, sulfation of *O*-glycans has been suggested to play a role in the pathology of RA, as sulfation was found to be the constituent of synovial lubricin found in the SF of arthritis patients [40]. In this study, both GlcNAc sulfation of core 2 (Paper II) and Gal, GalNAc sulfation of core 1 *O*-glycans were identified on synovial lubricin (Paper III). Core 1 sulfation was identified in RA patients while core 2 sulfation in OA patients. Sulfation of *O*-glycans is usually found on the C-3 and C-6 position of the monosaccharide residues [71, 72]. For core 1, both Gal and GalNAc have been identified to be sulfated (Paper III). Gal:3-*O*-sulfotransferases (Gal3ST) and 6-*O*-sulfotransferases (Gal6ST) are suggested to be responsible for the C-3 and C-6 sulfation of galactose respectively [71, 72]. This core 1 sulfation of Gal and GalNAc on synovial lubricin will be discussed further in the results section. Sulfation on core 2 structures has been shown to be on the 6 position of the GlcNAc as previously described by our group [40]. Sulfation on the 6-position of the GlcNAc on either extended core 1 or core 2 branched *O*-glycans makes lubricin a potential candidate for L-selectin binding [34, 40]. In the Golgi apparatus, two GlcNAc-6-sulfotransferases (GlcNAc6ST1 and 2) have been identified that can attach sulfate to the 6-position of the GlcNAc and are capable of generating L-selectin ligands [73]. In mice, the expression of GlcNAc6ST-1 has been shown in most tissues while GlcNAc6ST-2 was primarily shown in specialized blood vessels of the lymph nodes, where it plays a key role in lymphocyte circulation [74, 75]. In Chinese hamster ovary (CHO) and COS-7 cells, it has been shown that GlcNAc6ST-1 prefers to modify *N*-linked glycans while GlcNAc6ST-2 prefers to modify *O*-linked glycans [75]. Therefore, one may speculate that GlcNAc6ST-2 may be responsible for the GlcNAc sulfation on core 2 structures in synovial lubricin, as the molecule is identified to be heavily *O*-glycosylated mucin-like synovial glycoprotein.

Overall it is clear that changes in glycan transferases, leading to alterations in *O*-glycosylation, can have devastating and diverse effects on the body leading to a myriad of pathologies. This makes the improved understanding of these

glycans and enzymes essential, especially as we begin to understand more of glycan biosynthesis. Therefore, although this thesis is focusing on specific glycoprotein that are in themselves deserving of investigation, the development of methods to investigate these very highly *O*-glycosylated proteins is essential to a much broader field.

# 1.6   Electrospray ionization Mass spectrometry (ESI-MS)

Glycoproteomic analysis of mucins and mucin-like lubricin is a challenging task for the scientists in the glycoproteomic filed, as it involves the investigation of both the extensive *O*-glycosylation of the mucin domain and the protein component of the molecule. Considering the complexity associated with *O*-glycosylation and the importance of the lubricating role in the joint, there is a clear necessity for a suitable analytical technique capable of generating information both for the glycosylation as well as about the protein component of lubricin.

Mass spectrometry has developed into the most accepted analytical technique for the analysis of a number of biomolecules including glycoproteins [76]. A mass spectrometer (MS) is an analytical instrument used to determine the mass-to-charge ratio ($m/z$) of a charged molecule, where $m$ is the mass and $z$ is the charge state of the ion. It is one of the most sensitive and versatile analytical techniques since the interpretation of the data gives information both about the molecular weight of the ion as well as its composition and arrangement. The mass spectrometer is made up of three main parts (Fig. 4); an ion source used to ionize molecules, a mass analyzer operating in high vacuum used for gas phase separation of charged molecules and a detector to detect the charged molecule and record its $m/z$ value [77]. The mass spectrometer is also connected to a computer that allows the recording of the data, manipulating of the settings of the internal electronics and programming of the sample introduction modules. The work described in this report is based on the oligosaccharide and glycopeptide analysis of mucin-like synovial glycoproteins, lubricin, (Paper I-IV) by electrospray ionization coupled to a linear ion-trap-orbitrap (LTQ-Orbitrap) and QTRAP® mass spectrometer. The LTQ-Orbitrap is routinely used for qualitative analysis while the QTRAP® is mainly used for quantitative but can also be used for qualitative analysis, particularly as, the third quadrupole in the system serves both as a filter and as a linear ion-trap (LIT) [78, 79].
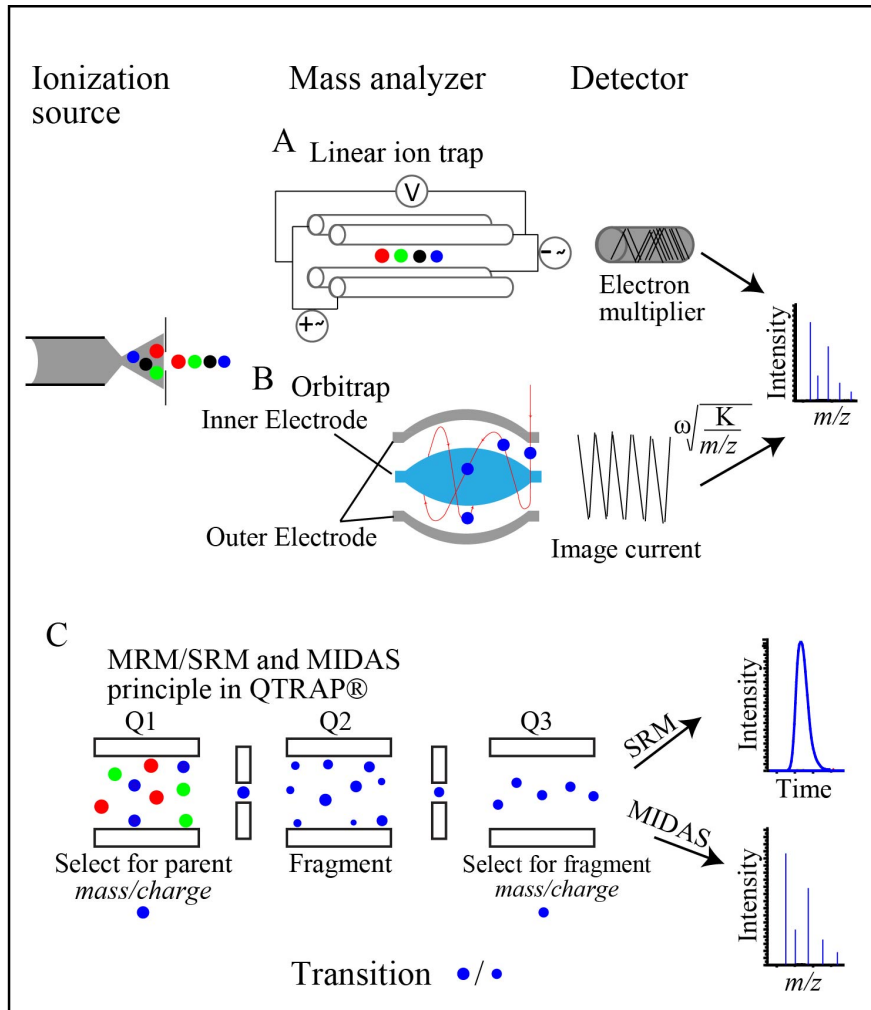
Figure 4. A mass spectrometer is made up of three main components: An ionization source, a mass analyzer, and a detector. An electrospray is a soft ionization technique suitable for biomolecules. The skimmer aligns the ions prior to entering the mass analyzer. The principle of (A) linear ion trap (B) orbitrap and (C) MRM/SRM with MIDAS in QTRAP® mass analyzer has been explained in the text. In linear ion trap, an increase in the ramped Rf voltage in-stabilize the trapped ions and the ions reach the detector (usually electron multiplier), which gives its m/z value. The orbitrap by itself acts as a detector as the ions oscillation is inversely proportional to its m/z value.

14

Ionization is a prerequisite for any molecule to be analyzed in a mass spectrometer, as it involves the addition or removal of a charge. Electrospray ionization (ESI) [80] and matrix assisted laser desorption ionization (MALDI) [81, 82] are the two most commonly used ionization techniques for generating gas phase ions in biological mass spectrometry. ESI was used as the ionization source in this report. In ESI, the ionization process occurs between the tip of the capillary column and the inlet of the mass spectrometer. A high potential difference (1-3 kV) is applied at the capillary tip, which results in the formation of a small liquid cone; referred to as a Taylor cone (Fig. 4). The liquid containing the analyte emerges from the capillary tip and is vaporized into a very fine spray of highly charged droplets sprayed towards the heated inlet of the instrument that results in the evaporation of the volatile mobile phase. The solvent evaporation leads to a decrease in droplet size and increase in the concentration of the charge at the droplet surface and the molecules become close to one another until the Rayleigh limit is reached. At the Rayleigh limit, the columbic repulsion overcomes the droplet surface tension and the droplet explodes into smaller droplets. This solvent evaporation and droplet explosion are continued until free ions are left which are introduced into the high vacuum region of the mass spectrometer [83]. Although this is the main theory of the process, an alternative theory suggests an ion evaporation model. This model assumes that when the droplets reach a certain size, charged gas phase ions are expelled from the droplets surface [83]. ESI leads to the formation of multiply charge ions by coupling of the gas or liquid chromatography separation column before the mass spectrometer [84].

## 1.6.1 Mass analyzer and detector

The mass analyzer separates charged molecules based on their *m/z* ratio. Before reaching the analyzer, the ions are first guided into a stable ion current, which is carried out by a set of parallel rods (4, 6 or 8) on which an oscillating potential is applied focusing the ions into its center trajectory. Linear ion trap compared to ion trap, has the advantage of significantly higher injection and trapping efficiencies, greater ion capacity and higher duty cycle. The LTQ-Orbitrap instrument used in the project is composed of two mass analyzer based on two different principles of ion separation, that is, a linear ion trap and an orbitrap. This type of instrument combines the benefits of fast scan rate, $MS^n$ scans, high duty cycle of the linear ion-trap and high resolution and high mass accuracy of the Orbitrap. Efficient ion transfers couple the two mass analyzers to each other.

15

In the linear ion trap, the ions are trapped in a four rod assembly (similar to a quadrupole) on which an oscillating electric potential is applied (Fig. 4). When sufficient ions are gathered in the ion trap, the ions are physically trapped inside the ion trap due to a fixed potential on the back plate of the trap and the potential on the front plate being raised. The ions within the trap oscillate at a stable trajectory. When a ramped RF voltage is applied, the oscillation is increased to instability causing the ions to leave the trap and reach the detector located peripherally to the trap region. Since selected RF voltages can be excluded, ions with a defined *m/z* can be isolated in the trap. Electron multiplier detectors are used to record the mass spectra, which operate on electron emission principle. The detector amplifies the impact of an ion into a bunch of secondary electrons that produce a small electric current. The number of secondary electrons released is dependent on the total number of ions of specific *m/z* hitting the detector simultaneously. The change in the RF potential allows the measurement of all trapped ions in a specific mass window resulting in a mass spectrum.

The orbitrap has become an instrument of choice in many biological applications due to its ability to deliver low-ppm mass accuracy and extremely high resolution; all within a time scale compatible with nano-liquid chromatography separation. In the orbitrap, the moving ions are trapped in an electrostatic field due to the applied potential on the center electrode. This electrostatic attraction is balanced by a centrifugal force arising from the initial velocity of the ion. The electrostatic field inside the orbitrap forces the ion to move in a spiral pattern and the ion start to cycle around the axial electrode in the center of a barrel shaped outer electrode (Fig. 4) [85]. The frequency of these oscillating ions is proportional to their *m/z*, which is detected as an image current which is transformed into mass spectra by Fourier transformation. The two mass spectrometers joined in the LTQ-Orbitrap can operate in parallel. The high resolution and high mass accuracy of the orbitrap can be used to identify the co-eluting precursor ion and the LTQ is used to sequence the peptide or glycopeptide by isolation and fragmentation approach. The selection of the ions for fragmentation is based on their intensity in the orbitrap. The average duty cycle for mass analysis and simultaneously fragmentation is around 1 second.

The analyzer region in a QTRAP® system [79] is based on the principle of a triple quadrupole (QqQ) routinely used for the quantification of target molecules in selected/multiple reaction monitoring mode (SRM/MRM) [86].

In an MRM scan, the first and third quadrupole act as a filter while the second acts as a collision cell used to fragment the ions filtered in the first quadrupole (Fig. 4C) [87]. The advantage of the MRM technique is increased selectivity, high S/N ratio, high accuracy and lower limit of quantification. However, in contrast to QqQ, the final quadrupole in the QTRAP® system can be operated both as a quadrupole and as a Linear Ion Trap (LIT) even during the same experiment. This additional functionality of the third quadrupole as LIT scan enhances the performance of the system for screening, identification and confirmation analysis as fragmentation spectra and $MRM^3$ can be performed on the same ions detected by the MRM scan. This feature is referred to as multiple reaction monitoring initiated detection and sequencing (MIDAS) which combines an MRM/SRM scan with a full scan linear ion trap MS/MS spectrum of the parent compound (Fig. 4C).

## 1.6.2 CID-$MS^n$ fragmentation of glycoconjugates

In biological MS, tandem mass spectrometry ($MS^2$) is routinely used for structural identification of molecules [88, 89]. $MS^2$ is the process of fragmentation of the parent ion that generates fragment or daughter ions. These daughter ions can be further fragmented and this is referred to as $MS^n$ (n= number of fragmentation). The generated $MS^n$ spectra are used to determine the composition and sequence of the molecule. This was the method of choice for structural and sequence analysis of oligosaccharides from mucins (Paper I and III) and mucin-like synovial lubricin (paper I-IV).

In MS, collision induced dissociation (CID) is routinely used for the fragmentation of oligosaccharides [90]. In CID, the parent ion isolated in the ion trap is accelerated by increasing their RF voltage, increasing the kinetic energy of the ion. This increase in the kinetic energy causes the ion to collide with the resident inert gas, usually helium. This collision converts some of the kinetic energy into internal energy resulting in fragmentation producing daughter ions. CID is particularly useful for sequence identification of glycans, either released (Paper I-IV) or attached to a peptide (Paper IV) and predominantly produces glycosidic and cross ring fragments. Domon and Costello developed a nomenclature to explain this fragmentation (Fig. 5). In this nomenclature, glycosidic fragments that contain the reducing end of the

oligosaccharides are designated $B_i$ or $C_j$ while fragments without the reducing end are designated $Z_j$ or $Y_j$ fragments.
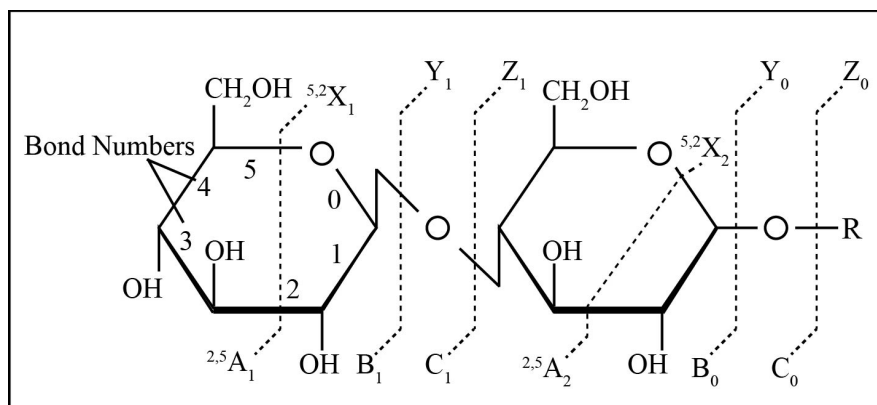


*Figure 5. Glycan fragmentation nomenclature described by Domon and Costello (adopted from Domon and Costello 1988).*

Subscripts i and j indicate the position of the subunit relative to the terminal ends of the oligosaccharides. Cross ring fragments with the reducing end of the oligosaccharides are labeled $A_i$ fragments while without the reducing end are labeled $X_j$ fragments. The specific cross cleavage is described as $^{k,i}A_i$ or $^{k,i}X_j$ where the superscript k and i indicate the particular carbon to carbon bonds that were cleaved across the cyclic sugar structure. In order to describe oligosaccharides branching; $\alpha$, $\beta$, $\gamma$ and $\delta$ are used where $\alpha$ is used for the largest branch, $\beta$ to the next largest branch and so on. The cross ring fragments (A/X ions) in the CID-MS$^n$ spectra has been used to explain the linkage and branching within the oligosaccharides [91]. These cross ring fragments were used to identify the terminal $\alpha$1-4GlcNAc epitope in paper II and sulfate position in paper III.

## 1.6.3 Identification of protein sequence and site-specific glycosylation

CID-MS$^n$ fragmentation of *O*-linked and *N*-linked glycopeptide generates sequence information both for the attached glycan (in MS$^2$) and for de-glycosylated peptide (in MS$^3$), but lacks the site-specific information of the modified amino acids [92]. This is due to extensive glycosidic bond cleavage of the precursor ion producing B/C and Y/Z ions (Domon and Costello

18

carbohydrate fragmentation nomenclature) [93] as illustrated in figure 6A. In $MS^3$, CID fragmentation of the de-glycosylated peptide induces peptide backbone cleavage producing mainly b- and y-type ions (Fig. 6A). In addition, the identification of the modified amino acids is even more difficult for peptides containing several Ser/Thr residues due to the lack of a consensus sequence for mucin-type *O*-glycosylation such as in lubricin. In MS, electron captured dissociation (ECD) [94] and electron transfer dissociation (ETD) are an alternative fragmentation techniques used for the site-specific characterization of protein post-translational modification including phosphorylation [95] and glycosylation [96]. Both techniques induce cleavage of the N-Cα bonds of the peptide backbone producing c- and z-type fragment ions while leaving the posttranslational modifications (PTMs) unaffected (Fig. 6B). ETD was used to fragment the glycopeptide generated from lubricin in order to identify the sites of *O*-glycosylation in the protein sequence (Paper IV) [76]. However, the advantage of ETD fragmentation in the present work was reduced due to the abundance of small repeats (EPAPTTPK) as its low mass reduced the higher charge state advantage of the ETD (Paper IV). This will be discussed further in the method section.
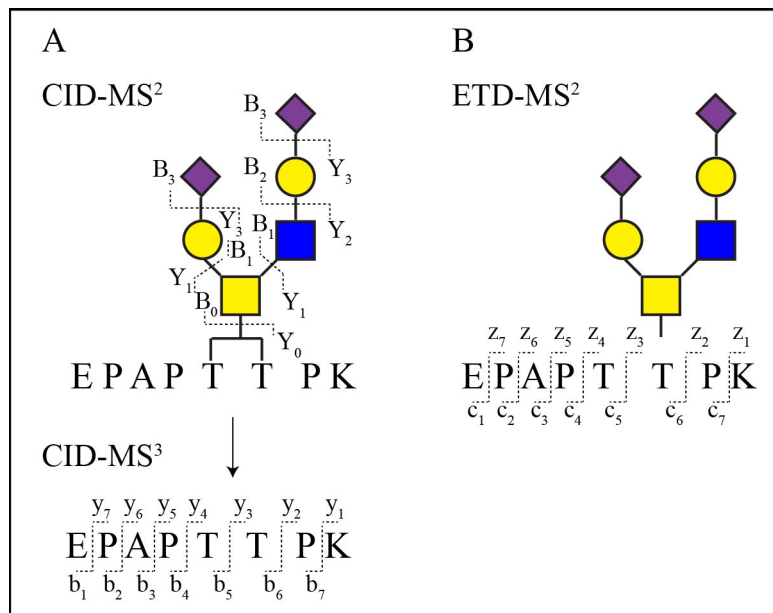


*Figure 6. CID/ETD-$MS^n$ fragmentations of an O-glycopeptide from lubricin. CID provides sequence information of the attached glycans in $MS^2$ and peptide in $MS^3$ while ETD provides site-specific glycosylation information in $MS^2$.*

19

In proteomics, a variety of software computational tools are available for automated processing of the $MS^n$ data that generates information about the protein. The online global proteome machine (GPM) [97], on-site version of the Mascot and Byonic software [98] with UniProt and NCBI human protein databases were used for MS data processing in order to identify the proteins in the negatively charged enriched fraction from human synovial fluid. In addition, Mascot and Byonic software were also used for $MS^n$ data (both CID and ETD) processing generated from glycopeptide analysis. In contrast to proteomics, structural assignment of glycans in glycomics still requires manual interpretation of the $MS^n$ spectra. The reason for this is the inherent complexity associated with glycans structures, which can have multiple branches and different linkage position. Similarly, the majority of the glycopeptide identifications were also based on the manual interpretation of the $MS^n$ spectra. This is due to a lack of sufficient glycopeptide information obtained when using software. Manual interpretation of the $MS^n$ data is an arduous and time consuming task that requires a high level of understanding to complete.

# 2 AIMS OF THE THESIS

## 2.1 Overall aim

The overall aim of this thesis was to develop methods for characterizing the difficult to analyze, mucin-like synovial glycoprotein, lubricin, and investigate its biolubricating ability in the joint.

## 2.2 Specific aims

1. Develop methods for structural elucidation of *O*-linked glycosylation to identify pathological glycobiomarker candidates.

2. Characterization of lubricin for identification of glyco- and peptide epitopes.

3. Site-specific glycopeptide characterization in order to identify the position of glycosylation in the protein sequence and its contribution to the amphoteric nature of lubricin.

# 3 METHODOLOGICAL CONSIDERATIONS

## 3.1 Materials and ethics

### 3.1.1 Human synovial fluid (Paper I-IV)

Human synovial fluid (SF) samples from arthritic patients were collected during therapeutic joint aspiration at the Rheumatology Clinic, Sahlgrenska University Hospital (Gothenburg, Sweden). All arthritic patients used in this study gave informed consent, and the procedure was approved by the Ethics Committee of the Gothenburg University Hospital. The arthritic patients fulfilled the American College of Rheumatology 1987 revised criteria for RA [99]. These samples were kindly provided by our collaborators, Anna-Karin H Ekwall and Lena Björkman, at the department of Rheumatology and Inflammation Research at Sahlgrenska Academy (Gothenburg, Sweden). The provided samples were clarified by centrifugation at 10,000 xg for 10 min and stored at -80 $^{o}$C before use.

Comments:

In the present work, human synovial fluid samples from arthritis patients (both RA and OA) were used to isolate lubricin for analysis. SF is composed of lubricating molecules, including lubricin, hyaluronic acid and phospholipids. SF provided by the collaborator were pre-labeled with patient number and disease stage but the patient identities were kept confidential. The advantage of using SF is that lubricin is an abundant mucin-like glycoprotein of SF present at a concentration of 0.05-0.35 mg/ml. Once isolated from SF, lubricin can be easily separated from hyaluronic acid and phospholipids through gel electrophoresis and gel filtration/size exclusion chromatography, since lubricin is much larger in size (~227-345 kDa) compared to the other two molecules.

### 3.1.2 Human tissues and cells (Paper IV)

Synovial tissue specimens were obtained from patients with rheumatoid (RA) (n= 2) and osteoarthritis (RA) (n= 2) during joint replacement surgery at Sahlgrenska University Hospital. Primary fibroblast-like synoviocytes (FLS)

22

cultures were established using collagenase/dispase with all cell counts for one million cells and used in passage 5. The cells were obtained in RNA*later®* solution. These RA and OA FLS cultures were kindly provided by our collaborator, Anna-Karin H Ekwall. The cell cultures were stored in -80 $^{o}$C before use.

Comments:

Primary fibroblast-like synoviocytes (FLS) cultures were used to investigate mRNA expression of the 20 known polypeptide *N*-acetylgalactosaminyltransferase (ppGalNAc-T) genes. The expression of mRNA in FLS cultures will suggest the types of *GALNTs* responsible for lubricin glycosylation, since FLS cultures have been shown to be one of the cell-types that secrete lubricin *in vivo*. One of the disadvantages was the passage number of the cells, since lubricin has only been identified in FLS medium up to passage 3 (data not shown). However, the mRNA expression of each *GALNT* may reveal the enzyme machinery in the cells and therefore the types and combination of *GALNTs* that are essential for glycosylation of proteins in these cells. Our collaborator Eric Paul Bennet performed real time PCR experiments on cDNA derived from RNA using high Capacity Reverse Transcription Kits.

### 3.1.3   Human saliva (Paper I and III)

A collaborator obtained the ethical approval from Gothenburg University Hospital for collecting RA patients' saliva. Whole saliva samples were collected from healthy and RA patients and from a patient diagnosed with sarcoidosis a minimum of 2 h after eating. It was centrifuged at 500 xg for 3 min, and saliva was removed from any remaining pellet and stored in -80 $^{o}$C before use.

Comments:

Saliva is a rich source of *O*-linked salivary mucin glycoproteins including MUC5B and MUC7. MUC5B is high molecular weight oligomeric glycoprotein and classified as a gel forming mucin while MUC7 is a low molecular weight monomeric glycoprotein. In this study, saliva is used both

as a source of mucin glycoproteins (Paper I and III) and also for investigating the types of glycosidases (Paper I) presented to the bacteria residing in the oral cavity where both harmful and beneficial bacteria use the oligosaccharide chains as nutrient source. Saliva also presented changes in the low abundant sulfated *O*-glycans of salivary mucins. This was further addressed in paper III where both healthy and arthritis patients salivary mucins were analyzed in order to investigate the inflammatory changes in the low abundant sulfated *O*-glycans of salivary MUC7 mucins.

## 3.2   Methodology

### 3.2.1   Purification of negatively charged proteins (Paper I-IV)

Purification of protein is the first step in determining the structure and function of any protein or glycoprotein such as lubricin, as it involves the isolation of proteins of interest from a complex mixture. In this study, acidic proteins, including lubricin were purified from SF as previously described by our group [34]. In detail, SF samples stored at -80 $^o$C were thawed and diluted in buffer A (20 mM Tris HCl pH 7.5, 10 mM EDTA) containing 250 mM NaCl (10 %, v/v i.e. 1 ml SF and 9 ml buffer). The diluted sample was applied to a 1 ml DEAE FF Hi-trap anion-exchange column (GE Healthcare Bio-sciences, Uppsala, Sweden) equilibrated with the same buffer using an AKTA FPLC chromatographic system (GE Healthcare). The negatively charged proteins were eluted with buffer A containing 1.0 M NaCl. The column was regenerated with buffer A containing 2.0 M NaCl and then re-equilibrated.

Lubricin containing fractions, identified by sandwich ELISA, were pooled and proteins precipitated with 80 % ethanol for 16 h at -20 $^o$C. The pellet after centrifugation at 16000 xg for 20 min was resuspended in phosphate buffered saline (PBS) pH 7.4. In addition to the precipitate, cut-of-filters were used for de-salting the pooled fractions, as it reduces the chance of protein denaturation by ethanol precipitation. Protein concentration was determined by BCA protein assay kit (Thermo Scientific, Rockford, IL, USA) using bovine serum albumin (BSA) as standard.

Comments:

One of the disadvantages of anion ion-exchange chromatography is that it purifies all negatively charged proteins and the protein of interest is usually present in a mixture of several other proteins. Therefore, enrichment of glycoproteins prior to further analysis (in particular to MS) is essential, since these unwanted proteins would make the analysis much more difficult and will also reduce the quality of the results. In this study, sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE) or SDS-agarose/polyacrylamide composite gel (SDS-AgPAGE) electrophoresis was performed to enriched lubricin (Paper I-IV) and high molecular weight salivary mucins MUC5B and MUC7 (Paper I and III) respectively. Gel electrophoresis separates proteins based on size, where the smaller molecular weight proteins run faster and travel further in the gel compared to large molecules resulting in protein bands in the gel. This efficiently enriched lubricin for further analysis.

## 3.2.2   LC-MS analysis of *O*-linked glycans (Paper I-III)

Changes in the *O*-glycosylation of mucin glycoproteins have shown to be associated with diseases. Therefore, in order to understand disease pathology, investigation of the *O*-glycosylation of mucin glycoproteins is essential. In contrast to *N*-linked glycosylation where enzymatic (PNGase F) release of glycans is possible, analysis of *O*-linked mucin glycosylation is generally performed after chemical release. The lack of a consensus sequence and the macro and microheterogeneity associated with extensive *O*-linked glycosylation often protect proteolytic cleavage sites. In this thesis, the oligosaccharides were released via a reductive β-elimination reaction. To perform the release, the enriched mucin glycoproteins were immobilized to a polyvinylidine fluoride (PVDF) membrane and incubated with 0.5-1 M $NaBH_4$ in 50 mM NaOH at 50 $^{o}C$ for 16 h. The released oligosaccharides were desalted with cation exchange chromatography to remove $Na^+$ ions and repeatedly extracted with 1 % acetic acid in methanol to remove borate salt as borate esters.

Unlike amino acid chains where there is only a single linkage position possible, both *N*- and *O*-linked oligosaccharides can form multiple branches (as many as 4 branches) and have different linkage positions (up to 6) with two anomeric forms (α or β). This results in the formation of isomeric structures making the separation of oligosaccharides prior ESI ionization and $MS^n$ fragmentation essential for structural elucidation. Online porous

25

graphitized carbon (PGC) chromatography is performed which separates oligosaccharides based on size and isomeric structure [100]. The PGC separated oligosaccharides, analyzed in negative ion mode, were isolated for CID-MS$^n$ fragmentation and the MS$^n$ spectra are manually interpreted for structural assignment. In order to evaluate the amount of degradation of the oligosaccharides during release (also known as peeling), major degradation products arising from the labile C-3 branch of GalNAc were monitored. The expected peeling products NeuAcα2-3Gal at *m/z* 470 (unreduced) and *m/z* 472 in negative ion mode were found to be close to the baseline indicating a negligible amount of glycan degradation during release.

Comments:

The oligosaccharides are reduced alditols when released via the reductive β-elimination reaction. One of the disadvantages with this method is that the released oligosaccharides cannot be labeled with, for example, 2-aminobenzamide (2-AB). However, it is still the preferred method as it reduces the destructive peeling of oligosaccharides that can occur with other release methods such as mild hydrazinolysis [101]. The reduced oligosaccharides are preferably analyzed in negative ion mode even though the MS$^n$ fragmentation of sialylated oligosaccharides gives less informative spectra due to labile sialic acid. The reason is the efficient ionization of oligosaccharide that is associated with their inherent negative charge sialic acid, hexuronic acid, and sulfates/phosphates residues. Oligosaccharides can also be analyzed in positive ion mode. In that case, derivatization of oligosaccharide such as per-*O*-methylation/acetylation is necessary as it generally facilitates the ionization process [102, 103]. Derivatization can also be used to investigate sialylated structures, as it can stabilize the labile sialic acid often lost during fragmentation in negative ion mode [102].

## 3.2.3 Structural assignment of oligosaccharides using MS$^n$ spectral matching (Paper I)

The MS$^n$ spectral annotation for structural assignment of oligosaccharides is an arduous and time consuming process as there are no effective software annotation tools available. However, in order to get high throughput structural characterization of oligosaccharides using MS, a quick annotation tool either software or manual is highly demanded. In this thesis, a reasonably quick MS$^2$

spectral intensity comparison method has been developed with the help of UniCarb-DB [104] (www.unicarb-db.com) for structural assignment. UniCarb-DB is a database of annotated *N*- and *O*-linked glycans. In this method, the relative intensity for each *m/z* value from the database peak list was downloaded for each isomeric structure with the same composition as the unknown and was matched with the corresponding relative intensity in the MS$^2$ spectra of the unknown within 0.5 Da. The sample peak lists were centroided with the Qual Browser 2.07 (Thermo Fisher) module and the matching exercise was performed manually using an excel spreadsheet containing MS$^2$ peak lists from the unknowns and from the database. The R$^2$ value (coefficient of determination) based on linear regression between matched intensity levels of MS$^2$ spectra of the unknown and from the unicarb database was used to score each match. The structures were ranked based on their score.

For identification of the terminal GlcNAc$\alpha$1-4 epitope, the matching protocol was performed for MS$^3$ spectral intensities of the $^{0,2}A_{1_\alpha}$ - H$_2$O fragment ion at *m/z* 304 (GlcNAc plus part of the cleaved Gal) obtained from un-identified oligosaccharide from porcine gastric mucin and from GlcNAc$\beta$1-4GlcNAc $\beta$1-4GlcNAc (Sigma Aldrich, St Louis, MO), and GalNAc$\beta$1-4Gal standards (DextraUK, Reading, UK) (Paper I).

The MS$^2$ spectral intensity comparison of the sialylated structures were inconclusive, as the different isomeric forms of the same structure gave similar R$^2$ value making the identification difficult. Negative ion LC-MS$^2$ spectra of the sialylated structures has been shown to be less informative due to labile sialic acid which may be the reason why they are not well assigned. This has been overcome by an alternative sialidase treatment and subsequent LC-MS$^2$ spectral intensity comparison of the generated neutral structure. The generated neutral structures were identified by an increase in the intensity after sialidase treatment indicating that the structures were the exoglycosidase product.

27

### 3.2.4 Investigation and relative quantitation of *O*-glycan isomers (Paper III)

The high throughput structural characterization of oligosaccharides is hindered by the time consuming nature and lack of resources for automatic processing of MS data. In addition, glycans can form isomeric structures that can further increase the complexity making the manual annotation much more difficult. This indicates that there is need of a more sensitive selected reaction monitoring (SRM) method for glycan analysis. Current glycan SRM methods have focused solely on *N*-glycans [105]. The work described in this thesis has focused on structurally annotating *O*-glycan isomers and relative quantitating sulfated and un-sulfated core 1 *O*-glycans from lubricin and salivary MUC7 of arthritis patients. PGC chromatography was performed to separate the isomers. Lubricin from a reactive arthritis (ReA) patient was used for creation and optimization of the sulfated core 1 transitions and *O*-glycans released from porcine gastric mucin (Sigma Aldrich) was used for the core 1 structure. Optimization of collision energy (CE), declustering potential (DP), and collision cell exit potential (CXP) was performed for each transition tested. The final method for SRM included the following transitions and specifications: 384.1/101.1 (CE -29 eV, DP -25, and CXP -15), 464/241.1 (CE -41 eV, DP -15, and CXP -20), and 464/302.1 (CE -40 eV, DP -15, and CXP -20) for core 1, Gal sulfation of core 1 and GalNAc sulfation of core 1 respectively. A dwell time of 35 ms was used for all. Enhanced product ion (EPI) spectra were acquired for the 384 and 464 ions, providing full scan MS/MS data from which sequence analysis could be performed. MultiQuant$^{TM}$ software version 2.1.1 (AB Sciex, Framingham, MA) was used for relative quantitation. My main contribution as a co-author to this publication was the discovery of sulfated core 1 structures in saliva and SF in acute inflammation.

### 3.2.5 CID/ETD-MS$^2$ glycopeptide characterization of lubricin (Paper IV)

The enzymatic/chemical release methodology is well established both for investigating glycosylation (both *N*- and *O*-linked) as well as for inflammation/disease specific changes in the glycosylation but give no, or very little information about the protein core and no information about the glycosylation site in the protein sequence. In contrast, traditional proteomic analysis will maximize protein information but will deduce minimal information about the glycan component of a glycoprotein. In order to

understand the nature of a glycoprotein, the glycopeptide approach performed in this study (Paper IV) is essential, as it not only defines the protein component or the glycan characteristics but also how these two essential component interact [106]. In this approach, the enriched SF fractions from arthritis patients (reduced and alkylated in SDS sample buffer) were separated in SDS-PAGE gel and the gel was stained with Coomassie brilliant blue. Lubricin protein bands (identified by anti-lubricin monoclonal antibody) were excised and subjected to in-gel endoproteinase lysine-C (Lys-C) and trypsin digestion at 37 $^{\circ}$C for 16 h each. Acetonitrile extracted peptides and glycopeptides from the gel were concentrated in a vacuum centrifuge.

Glycopeptides are usually a minor component of the peptide mixture in any glycopeptide experiment suggesting that enrichment is a pre-requisite before MS analysis. In this study, offline cotton wool HILIC SPE chromatography was performed as previously described by the Manfred Wuhrer group, which efficiently enriched the generated *O*-glycopeptides [107]. LC-MS$^2$ investigations were performed both before and after HILIC fractionation as well as on the HILIC flowthrough in order to identify the glycosylated and non-glycosylated peptide sharing the same peptide moiety. The glycopeptides were separated by online gradient C18 reverse phase chromatography. C18 separated glycopeptides, analyzed in positive ion mode, were isolated for both CID/ETD-MS$^2$ fragmentation and the MS$^2$ data was processed by the Mascot and Byonic software for sequence (both glycan and peptide) and site localization of glycan. The CID/ETD-MS$^2$ data was also interpreted manually, since very little information was obtained from software annotation.

Comments:

CID-MS$^2$ fragmentation used in this study produced glycan sequence information both for the different glycopeptides as well as for the different glycoforms of the same tryptic peptide (EPAPTTPK) indicating that lubricin display both macro- (two separate glycans) and site-specific micro-heterogeneity (different glycan at a single amino acid position). However, one of the disadvantages with CID-MS$^2$ fragmentation was the site-specific location of the glycan was not always possible to identify for peptides with more than one Thr or Ser, as it resulted in extensive glycosidic fragment ions. ETD was used as an alternative fragmentation technique for the site-specific identification of glycans, as it induces peptide backbone cleavage leaving the glycan unaffected [94, 108]. However, the abundance of the small tandem repeats (EPAPTTPK) in the protein sequence reduced the higher charge state

advantages of the ETD-MS$^2$ fragmentation in this study. Therefore, it was the novel combined use of CID and ETD that allowed the site-specific glycan localization as well as glycan determination of this difficult protein.

## 3.2.6 Investigating the role of sialic acid in the amphoteric nature of lubricin (Paper IV)

The predicted isoelectric point (pI) of apolubricin is very high and is close to 9.8. It is due to the high content of positively charged arginine and lysine residues in the protein sequence making the molecule basic. However, processed lubricin in the SF was purified as negatively charged proteins on DEAE column [34] indicating that glycosylation may have transformed the molecule into an acidic protein. Lubricin is a heavily glycosylated mucin-like glycoprotein and has been shown to have both *O*-glycan sulfation and sialylation. These sialylated and sulfated *O*-glycans can change the charge of the molecule suggesting that the glycosyltransferases, in particular sialyltransferases and sulfotransferases residing in the Golgi apparatus, may be responsible for this transformation of the molecule. Therefore, in order to understand the nature of the molecule and its biolubricating ability, it is essential to investigate the role of these negatively charged residues, in particular sialic acid, as lubricin is predominately composed of the sialylated core 1 structure.

In this study, the pI dependency on the sialic acid of the molecule before and after de-sialylation was determined. In brief, enriched SF proteins both before and after de-sialylation were separated in the immobilized pH gradient (IPG) gel of pH 3-10 (BioRad). Traditional 2D gel electrophoresis routinely used for investigating proteins in a complex mixture is not possible for mucin-like glycoproteins such as lubricin. The reason for this is the negatively charged glycan and the high content of acrylamide in the IPG gel that immobilize the molecule in the gel making the transfer to the second dimension difficult. Therefore, in order to examine the effect of sialic acid, proteins from the IPG gel were transferred to PVDF membrane via passive diffusion as previously described and subsequently probed with lubricin specific antibody (mouse anti-lubricin mAb13, Pfizer Research, Cambridge, MA, USA) for immunostaining.

### 3.2.7 Immuno/lectin investigation of lubricin glycosylation (Paper II)

Immunoassay with lectins or antibodies is a widely used technique for investigating the glycan content or specific protein in a sample. In this study, immunoassays with lectins and anti-carbohydrate antibodies were performed in order to identify the different glyco-epitopes of synovial lubricin. The SDS-PAGE separated negatively charged enriched SF proteins were electroblotted on to a PVDF membrane. The membrane was blocked and incubated with the appropriate concentration of primary antibodies or biotinylated lectins and subsequent incubation with secondary antibodies or sterptavidin labeled with HRP for detection.

Although, the immune/lectin-blotting method is convenient for investigating glycoepitopes, detailed structural information is commonly lacking. In addition, some glycoepitopes may not be detectable due to hindrance in space or lack of specific antibodies. For example, there is currently no antibody available that can distinguish 3-*O*-sulfation from 6-*O*-sulfation. Therefore, in order to maximize structural information, the glycosylation of lubricin was primarily analyzed by mass spectrometry.

### 3.2.8 mRNA expression analysis (Paper IV)

The prediction of the GalNAc-type (or mucin type) *O*-glycosylation such as on lubricin is difficult, as discussed earlier. Therefore, In order to understand the biosynthesis of lubricin glycosylation, it is essential to investigate the mRNA expression of the Golgi residing glycosyltransferase enzymes responsible for initiating *O*-glycosylation. Primary human fibroblast-like synoviocytes (FLS) cultures from RA (n= 2) and OA (n= 2) patients were used for RNA extraction and profiling of all the 20 genes was performed by our collaborator, Eric Paul Bennett. The FLS cultures are known to produce lubricin.

# 4 RESULTS AND COMMENTS

## 4.1 Identification of sialic acid configuration in lubricin (Paper I)

Over the last two decades, mass spectrometry has developed into the most accepted method for the analysis of a number of biomolecules including glycosylation [109]. Increased sensitivity combined with detailed high throughput structural characterization of oligosaccharides allows mass spectrometry to be routinely used for the structural elucidation of both *N*- and *O*-linked glycosylation [91, 110]. However, structural characterization of sialylated glycans (in released form) using negative ion LC-MS$^2$ is still challenging as the MS$^2$ spectra of sialylated glycans is less informative. The reason is the labile sialic acid leaving behind very little information suggesting that linkage specific sialidases should be used in order to increase the available information [111, 112]. In addition, the spectral annotation of the sialylated and sialidase-generated structure is a time consuming process, as there is no software annotation tool available. Therefore, LC-MS$^2$ spectral matching with the spectra reported in the UniCarb-DB in combination with specific sialidase method was used for high throughput structural and configurational analysis of sialylated structures in human synovial lubricin.

The enriched SF proteins were SDS-PAGE separated; the *O*-glycans from the lubricin band were released by reductive β-elimination and subsequently analyzed by LC-MS$^2$ in negative ion mode. The spectra were dominated by mono- and di-sialylated structures. However, in order to identify sialic acid linkage and subsequent LC-MS$^2$ structural investigation, the released glycans were incubated with sialidase S (*Streptococcus pneumonia*) specific for α2-3 linked sialic acid at 37 $^\circ$C for 16 h. A complete degradation of the mono- and di-sialylated core 2 structures was observed. This was accompanied with an increase in the intensity of the neutral ion at *m/z* 749 (Hex$_2$HexNAc$_1$HexNAcol) indicating that this is the sialidase-generated product. This was a core 2 structure with Galβ1-3(Galβ1-4GlcNAcβ1-6)GalNAcα1- when compared with spectra reported in the MS$^2$ database UniCarb-DB suggesting that this can be terminated with one sialic acid (on either branch) and with 2 sialic acid residues (one on each branche). The complete degradation of the mono- and di-sialylated core 2 structures with sialidase S revealed the configuration of the NeuAc moiety in both the

structures to be $\alpha$2-3 linked. The complete degradation of the isomeric sialylated core 1 (NeuAc$\alpha$2-3Gal$\beta$1-3GalNAc$\alpha$1-) structures was also shown (data not shown). However, the second isomer of the sialylated core 1 remained un-affected indicating that the core 1 structure has both $\alpha$2-3 (NeuAc$\alpha$2-3Gal$\beta$1-3GalNAc$\alpha$1-) and $\alpha$2-6 [Gal$\beta$1-3(NeuAc$\alpha$2-6)GalNAc$\alpha$1-)] linked NeuAc moiety. Thus it can be concluded that lubricin has both $\alpha$2-3 and $\alpha$2-6 linked sialylated structures.

The identification of $\alpha$2-3 and $\alpha$2-6 sialylated structures suggested that both the $\alpha$2-3 and $\alpha$2-6 linked sialyltransferases might be active on lubricin. To date, six mammalian genes have been identified forming the $\alpha$2-3 sialic acid linkage (ST3Gal-I-VI) that are potentially involved in the formation of the selectin ligands [113]. In some cancer tissues, the expression of ST3Gal-III has been shown to be associated with the formation of sLe$^x$ [114]. In studies using ST3Gal-I knock out mice, it has been shown that sialylation of core 1 galactose (Gal$\beta$1-3GalNAc$\alpha$1-) is primarily performed by ST3Gal-I [115] resulting in the synthesis of sialyl-T antigen (NeuAc$\alpha$2-3Gal$\beta$1-3GalNAc$\alpha$1-). In CD43 and CD45 T cells, ST3Gal-I expression has been shown to reduce the formation of core 2 *O*-glycans [116] essential for the interaction of T cell surface glycoproteins with endothelial selectins.

In contrast, the $\alpha$2-6 sialyltransferases (ST6GalNAc-I and II) transfers NeuAc to the 6 position of GalNAc$\alpha$1- resulting in structures that lack the core 2 branch such as [Gal$\beta$1,3(NeuAc$\alpha$2,6)GalNAc$\alpha$1-] and sialyl-Tn antigen (NeuAc$\alpha$2,6GalNAc$\alpha$1-) [64]. The ST6GalNAc-I has been suggested to be the predominant sialyltransferase in the synthesis of sialyl-Tn [64]. In gastric tissues, ST6GalNAc-I has been shown to be the major enzyme that controls the expression of the cancer associated sialyl-Tn antigen [117]. ST3GalNAc-I and ST3GalNAc-II show similar *in vitro* activity, when Tn antigen alone is available. However, in the presence of both T and Tn antigen, ST3GalNAc-I prefers Tn while ST3GalNAc-II prefers T antigen [64]. Aberrant expression of sialyl-T and sialyl-Tn antigen is a characteristic feature observed in gastric [65], breast [66], ovarian and pancreatic human carcinoma [67]. In addition, altered glycosylation of the MUC1 mucin in carcinoma has been shown to be associated with the expression of sialyl-Tn antigen. Both MUC1 and sialyl-Tn are used as targets for cancer immunotherapy [68, 69].

## 4.2 Identification of lubricin fragments as possible biomarker candidates (Paper II)

Although osteoarthritis (OA) (mechanical degradation) and rheumatoid arthritis (RA) (chemical degradation) are different in disease etiology, the degeneration of the articular cartilage is the hallmark of both diseases [118]. This degeneration, detected by glycoprotein fragments in the SF, is suggested to be due to the detachment/degradation of the biomolecules that are responsible for lubricating the joint surfaces. Lubricin is suggested to be one of the molecules responsible for lubricating the joint [119]. Despite its important lubricating function, lubricin has not been investigated as a method to monitor the state of the joint. In RA, a decrease in expression and increased degradation of lubricin has been shown. A decreased level of lubricin, particularly in the early stage of the disease, has also been shown in studies using animal models of OA [28]. This reduced levels and increased degradation of lubricin may accelerate the destruction of the joint, exacerbating the disease. Therefore, proteomic characterization of lubricin for identification of possible proteolytic fragments in the SF is essential and may serve as a potential marker for early diagnosis of OA and RA diseases.

Although the proteomic characterization of the negatively enriched SF component has been investigated previously, lubricin (or its proteolytic fragments) has only been reported in a very few studies [34, 120]. In this study, lubricin in the enriched SF proteins were investigated by proteomic methods. For this, the dominating band in the SDS-PAGE (~227-345 kDa area) was in-gel trypsin digested and subsequently LC-MS$^2$ analyzed. The proteomic analysis identified 28.5% of the protein sequence and is believed to be the full-length processed secretory lubricin. It is due to the heavily $O$-glycosylation of the processed protein that account for 50 % of the molecular weight of the protein (apolubricin is around 151 kDa). The sequence of all exons were detected except exon 1, consisting of the 24 amino acid signal peptide. However, in order to identify proteolytic fragments of the molecule, it is essential to not only investigate the major lubricin band but also other areas of the gel, in particular the low mass region. In this study, the entire gel lane was trypsin digested and analyzed by LC-MS$^2$. Interestingly, the molecule was also identified in all other areas of the gel. In lubricin, 4 possible native cleavage sites were indicated two on either side of the mucin domain.

Is there any association between degradation of lubricin and inflammation of joint diseases? Or is the fragmentation of lubricin in healthy individuals, if any, different from arthritis patients? SF of OA patients has been reported to contain fragments of lubricin [121]. In addition, the *C*-terminal fragments of the molecule were also identified in OA SF. However, in the same study no fragmentation was observed when the SF of healthy individuals was investigated [120]. The fragments of lubricin identified in this study were from RA SF indicating that this may very well be due to the inflammation and pathological degradation of the molecule in the joints of arthritis patients. Therefore, the identification of lubricin fragments in this report may open up a new possibility for disease specific biomarker discovery.

# 4.3 Identification of novel isomeric core 1 sulfated structures on Lubricin and salivary MUC7 (Paper III)

Altered protein glycosylation (both *N*- and *O*-linked) has been observed in a large number of disease states including cancer that can affect glycoconjugates, including the heavily *O*-glycosylated mucins and mucin-like synovial glycoprotein, lubricin. Mucins are usually found on the mucosal and sliding surfaces where the *O*-glycosylation serves to protect the underlying layer and provide lubrication. In a demanding environment, mucins undergo a change in glycosylation, as it is the first line of defense against pathogens and the surrounding environment. This change has been observed in several types of diseases [122], such as breast cancer [123], gastric cancer [124], RA and OA [125]. However, in order to understand this change in glycosylation, investigation of the glycosylation of a protein in different disease states, including the healthy, is essential.
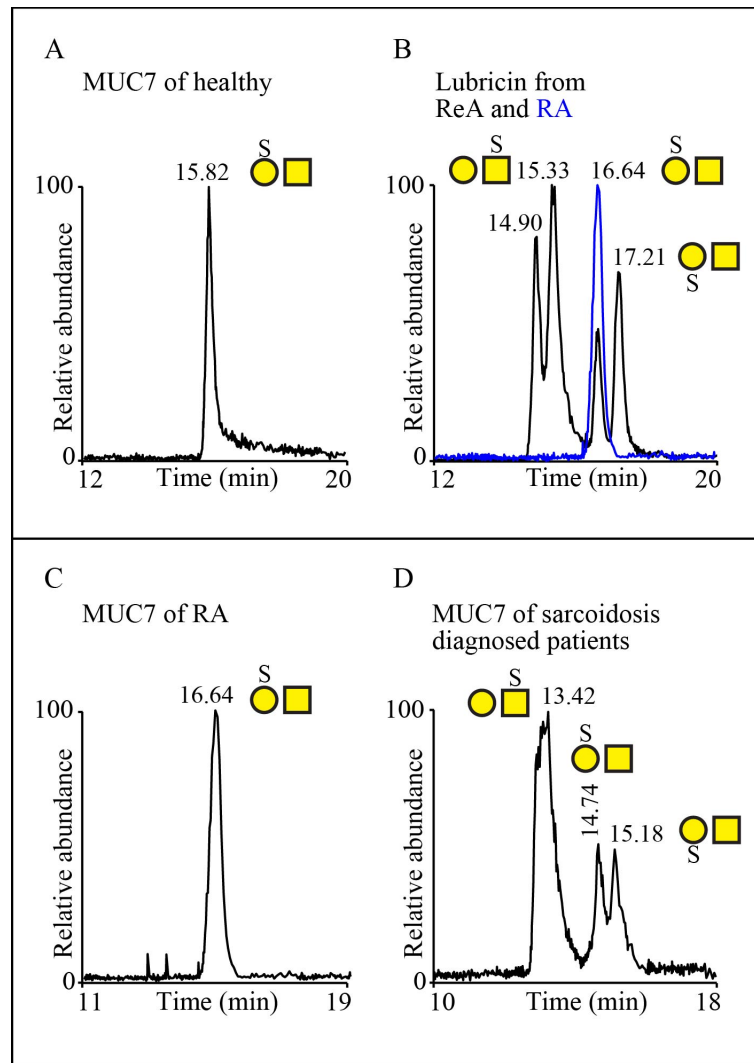
*Figure 7. Core 1 sulfation of O-glycans on synovial lubricin and salivary MUC7 glycoproteins. Lubricin isolated from SF of ReA (B) and salivary MUC7 from patient diagnosed with sarcoidosis (D) displayed more complex pattern of core 1 sulfation compared to single isomer in RA lubricin (B, blue) and in healthy (A) and RA patient salivary MUC7 (C).*

In this study, the *O*-glycosylation of synovial lubricin isolated from arthritic patient, including rheumatoid arthritis (RA) and reactive arthritis (ReA) was

investigated. RA is a systemic chronic autoimmune inflammatory disease while ReA is an acute autoimmune condition and is agreed to be caused by molecular mimicry, resulting in cross-reactive antibodies from gastrointestinal or genitourinary tract bacterial infections [126]. The $O$-glycans were released by reductive β-elimination and subsequently analyzed by LC-MS$^2$ using PGC chromatography. In addition to the previously reported $O$-glycans [34], novel isomeric sulfated core 1 $O$-glycans on lubricin were identified in this study, which were investigated further. Sulfation of core 2 $O$-glycans on lubricin has been reported previously and this was suggested to play a role in the pathology of RA [40]. However, sulfation of core 1 $O$-glycan was identified for the first time on synovial lubricin in this report (Fig. 7). In contrast to a single sulfated core 1 isomer in RA, a more complex pattern of core 1 sulfation was observed in ReA patient, as PGC (separates isomeric structures) chromatography identified 4 different isomers in ReA patient. This indicates that sulfation may have different roles in the acute ReA and chronic RA conditions investigated [127]. Sulfation of $O$-glycans is usually present on the C-3 and C-6 position of the monosaccharide residues [72]. The diagnostic cross ring fragment ions in the MS/MS spectra of the sulfated core 1 structure in RA patient indicated the C-3 sulfation of galactose. C-3 sulfation of Gal is usually catalyzed by the Gal3STs (4 Gal3STs have been cloned so far) [128]. However, in addition to C-3 sulfation of galactose, ReA patient also indicated C-6 sulfation of galactose and GalNAc, including an additional un-assigned position. C-6 sulfation of Gal/GalNAc core 1 $O$-glycans has been reported previously on MUC1 glycoprotein secreted from human breast cancer cell line [72]. In the same study, Gal6ST, the enzymes responsible for C-6 sulfation of terminal Gal, was suggested to be involved in the C-6 GalNAc sulfation of core 1 $O$-glycans. This suggests that this Gal6ST may also be responsible for the C-6 sulfation of GalNAc on synovial lubricin, as C-6 sulfation of core 1 is also identified on lubricin indicating that both the Gal3STs and Gal6ST may be active. The LC-MS data also contained an un-assigned isomer. The most likely explanation for this un-assigned sulfated core 1 isomer may be the C-4 position of the GalNAc, since the C-3 position of the GalNAc is already occupied. C-4 sulfation of GalNAc is known in chondroitin sulfate (CS) and dermatan sulfate (DS) [129, 130]. However, lubricin is identified to be a mucin-like glycoprotein and is devoid of CS/DS [40]. However, the location of sulfate in this isomer was not assigned, and further investigation will be required to identify the location of the sulfate as well as associated enzymes.

Given the identification of the different sulfated core 1 $O$-glycans in the two arthritic conditions, the investigation of this structure was further studied in other tissues and fluids of the body. Saliva, due to the non-invasive sampling procedure and rich source of $O$-linked mucin glycoproteins, was considered

37

suitable for analyzing this sulfation further. In addition, sulfation has been shown to be a constituent of the salivary *O*-glycoproteins suggested to be important for the clearance of oral bacteria and for the recruitment of neutrophils [131]. Therefore, saliva was collected from healthy and RA patient and from a patient diagnosed with sarcoidosis. At the time of sample collection, the sarcoidosis-diagnosed patient was using anti-inflammatory steroids as treatment supplement but the treatment of the ReA patient was unknown. Whole saliva samples were SDS-AgPAGE separated and *O*-glycans were released from salivary MUC7 glycoprotein. The identification of the sulfated core 1 *O*-glycans on MUC7 confirmed this glycoprotein as an appropriate sample to investigate this core 1 sulfation. In healthy and RA patients, only C-3 sulfation of galactose was identified in sulfated core 1 *O*-glycans. However, similar to lubricin isolated from ReA, MUC7 of sarcoidosis patient also indicated more complex pattern of core 1 sulfation. It showed all the three assigned isomers (C-3 sulfation of Gal and C-3 and C-6 sulfation of GalNAc) of sulfated core 1 *O*-glycans on lubricin from ReA patient. However, the fourth isomer was not identified in MUC7 or may be co-eluting in the first peak. This similarity of core 1 sulfation between acute ReA and sarcoidosis-diagnosed patient may be due to the immune reactive nature of core 1 sulfation. However, in order to find an association between core 1 sulfation and immunological response, this study needs to be expanded further to a larger number of patients.

# 4.4 Lubricin: an amphoteric mucin-like molecule (Paper IV)

Negatively charged sialylated and sulfated glycans can alter the charge of heavily glycosylated mucins and the mucin-like synovial glycoprotein, lubricin. Apomucins are usually neutral or acidic, secreted with a predicted pI of 2-4.7 [36]. However, apolubricin is exceptional with a pI predicted to be as high as 9.8. This is due to the large number of positively charged arginine and lysine residues in the *N*- and *C*-terminal regions making the polypeptide backbone positively charged. However, the secreted protein is purified as an acidic glycoprotein indicating that glycosylation, in particular sialylation and to some extent sulfation, in the central Ser, Thr and Pro (STP) rich region have changed the molecule, as lubricin is suggested to have a majority of sialylated *O*-glycans [40]. The glycosylation (in particular sialylation)

dependency for the charge and pI of lubricin model suggested that, at the physiological pH (7.2-7.4) of the SF, the STP rich region required approximately 60 sialic acids in order to be negatively charged. Beyond 80 sialic acid residues, lubricin and its STP rich region are both negatively charged and capable of maintaining its negative charge during pH shifts of SF and/ or limited chemical/enzymatical agents that partially lower the sialic acid content of lubricin or degrade the protein. This could possibly be the number of sialic acid residues that is required for its function on the cartilage.


Isoelectric focusing before and after de-sialylation was carried out in order to understand the contribution of sialic acid to the physical properties of lubricin. The pI of the molecule before de-sialylation ranged from 4-7.5, while after de-sialylation the pI was found to be approximately 7.5. This also indicates that de-sialylation has changed the molecule from highly acidic to basic where in addition to *N*- and *C*-terminal, the mucin domain was also positively charged. This suggests that lubricin is an amphoteric mucin-like molecule in its fully glycosylated state. The lower pI compared to the predicted apolubricin may possibly be due to the remaining sulfated residues on lubricin glycans. Lubricin is suggested to be a good lubricant for negatively charged surfaces such as the outer most articular cartilage layers (lamina splendens) [132]. This may be due to an increase in repellent charge force between the negatively charged STP rich region and negatively charged components (HA, lipids and proteoglycans) of the outer most cartilage layer. However, lubricin lacking the terminal somatomedin B-like and hemopexin-like domains has been suggested to cause in-efficient lubrication due to weak binding at the cartilage surface [133]. Therefore, one can speculate that the amphoteric nature of lubricin may be essential for efficient lubrication.
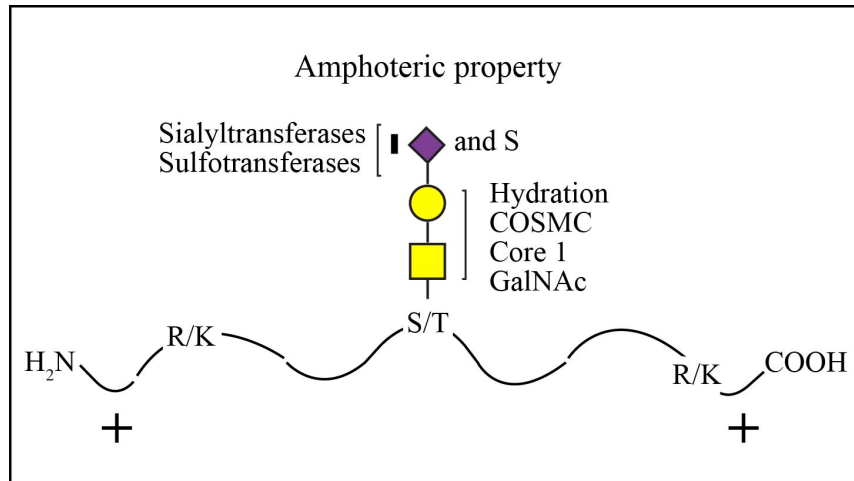
*Figure 8. Lubricin is an amphoteric mucin-like molecule, with both the number of glycosylation sites and the amount of sialic acid and sulfate is essential for its amphoteric nature.*

The amphoteric property of lubricin is controlled by the number of glycosylation sites and the amount of negatively charged sialic acid and sulfate attached (Fig. 8). The glycosylation sites are controlled by the ppGalNAc-Ts responsible for initiating mucin-type *O*-glycosylation [42]. In addition, the active expression of core β1-3galactosyltransferase and ER resident molecular chaperone Cosmc is also essential, as lubricin is predominantly shown to have core 1 *O*-glycans that can be sialic acid terminated [34]. The sulfation of both core 1 and core 2 *O*-glycans has also been identified on lubricin. This suggests that both sialyltransferases and sulfotransferases are active in lubricin producing cells. This active expression is essential, as they can alter the property of the molecule by imparting negative charges. The sialylated core 1 *O*-glycans on lubricin has been shown to be responsible for lubricating the joint [119]. In addition, the abundance of negatively charged glycans on lubricin will increase the hydration of lubricin and will contribute to SF lubrication [134]. This hydration generates strong repulsive hydration forces that can also contribute to the protein boundary lubrication, as this repulsion reduces friction by preventing surface contact during motion.

The substantial change in the pI and the drastic alteration of the charge of lubricin around 60-70 sialic acids indicates that there is a critical point where the number of glycosylation sites and the amount of sialylation, and to some extent sulfation, will significantly alter the properties of lubricin. All together this shows that the pathological alteration of the glycosylation of lubricin may contribute to an altered lubricating surface of articular joints.

Overall, the approach adopted in this thesis has successfully investigated both the glycosylation and the protein component of the heavily glycosylated mucin-like glycoprotein, lubricin. This approach has not only identified the types of glycosylation and its position in the protein sequence, but also indicated its likely contribution to the physical property of the molecule. The identification of novel sulfated core 1 *O*-glycans suggests that lubricin is an immune reactive molecule and can respond to external stimuli. A substantial increase in the pI with de-sialylation indicated that lubricin is an amphoteric mucin-like molecule suggesting that sialylation and to some extent sulfation is essential for its amphoteric nature.

41

# 5  OVERALL DISCUSSION

Lubricin, in combination with other SF macromolecules, has been suggested to be one of the molecules responsible for lubricating the joints [119]. The joints of rheumatoid arthritis (RA) and osteoarthritis (OA) patients have shown a decrease in expression and changes in glycosylation of lubricin [26]. This decrease in lubricin expression exacerbates the disease by accelerating joint destruction indicating that certain characteristics of lubricin may be an indicator of disease progression in RA and OA. Despite being an important lubricator in the joint, the detailed structural characterization of lubricin has not been performed. This is due to its large size and extensive mucin-like *O*-glycosylation that is suggested to protect the proteolytic cleavage sites. In this thesis, the structural investigation of lubricin and its glycosylation was performed that may explain the lubricative properties of the molecule and its response to pathology.

Structural investigation revealed that lubricin predominately contains mono-sialylated core 1 *O*-glycans, as previously reported [34]. A small proportion of core 2 and mono-sialylated core 2 *O*-glycans have also been confirmed on lubricin. In addition, different isomeric structure of the novel sulfated core 1 *O*-glycans were identified in acute reactive arthritis (ReA) compared to a single isomer in chronic rheumatoid arthritis (RA). This indicates that lubricin is an immune reactive molecule and can respond to immunological stimuli. The molecule has been shown to be a good lubricant for negatively charged surfaces (lamina splenden) [132]. This is due to repulsive forces between negatively charged glycan residues and negatively charged component of the cartilage (HA, lipids and proteoglycans). Water molecules in the SF will hydrate the negatively charged sialylated and sulfated residues (though differently as sulfate does not have the poly-ol groups) resulting in the formation of hydration layers around the charged residues. These hydration layers generate strong steric hydration repulsion when the molecules surrounded by the hydration layers come close to each other [134]. These hydration repulsive forces will result in reduced friction and wear of the cartilage during motion. The large number of positively charged lysine and arginine in the protein sequence makes the polypeptide backbone positively charged. This positive charge may be involved in the adherence of lubricin to negatively charged components (hyaluronic acid, proteoglycans) of the outermost cartilage layer, as the *N*- and *C*-terminal recombinant construct has been shown to promote the attachment of cells to the cartilage [30].
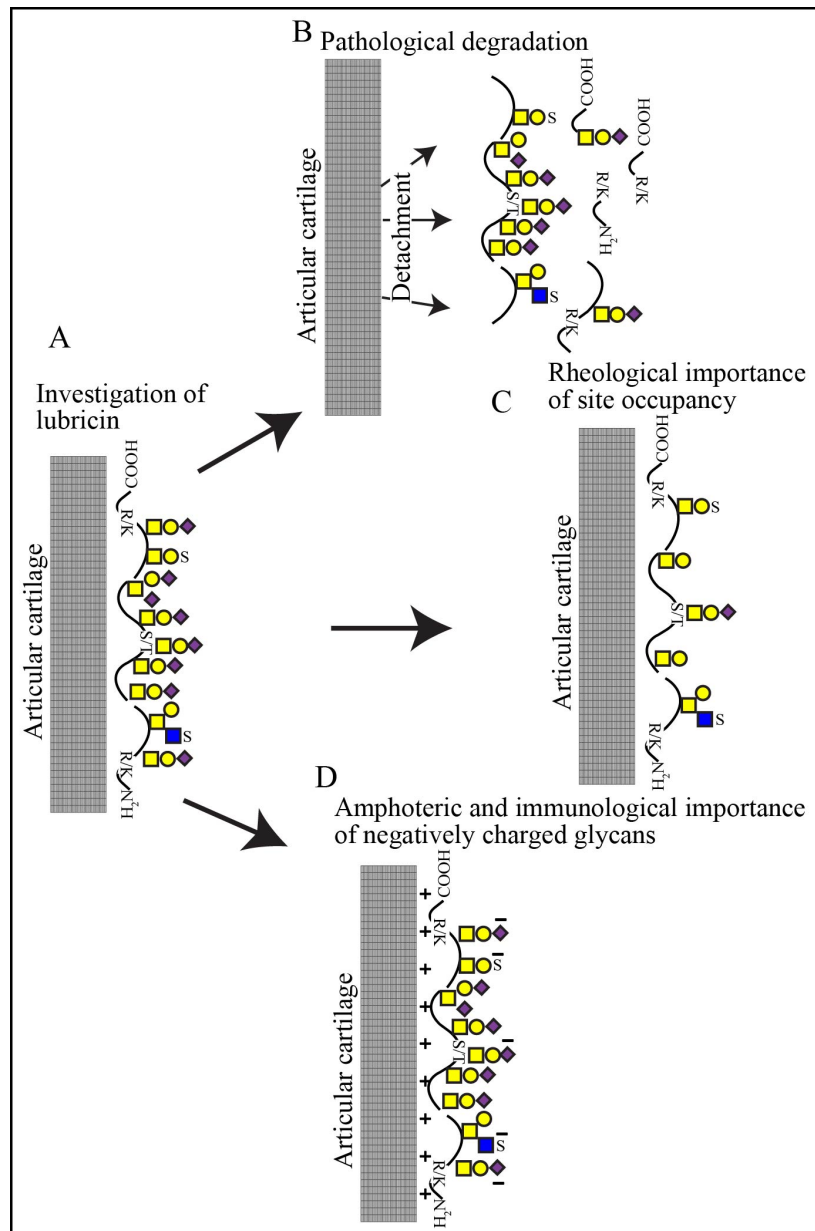
*Figure 9. The investigation of lubricin suggested these three likely hypotheses. A) The pathological degradation of articular cartilage may be due to the pathological detachment of lubricin from the cartilage surface. B) It is the number of glycosylation sites that may control the rheological property of the molecule. C) The negatively charged sialylated and sulfated glycans may not only be important for the amphoteric nature of the molecule but may also have immunological functions.*

The increase in stress and protease activity associated with the pathology of arthritic diseases, both RA (chemical degradation) and OA (mechanical degradation), are suggested to result in the degradation of the articular cartilage [19, 20]. This degradation is due to the detachment of the cartilage adhered lubricin and other lubricating molecules that are responsible for joint lubrication (paper II). In this thesis, the identification of lubricin fragments in the SF of RA patients may open up a new possibility for disease-specific biomarkers (Fig. 9B). In mucins, the heavily *O*-glycosylation of the mucin domain is suggested to protect the proteolytic cleavage sites. However, in the case of lubricin, it was extensively degraded by trypsin. Other proteases including neutrophil elastase, papain and pronase have also been shown to degrade lubricin [17, 135]. The identification of tryptic peptides in the low mass 30-65 kDa region may be the *N*- or *C*-terminal of the molecule with a small portion of the mucin-like domain, as the non-glycosylated *N*- and *C*-terminal are approximately 35 kDa each. However, it is not clear whether they were from unique cleavages along the protein sequence or randomly excised *in vivo*. The identification of *C*-terminal fragments of lubricin has also been shown in the SF of OA patients [120]. However, in the same study, no fragments were observed when the SF of healthy individuals was investigated. This suggests that this fragmentation of lubricin may be associated with the inflammation or pathology of joint diseases.

The drastic alteration in the charge of lubricin with the addition of 60-70 sialic acid residues suggests that sialic acid may play an important role in the property of lubricin. The lubricative property of the molecule has shown to be not solely dependent on sialic acid, as it has shown to be less affected with de-sialylation [119]. In the same study, the lubricative property of lubricin was shown to be predominately associated with the galactose of core 1 structure, as a drastic decrease in lubrication was observed with the removal of this residue. However, an alternative explanation to this may be that it is the bulk of glycosylation, which may be important for the lubricative property of the molecule (Fig. 9C), as lubricin is suggested to be predominately composed of the galactosylated core 1 structure. This may also explain the steady change in the charge of lubricin up to 50 sialic acid residues indicating that a minimum number of glycosylation sites are essential in order to observe a visible change in the charge of lubricin. The large number of glycosylation sites on lubricin (168 sites were identified in this thesis) will allow an increase of the negative charge density resulting in changing the charge of the molecule, as all these glycosylation sites can potentially terminate with sialic acid or sulfate. The greater number of negatively charged sialic acid residues (and sulfate also)

will attract more water molecules in the SF resulting in the increased hydration lubrication mechanisms. This increase hydration will generate strong steric hydration repulsive forces keeping the surface apart and resulting in reduced friction during motion [134].

Isoelectric focusing (pI) before and after de-sialylation was carried out in order to further understand the contribution of sialic acid to the properties of lubricin. The pI of the molecule before de-sialylation ranged from 4-7.5 while after it was approximately 7.5 indicating that de-sialylation changed the molecule from acidic to a basic molecule where in addition to the *N*- and *C*-terminal, the mucin domain of lubricin also bacame positively charged. This is due to the large number of positively charged lysine and to some extent arginine residues in the polypeptide backbone. The substantial change in the pI suggests that lubricin is an amphoteric mucin-like molecule (Fig. 9D) where the mucin domain, due to its hydration caused by the substantial glycosylation of its negatively charged glycans, may be important for lubrication while the positively charged *N*- and *C*-terminal may be important for adherence to the cartilage surface [132, 136]. The negatively charged sialylated and sulfated glycans are also important for the immunological properties of lubricin, as they can result in the formation of immune reactive epitopes such as sLe$^x$ and sulfation of *O*-glycans [40]. These epitopes make lubricin a potential candidate for L-selectin binding [40].

# 6 CONCLUSION

The methods developed in this thesis have successfully investigated both the glycosylation as well as the protein component of a heavily glycosylated mucin-like synovial glycoprotein, lubricin. Screening of sialic acid linkage configuration suggested that lubricin has predominately 3-linked sialylated *O*-glycans. However, a small proportion of 6-linked sialylated structures have also been identified. The identification of novel isomeric sulfated *O*-glycan in arthritic diseases indicated that lubricin is an immune reactive molecule and can change in response to a pathological environment. This also suggested that these immune reactive sulfated *O*-glycans have a different role in the two arthritic conditions investigated. The identification of proteolytic fragments in the SF of RA patients indicated the chemical/immunological stress and protease activity associated pathological degradation may result in the detachment of the molecule from the cartilage surface. The investigation performed in this thesis has redefined an extended serine, threonine and proline (STP) rich region in lubricin compared to the previously defined mucin domain. The developed methods have successfully mapped the glycosylation profile of lubricin within this STP rich region indicating that lubricin glycosylation displays both micro- and macroheterogeneity. Overall the current thesis showed that lubricin is an amphoteric mucin-like molecule, where both the number of glycosylation sites and sialylation are important for the amphoteric nature and may very well be critical for its function as an amphoteric biolubricator.

# 7  FUTURE PERSPECTIVES

Although the results from this thesis have shed some light on the characterization and biolubricative properties of synovial lubricin, including its response to diverse pathological environments, several new questions arise that need to be addressed.

In this study, novel isomeric sulfated *O*-glycans (sulfation of both Gal and GalNAc) on lubricin have been identified in the two different arthritic conditions. Altered expression of sulfotransferases has been reported in various inflammatory diseases. Therefore, in order to understand the pathological association between sulfotransferases and inflammation, the investigation of the type of sulfotransferases involved in *O*-glycan sulfation on lubricin and their expression in inflammation is essential and needs to be investigated further.

The mRNA expression analysis of the polypeptide *N*-acetylgalactosaminyltransferase genes (*GALNT*) suggested that *GALNT15* is the most abundant enzymes in primary human fibroblast-like synoviocytes, one of the cell types that secrete lubricin *in vivo*. Little is known about this enzyme in regards to substrate specificity. Therefore, further investigation of this and other expressed enzymes is essential in order to understand the type of enzymes that are involved in the glycosylation of lubricin.

The investigation of lubricin in the current thesis suggested the likely three hypotheses (Fig. 9). These include the pathological degradation of the molecule, the rheological importance of site occupancy and the immunological importance of negatively charged glycans and their contribution to the amphoteric property of lubricin. However, further investigation is required in order to understand these hypotheses.

# ADDITIONAL BIBLIOGRAPHY

1. Kenny, D.T., Ali, L., Issa, S., and Karlsson, N. G. (2011) **Glycomic Analysis of Membrane-Associated Proteins, In Sample Preparation in Biological mass Spectrometry**. A.R. Ivanov and A.V. Lazarev, Editors. *Springer*, 498-513.

2. Miliotis, T., Ali, L., Palm, J. E., Lundqvist, A. J., Ahnoff, M., Andersson, T. B., and Hilgendorf, C. (2011) **Development of a Highly Sensitive Method Using Liquid Chromatography-Multiple Reaction Monitoring to Quantify Membrane P-Glycoprotein in biological Matrices and relationship to Transport Function**. *Drug. Metab. Dispos* 39, 2440-49.

3. Wang, M., Raj, A., Liu, C., Claesson, P. M., Ali, L., Karlsson, N. G., and Dedinaite, A. **Molecular Synergies in biolubrication**. Submitted to Soft Matter.

4. Flowers, S. A., Kalamajski, S., Ali, L., Liu, C., Raj, A., Björkman, L. I., Eisler, T., Raj, J. R., Aspberg, A., Heinegård, D., Dedinaite, A., Karlsson, N. G., and Jin, C. **Cartilage oligomeric matrix protein forms protein complexes with synovial lubricin on the cartilage surface**. Submitted to Arthritis & Rheumatology.

48

# ACKNOWLEDGEMENT

This thesis would have not been possible without the support and contribution of a large number of people that I would like to take the opportunity to thank.

First of all thank to **Almighty Allah**, the creator and sustainer of the universe for giving me the strength and patience to fulfill my commitment.

**Niclas Karlsson** (**special thanks)** for making me part of your group and for intruducing me to the world of Glycomics/Glycoproteomics. It was your tutorship and guidance which taught me that ''the road may be rough, the journey may be tough and the experience may be difficult, but they are just stairs to a bright future''. Thanks for all the help in writing papers and thesis, this was a memorable time being a student. Apart from science, I thank you for all the encouragment and laughs that you provided.

My co-supervisor **Gunnar Hansson** and **Malin Backström** thank you for your help and support during these years. You guys have provided all the support whenever I asked for.

**Niclas Karlsson's** group with past and present members: Thank you all for making this a memorable time**. Sarah Flowers** (**special thanks**) for all your help, support and for your guidance of writing nice papers and off course for all the help in writing and correcting my papers and thesis during your holidays. Thanks for providing a nice cover image for this book (It is very pretty), for all the nice discussion and for all your concern whenever I leave for my home country. You are the best friend. **Diarmuid**, **Shake ´n Bake**, for your help in preparing my manuscript and for all the discussion and for being such a nice and caring friend. **Jin** for helping in the lab, preparing manuscript and always been willing to help when asked. **Samah** for teaching phosphoproteomics sample prep and for all the help in Denmark. **Catherine Hayes** for helping in manuscripts, nice discussion and good music in outside activities during these years. **Barbara** for your encouragment, for providing free transport with nice discussion and for always being willing to help. **Varvara** for all the nice discussion, coffee breaks and for your willing to help, and off course for booking a place for party. **Jayesh** (not for discussing cricket!! just kiding), **Miguel**, **Stefan**, **Jaison** and **Rebecca** for their encouragment. **Jessica** for all your encouragment, good wishes and off course for translating the abstract into Swedish language.

# REFERENCES

1.          Mow V. C. and Guo X. E. (2002) Mechano-electrochemical properties of articular cartilage: their inhomogeneities and anisotropies. *Annu Rev Biomed Eng* 4:175-209

2.          Celli J., Gregor B., Turner B., Afdhal N. H., Bansil R. and Erramilli S. (2005) Viscoelastic properties and dynamics of porcine gastric mucin. *Biomacromolecules* 6:1329-33

3.          Lee S. and Spencer N. D. (2008) Materials science. Sweet, hairy, soft, and slippery. *Science* 319:575-6

4.          Humphrey S. P. and Williamson R. T. (2001) A review of saliva: normal composition, flow, and function. *J Prosthet Dent* 85:162-9

5.          Wickstrom C., Christersson C., Davies J. R. and Carlstedt I. (2000) Macromolecular organization of saliva: identification of 'insoluble' MUC5B assemblies and non-mucin proteins in the gel phase. *Biochem J* 351 Pt 2:421-8

6.          Christersson C. E., Lindh L. and Arnebrant T. (2000) Film-forming properties and viscosities of saliva substitutes and human whole saliva. *Eur J Oral Sci* 108:418-25

7.          Hills B. A. (2000) Boundary lubrication in vivo. *Proc Inst Mech Eng. H J Eng Med* 214:83-94

8.          Lee D. W., Banquy X. and Israelachvili J. N. (2013) Stick-slip friction and wear of articular joints. *Proc Natl Acad Sci USA* 110:E567-74

9.          Jay G. D., Torres J. R., Warman M. L., Laderer M. C. and Breuer K. S. (2007) The role of lubricin in the mechanical behavior of synovial fluid. *Proc Natl Acad Sci USA* 104:6194-9

10.          Oates K. M., Krause W. E., Jones R. L. and Colby R. H. (2006) Rheopexy of synovial fluid and protein aggregation. *J R Soc Interface* 3:167-74

11.          Swann D. A., Silver F. H., Slayter H. S., Stafford W. and Shore E. (1985) The molecular structure and lubricating activity of lubricin isolated from bovine and human synovial fluids. *Biochem J* 225:195-201

12.          Mazzucco D. and Spector M. (2004) The John Charnley Award Paper. The role of joint fluid in the tribology of total joint arthroplasty. *Clin Orthop Relat Res*:17-32

13.          Schwarz I. M. and Hills B. A. (1998) Surface-active phospholipid as the lubricating component of lubricin. *Brit J Rheumatol* 37:21-6

51

14.        Momberger T. S., Levick J. R. and Mason R. M. (2005) Hyaluronan secretion by synoviocytes is mechanosensitive. *Matrix Biol*: 24:510-9

15.        Worrall J. G., Wilkinson L. S., Bayliss M. T. and Edwards J. C. (1994) Zonal distribution of chondroitin-4-sulphate/dermatan sulphate and chondroitin-6-sulphate in normal and diseased human synovium. *Ann Rheum Dis* 53:35-8

16.        Coleman P., Kavanagh E., Mason R. M., Levick J. R. and Ashhurst D. E. (1998) The proteoglycans and glycosaminoglycan chains of rabbit synovium. *Histochem J* 30:519-24

17.        Elsaid K. A., Jay G. D., Warman M. L., Rhee D. K. and Chichester C. O. (2005) Association of articular cartilage degradation and loss of boundary-lubricating ability of synovial fluid following injury and inflammatory arthritis. *Arthritis Rheum* 52:1746-55

18.        Mazzucco D., Scott R. and Spector M. (2004) Composition of joint fluid in patients undergoing total knee replacement and revision arthroplasty: correlation with flow properties. *Biomaterials* 25:4433-45

19.        Lorenz H. and Richter W. (2006) Osteoarthritis: cellular and molecular changes in degenerating cartilage. *Prog Histochem Cyto* 40:135-63

20.        Ungethuem U., Haeupl T., Witt H., Koczan D., Krenn V., Huber H., von Helversen T. M., Drungowski M., Seyfert C., Zacher J., Pruss A., Neidel J., Lehrach H., Thiesen H. J., Ruiz P. and Blass S. (2010) Molecular signatures and new candidates to target the pathogenesis of rheumatoid arthritis. *Physiol Genomic*s 42A:267-82

21.        Goggs R., Carter S. D., Schulze-Tanzil G., Shakibaei M. and Mobasheri A. (2003) Apoptosis and the loss of chondrocyte survival signals contribute to articular cartilage degradation in osteoarthritis. *Vet J* 166:140-58

22.        Siwik D. A., Chang D. L. and Colucci W. S. (2000) Interleukin-1beta and tumor necrosis factor-alpha decrease collagen synthesis and increase matrix metalloproteinase activity in cardiac fibroblasts in vitro. *Circ Res* 86:1259-65

23.        Goldring S. R. (2003) Pathogenesis of bone and cartilage destruction in rheumatoid arthritis. *Rheumatology* 2:11-6

24.        Van den Berg W. B. (2002) Lessons from animal models of arthritis. *Curr Rheumatol Rep* 4:232-9

25.        Danfelter M., Onnerfjord P. and Heinegard D. (2007) Fragmentation of proteins in cartilage treated with interleukin-1: specific cleavage of type IX collagen by matrix metalloproteinase 13 releases the NC4 domain. *J Biol Chem* 282:36933-41

26.        Elsaid K. A., Fleming B. C., Oksendahl H. L., Machan J. T., Fadale P. D., Hulstyn M. J., Shalvoy R. and Jay G. D. (2008) Decreased lubricin concentrations and markers of joint inflammation in the synovial fluid of patients with anterior cruciate ligament injury. *Arthritis Rheum* 58:1707-15

27.        Young A. A., McLennan S., Smith M. M., Smith S. M., Cake M. A., Read R. A., Melrose J., Sonnabend D. H., Flannery C. R. and Little C. B. (2006) Proteoglycan 4 downregulation in a sheep meniscectomy model of early osteoarthritis. *Arthritis Res Ther* 8:R41

28.        Elsaid K. A., Jay G. D. and Chichester C. O. (2007) Reduced expression and proteolytic susceptibility of lubricin/superficial zone protein may explain early elevation in the coefficient of friction in the joints of rats with antigen-induced arthritis. *Arthritis Rheum* 56:108-16

29.        Bao J. P., Chen W. P. and Wu L. D. (2011) Lubricin: a novel potential biotherapeutic approaches for the treatment of osteoarthritis. *Mol Biol Rep* 38:2879-85

30.        Rhee D. K., Marcelino J., Baker M., Gong Y., Smits P., Lefebvre V., Jay G. D., Stewart M., Wang H., Warman M. L. and Carpten J. D. (2005) The secreted glycoprotein lubricin protects cartilage surfaces and inhibits synovial cell overgrowth. *J Clin Invest* 115:622-31

31.        Onnerfjord P., Khabut A., Reinholt F. P., Svensson O. and Heinegard D. (2012) Quantitative proteomic analysis of eight cartilaginous tissues reveals characteristic differences as well as similarities between subgroups. *J Biol Chem* 287:18913-24

32.        Schumacher B. L., Hughes C. E., Kuettner K. E., Caterson B. and Aydelotte M. B. (1999) Immunodetection and partial cDNA sequence of the proteoglycan, superficial zone protein, synthesized by cells lining synovial joints. *J Orthop Res* 17:110-20

33.        Schumacher B. L., Schmidt T. A., Voegtline M. S., Chen A. C. and Sah R. L. (2005) Proteoglycan 4 (PRG4) synthesis and immunolocalization in bovine meniscus. *J Orthop Res* 23:562-8

34.        Estrella R. P., Whitelock J. M., Packer N. H. and Karlsson N. G. (2010) The glycosylation of human synovial lubricin: implications for its role in inflammation. *Biochem J* 429:359-67

35.        Jay G. D. (1992) Characterization of a bovine synovial fluid lubricating factor. I. Chemical, surface activity and lubricating properties. *Connect Tissue Res* 28:71-88

36.        Lee S., Muller M., Rezwan K. and Spencer N. D. (2005) Porcine gastric mucin (PGM) at the water/poly(dimethylsiloxane) (PDMS) interface: influence of pH and ionic strength on its conformation, adsorption, and aqueous lubrication properties. *Langmuir* 21:8344-53

37.        Liao H., Wu J., Kuhn E., Chin W., Chang B., Jones M. D., O'Neil S., Clauser K. R., Karl J., Hasler F., Roubenoff R., Zolg W. and Guild B. C. (2004) Use of mass spectrometry to identify protein biomarkers of disease severity in the synovial fluid and serum of patients with rheumatoid arthritis. *Arthritis Rheum* 50:3792-803

38.        Hasegawa M., Hirata H., Sudo A., Kato K., Kawase D., Kinoshita N., Yoshida T. and Uchida A. (2004) Tenascin-C concentration in synovial fluid correlates with radiographic progression of knee osteoarthritis. *J Rheumatol* 31:2021-6

39.        Kuhn E., Wu J., Karl J., Liao H., Zolg W. and Guild B. (2004) Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and 13C-labeled peptide standards. *Proteomics* 4:1175-86

40.        Jin C., Ekwall A. K., Bylund J., Bjorkman L., Estrella R. P., Whitelock J. M., Eisler T., Bokarewa M. and Karlsson N. G. (2012) Human synovial lubricin expresses sialyl Lewis x determinant and has L-selectin ligand activity. *J Biol Chem* 287:35922-33

41.        Wandall H. H., Irazoqui F., Tarp M. A., Bennett E. P., Mandel U., Takeuchi H., Kato K., Irimura T., Suryanarayanan G., Hollingsworth M. A. and Clausen H. (2007) The lectin domains of polypeptide GalNAc-transferases exhibit carbohydrate-binding specificity for GalNAc: lectin binding to GalNAc-glycopeptide substrates is required for high density GalNAc-O-glycosylation. *Glycobiology* 17:374-87

42.        Bennett E. P., Mandel U., Clausen H., Gerken T. A., Fritz T. A. and Tabak L. A. (2012) Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* 22:736-56

43.        Gotting C., Kuhn J. and Kleesiek K. (2007) Human xylosyltransferases in health and disease. *Cell Mol Life Sci* 64:1498-517

44.        Matsuura A., Ito M., Sakaidani Y., Kondo T., Murakami K., Furukawa K., Nadano D., Matsuda T. and Okajima T. (2008) O-linked N-acetylglucosamine is present on the extracellular domain of notch receptors. *J Biol Chem* 283:35486-95

45.        Hu P., Shimoji S. and Hart G. W. (2010) Site-specific interplay between O-GlcNAcylation and phosphorylation in cellular regulation. *FEBS lett* 584:2526-38

46.        Lommel M., Willer T., Cruces J. and Strahl S. (2010) POMT1 is essential for protein O-mannosylation in mammals. *Methods Enzymol* 479:323-42

47.         Condac E., Silasi-Mansat R., Kosanke S., Schoeb T., Towner R., Lupu F., Cummings R. D. and Hinsdale M. E. (2007) Polycystic disease caused by deficiency in xylosyltransferase 2, an initiating enzyme of glycosaminoglycan biosynthesis. *Proc Natl Acad Sci USA* 104:9416-21

48.         Marth J. D. (1996) Complexity in O-linked oligosaccharide biosynthesis    engendered    by    multiple    polypeptide    N-acetylgalactosaminyltransferases. *Glycobiology* 6:701-5

49.         Manzi A. E., Norgard-Sumnicht K., Argade S., Marth J. D., van Halbeek H. and Varki A. (2000) Exploring the glycan repertoire of genetically modified mice by isolation and profiling of the major glycan classes and nano-NMR analysis of glycan mixtures. *Glycobiology* 10:669-89

50.         Tenno M., Ohtsubo K., Hagen F. K., Ditto D., Zarbock A., Schaerli P., von Andrian U. H., Ley K., Le D., Tabak L. A. and Marth J. D. (2007) Initiation of protein O glycosylation by the polypeptide GalNAcT-1 in vascular biology and humoral immunity. *Mol Cell Biol* 27:8783-96

51.         Tian E., Hoffman M. P. and Ten Hagen K. G. (2012) O-glycosylation modulates integrin and FGF signalling by influencing the secretion of basement membrane components. *Nat Commun* 3:869

52.         Zhang Y., Iwasaki H., Wang H., Kudo T., Kalka T. B., Hennet T., Kubota T., Cheng L., Inaba N., Gotoh M., Togayachi A., Guo J., Hisatomi H., Nakajima K., Nishihara S., Nakamura M., Marth J. D. and Narimatsu H. (2003) Cloning and characterization of a new human UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase, designated pp-GalNAc-T13, that is specifically expressed in neurons and synthesizes GalNAc alpha-serine/threonine antigen. *J Biol Chem* 278:573-84

53.         Brockhausen I. (1997) Biosynthesis and functions of O-glycans and regulation of mucin antigen expression in cancer. *Biochem Soc Trans* 25:871-4

54.         Ju T., Brewer K., D'Souza A., Cummings R. D. and Canfield W. M. (2002) Cloning and expression of human core 1 beta1,3-galactosyltransferase. *J biol Chem* 277:178-86

55.         Ju T. and Cummings R. D. (2002) A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase. *Proc Natl Acad Sci USA* 99:16613-8

56.         Wang Y., Ju T., Ding X., Xia B., Wang W., Xia L., He M. and Cummings R. D. (2010) Cosmc is an essential chaperone for correct protein O-glycosylation. *Proc Natl Acad Sci USA* 107:9228-33

57.         Aryal R. P., Ju T. and Cummings R. D. (2010) The endoplasmic reticulum chaperone Cosmc directly promotes in vitro folding of T-synthase. *J Biol Chem* 285:2456-62

58.	Bierhuizen M. F. and Fukuda M. (1992) Expression cloning of a cDNA encoding UDP-GlcNAc:Gal beta 1-3-GalNAc-R (GlcNAc to GalNAc) beta 1-6GlcNAc transferase by gene transfer into CHO cells expressing polyoma large tumor antigen. *Proc Natl Acad Sci USA* 89:9326-330

59.	Schwientek T., Nomoto M., Levery S. B., Merkx G., van Kessel A. G., Bennett E. P., Hollingsworth M. A. and Clausen H. (1999) Control of O-glycan branch formation. Molecular cloning of human cDNA encoding a novel beta1,6-N-acetylglucosaminyltransferase forming core 2 and core 4. *J Biol Chem* 274:4504-12

60.	Wang H., Tang R., Zhang W., Amirikian K., Geng Z., Geng J., Hebbel R. P., Xia L., Marth J. D., Fukuda M., Katoh S. and Huo Y. (2009) Core2 1-6-N-glucosaminyltransferase-I is crucial for the formation of atherosclerotic lesions in apolipoprotein E-deficient mice. *Arterioscler, Thromb, Vasc Biol* 29:180-7

61.	Gauguet J. M., Rosen S. D., Marth J. D. and von Andrian U. H. (2004) Core 2 branching beta1,6-N-acetylglucosaminyltransferase and high endothelial cell N-acetylglucosamine-6-sulfotransferase exert differential control over B- and T-lymphocyte homing to peripheral lymph nodes. *Blood* 104:4104-12

62.	Stone E. L., Ismail M. N., Lee S. H., Luu Y., Ramirez K., Haslam S. M., Ho S. B., Dell A., Fukuda M. and Marth J. D. (2009) Glycosyltransferase function in core 2-type protein O glycosylation. *Mol Cell Biol* 29:3770-82

63.	Kuhns W., Rutz V., Paulsen H., Matta K. L., Baker M. A., Barner M., Granovsky M. and Brockhausen I. (1993) Processing O-glycan core 1, Gal beta 1-3GalNAc alpha-R. Specificities of core 2, UDP-GlcNAc: Gal beta 1-3 GalNAc-R(GlcNAc to GalNAc) beta 6-N-acetylglucosaminyltransferase and CMP-sialic acid: Gal beta 1-3GalNAc-R alpha 3-sialyltransferase. *Glycoconj J* 10:381-94

64.	Marcos N. T., Pinho S., Grandela C., Cruz A., Samyn-Petit B., Harduin-Lepers A., Almeida R., Silva F., Morais V., Costa J., Kihlberg J., Clausen H. and Reis C. A. (2004) Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. *Cancer Res* 64:7050-7

65.	Baldus S. E. and Hanisch F. G. (2000) Biochemistry and pathological importance of mucin-associated antigens in gastrointestinal neoplasia. *Adv Cancer Res* 79:201-48

66.	Leivonen M., Nordling S., Lundin J., von Boguslawski K. and Haglund C. (2001) STn and prognosis in breast cancer. *Oncology* 61:299-305

67.          Kim G. E., Bae H. I., Park H. U., Kuan S. F., Crawley S. C., Ho J. J. and Kim Y. S. (2002) Aberrant expression of MUC5AC and MUC6 gastric mucins and sialyl Tn antigen in intraepithelial neoplasms of the pancreas. *Gastroenterology* 123:1052-60

68.          Ragupathi G., Koide F., Sathyan N., Kagan E., Spassova M., Bornmann W., Gregor P., Reis C. A., Clausen H., Danishefsky S. J. and Livingston P. O. (2003) A preclinical study comparing approaches for augmenting the immunogenicity of a heptavalent KLH-conjugate vaccine against epithelial cancers. *Cancer Immunol Immunother* 52:608-16

69.          Vlad A. M., Muller S., Cudic M., Paulsen H., Otvos L., Jr., Hanisch F. G. and Finn O. J. (2002) Complex carbohydrates are not removed during processing of glycoproteins by dendritic cells: processing of tumor antigen MUC1 glycopeptides for presentation to major histocompatibility complex class II-restricted T cells. *J Exp Med* 196:1435-46

70.          Honke K. and Taniguchi N. (2002) Sulfotransferases and sulfated oligosaccharides. *Med Res Rev* 22:637-54

71.          Chandrasekaran E. V., Jain R. K., Vig R. and Matta K. L. (1997) The enzymatic sulfation of glycoprotein carbohydrate units: blood group T-hapten specific and two other distinct Gal:3-O-sulfotransferases as evident from specificities and kinetics and the influence of sulfate and fucose residues occurring in the carbohydrate chain on C-3 sulfation of terminal Gal. *Glycobiology* 7:753-68

72.          Seko A., Ohkura T., Ideo H. and Yamashita K. (2012) Novel O-linked glycans containing 6'-sulfo-Gal/GalNAc of MUC1 secreted from human breast cancer YMB-S cells: possible carbohydrate epitopes of KL-6(MUC1) monoclonal antibody. *Glycobiology* 22:181-95

73.          van Zante A. and Rosen S. D. (2003) Sulphated endothelial ligands for L-selectin in lymphocyte homing and inflammation. *Biochem Soc Trans* 31:313-7

74.          Hiraoka N., Petryniak B., Nakayama J., Tsuboi S., Suzuki M., Yeh J. C., Izawa D., Tanaka T., Miyasaka M., Lowe J. B. and Fukuda M. (1999) A novel, high endothelial venule-specific sulfotransferase expresses 6-sulfo sialyl Lewis(x), an L-selectin ligand displayed by CD34. *Immunity* 11:79-89

75.          de Graffenried C. L. and Bertozzi C. R. (2004) The stem region of the sulfotransferase GlcNAc6ST-1 is a determinant of substrate specificity. *J Biol Chem* 279:40035-43

76.          Zauner G., Kozak R. P., Gardner R. A., Fernandes D. L., Deelder A. M. and Wuhrer M. (2012) Protein O-glycosylation analysis. *Biol Chem* 393:687-708

77.	Marshall A. G. and Hendrickson C. L. (2008) High-resolution mass spectrometers. *Annu Rev Anal Chem* 1:579-99

78.	Hu Q., Noll R. J., Li H., Makarov A., Hardman M. and Graham Cooks R. (2005) The Orbitrap: a new mass spectrometer. *J Mass Spectrom* 40:430-43

79.	Anderson L. and Hunter C. L. (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol Cell Proteomics* 5:573-88

80.	Fenn J. B., Mann M., Meng C. K., Wong S. F. and Whitehouse C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246:64-71

81.	Karas M. and Hillenkamp F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60:2299-301

82.	Wuhrer M., de Boer A. R. and Deelder A. M. (2009) Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry. *Mass Spectrom Rev* 28:192-206

83.	Kebarle P. (2000) A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J Mass Spectrom* 35:804-17

84.	Quenzer T. L., Emmett M. R., Hendrickson C. L., Kelly P. H. and Marshall A. G. (2001) High sensitivity Fourier transform ion cyclotron resonance mass spectrometry for biological analysis with nano-LC and microelectrospray ionization. *Anal Chem* 73:1721-5

85.	Scigelova M. and Makarov A. (2006) Orbitrap mass analyzer--overview and applications in proteomics. *Proteomics* 2:16-21

86.	Chung L. M., Colangelo C. M. and Zhao H. (2014) Data Pre-Processing for Label-Free Multiple Reaction Monitoring (MRM) Experiments. *Biology* 3:383-402

87.	Dakin S. G., Smith R. K., Heinegard D., Onnerfjord P., Khabut A. and Dudhia J. (2014) Proteomic analysis of tendon extracellular matrix reveals disease stage-specific fragmentation and differential cleavage of COMP (cartilage oligomeric matrix protein). *J Biol Chem* 289:4919-27

88.	Karlsson N. G., Wilson N. L., Wirth H. J., Dawes P., Joshi H. and Packer N. H. (2004) Negative ion graphitised carbon nano-liquid chromatography/mass spectrometry increases sensitivity for glycoprotein oligosaccharide analysis. *Rapid Commun Mass Spectrom* 18:2282-92

89.	Wuhrer M. (2013) Glycomics using mass spectrometry. *Glycoconj J* 30:11-22

90.		Kenny D. T., Skoog E. C., Linden S. K., Struwe W. B., Rudd P. M. and Karlsson N. G. (2012) Presence of terminal N-acetylgalactosaminebeta1-4N-acetylglucosamine residues on O-linked oligosaccharides from gastric MUC5AC: involvement in Helicobacter pylori colonization? *Glycobiology* 22:1077-85

91.		Karlsson N. G., Schulz B. L. and Packer N. H. (2004) Structural determination of neutral O-linked oligosaccharide alditols by negative ion LC-electrospray-MSn. *J Am Soc Mass Spectrom* 15:659-72

92.		Halim A., Nilsson J., Ruetschi U., Hesse C. and Larson G. (2012) Human urinary glycoproteomics; attachment site specific analysis of N- and O-linked glycosylations by CID and ECD. *Mol Cell Proteomics* 11:M111 013649

93.		Domon B. and Costello C. E. (1988) Structure elucidation of glycosphingolipids and gangliosides using high-performance tandem mass spectrometry. *Biochemistry* 27:1534-43

94.		Zauner G., Selman M. H., Bondt A., Rombouts Y., Blank D., Deelder A. M. and Wuhrer M. (2013) Glycoproteomic analysis of antibodies. *Mol Cell Preoteomics* 12:856-65

95.		Lu H., Zong C., Wang Y., Young G. W., Deng N., Souda P., Li X., Whitelegge J., Drews O., Yang P. Y. and Ping P. (2008) Revealing the dynamics of the 20 S proteasome phosphoproteome: a combined CID and electron transfer dissociation approach. *Mol Cell Preoteomics* 7:2073-89

96.		Hanisch F. G. (2012) O-glycoproteomics: site-specific O-glycoprotein analysis by CID/ETD electrospray ionization tandem mass spectrometry and top-down glycoprotein sequencing by in-source decay MALDI mass spectrometry. *Methods Mol Biol* 842:179-89

97.		Zhang C. C., Rogalski J. C., Evans D. M., Klockenbusch C., Beavis R. C. and Kast J. (2011) In silico protein interaction analysis using the global proteome machine database. *J Proteome Res* 10:656-68

98.		McClintock C. S., Parks J. M., Bern M., Ghattyvenkatakrishna P. K. and Hettich R. L. (2013) Comparative informatics analysis to evaluate site-specific protein oxidation in multidimensional LC-MS/MS data. *J Proteome Res* 12:3307-16

99.		Arnett F. C., Edworthy S. M., Bloch D. A., McShane D. J., Fries J. F., Cooper N. S., Healey L. A., Kaplan S. R., Liang M. H., Luthra H. S. and et al. (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 31:315-24

100.     Ruhaak L. R., Deelder A. M. and Wuhrer M. (2009) Oligosaccharide analysis by graphitized carbon liquid chromatography-mass spectrometry. *Anal Bioanal Chem* 394:163-74

101.     Chiba A., Matsumura K., Yamada H., Inazu T., Shimizu T., Kusunoki S., Kanazawa I., Kobata A. and Endo T. (1997) Structures of sialylated O-linked oligosaccharides of bovine peripheral nerve alpha-dystroglycan. The role of a novel O-mannosyl-type oligosaccharide in the binding of alpha-dystroglycan with laminin. *J Biol Chem* 272:2156-62

102.     Kang P., Mechref Y., Klouckova I. and Novotny M. V. (2005) Solid-phase permethylation of glycans for mass spectrometric analysis. *Rapid Commun Mass Spectrom* 19:3421-8

103.     Hakomori S. (1964) A rapid permethylation of Glycolipid, and Polysaccharide catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide.. *J Biochem* 55:205-8

104.     Hayes C. A., Karlsson N. G., Struwe W. B., Lisacek F., Rudd P. M., Packer N. H. and Campbell M. P. (2011) UniCarb-DB: a database resource for glycomic discovery. *Bioinformatics* 27:1343-4

105.     Zhang H., Wang Z., Stupak J., Ghribi O., Geiger J. D., Liu Q. Y. and Li J. (2012) Targeted glycomics by selected reaction monitoring for highly sensitive glycan compositional analysis. *Proteomics* 12:2510-22

106.     Zauner G., Koeleman C. A., Deelder A. M. and Wuhrer M. (2010) Protein glycosylation analysis by HILIC-LC-MS of Proteinase K-generated N- and O-glycopeptides. *J Sep Sci* 33:903-10

107.     Selman M. H., Hemayatkar M., Deelder A. M. and Wuhrer M. (2011) Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides. *Anal Chem* 83:2492-9

108.     Catalina M. I., Koeleman C. A., Deelder A. M. and Wuhrer M. (2007) Electron transfer dissociation of N-glycopeptides: loss of the entire N-glycosylated asparagine side chain. *Rapid Commun Mass Spectrom* 21:1053-61

109.     Wuhrer M., Deelder A. M. and Hokke C. H. (2005) Protein glycosylation analysis by liquid chromatography-mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 825:124-33

110.     Reiding K. R., Blank D., Kuijper D. M., Deelder A. M. and Wuhrer M. (2014) High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification. *Anal Chem* 86:5784-93

111.     Marino K., Bones J., Kattla J. J. and Rudd P. M. (2010) A systematic approach to protein glycosylation analysis: a path through the maze. *Nat Chem Biol* 6:713-23

112.        Guzman-Aranguez A., Mantelli F. and Argueso P. (2009) Mucin-type O-glycans in tears of normal subjects and patients with non-Sjogren's dry eye. *Invest Ophthalmol Vis Sci* 50:4581-7

113.        Kono M., Ohyama Y., Lee Y. C., Hamamoto T., Kojima N. and Tsuji S. (1997) Mouse beta-galactoside alpha 2,3-sialyltransferases: comparison of in vitro substrate specificities and tissue specific expression. *Glycobiology* 7:469-79

114.        Ogawa J. I., Inoue H. and Koide S. (1997) alpha-2,3-Sialyltransferase type 3N and alpha-1,3-fucosyltransferase type VII are related to sialyl Lewis(x) synthesis and patient survival from lung carcinoma. *Cancer* 79:1678-85

115.        Priatel J. J., Chui D., Hiraoka N., Simmons C. J., Richardson K. B., Page D. M., Fukuda M., Varki N. M. and Marth J. D. (2000) The ST3Gal-I sialyltransferase controls CD8+ T lymphocyte homeostasis by modulating O-glycan biosynthesis. *Immunity* 12:273-83

116.        Grabie N., Delfs M. W., Lim Y. C., Westrich J. R., Luscinskas F. W. and Lichtman A. H. (2002) Beta-galactoside alpha2,3-sialyltransferase-I gene expression during Th2 but not Th1 differentiation: implications for core2-glycan formation on cell surface proteins. *Eur J Immunol* 32:2766-72

117.        Marcos N. T., Bennett E. P., Gomes J., Magalhaes A., Gomes C., David L., Dar I., Jeanneau C., DeFrees S., Krustrup D., Vogel L. K., Kure E. H., Burchell J., Taylor-Papadimitriou J., Clausen H., Mandel U. and Reis C. A. (2011) ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues. *Front Biosci* 3:1443-55

118.        Ahrman E., Lorenzo P., Holmgren K., Grodzinsky A. J., Dahlberg L. E., Saxne T., Heinegard D. and Onnerfjord P. (2014) Novel Cartilage Oligomeric Matrix Protein (COMP) Neoepitopes Identified in Synovial Fluids from Patients with Joint Diseases Using Affinity Chromatography and Mass Spectrometry. *J Biol Chem* 289:20908-20916

119.        Jay G. D., Harris D. A. and Cha C. J. (2001) Boundary lubrication by lubricin is mediated by O-linked beta(1-3)Gal-GalNAc oligosaccharides. *Glycoconj J* 18:807-15

120.        Kamphorst J. J., van der Heijden R., DeGroot J., Lafeber F. P., Reijmers T. H., van El B., Tjaden U. R., van der Greef J. and Hankemeier T. (2007) Profiling of endogenous peptides in human synovial fluid by NanoLC-MS: method validation and peptide identification. *J Proteome Res* 6:4388-96

121.        Gobezie R., Kho A., Krastins B., Sarracino D. A., Thornhill T. S., Chase M., Millett P. J. and Lee D. M. (2007) High abundance synovial fluid proteome: distinct profiles in health and osteoarthritis. *Arthritis Res Ther* 9:R36

122.        Hollingsworth M. A. and Swanson B. J. (2004) Mucins in cancer: protection and control of the cell surface. *Nat Rev Cancer* 4:45-60

123.        Ghosh S. K., Pantazopoulos P., Medarova Z. and Moore A. (2013) Expression of underglycosylated MUC1 antigen in cancerous and adjacent normal breast tissues. *Clin Breast Cancer* 13:109-18

124.        Conze T., Carvalho A. S., Landegren U., Almeida R., Reis C. A., David L. and Soderberg O. (2010) MUC2 mucin is a major carrier of the cancer-associated sialyl-Tn antigen in intestinal metaplasia and gastric carcinomas. *Glycobiology* 20:199-206

125.        Firestein G. S. (2003) Evolving concepts of rheumatoid arthritis. *Nature* 423:356-61

126.        Morris D. and Inman R. D. (2012) Reactive arthritis: developments and challenges in diagnosis and treatment. *Curr Rheumatol Rep* 14:390-4

127.        Flowers S. A., Ali L., Lane C. S., Olin M. and Karlsson N. G. (2013) Selected reaction monitoring to differentiate and relatively quantitate isomers of sulfated and unsulfated core 1 O-glycans from salivary MUC7 protein in rheumatoid arthritis. *Mol Cell Proteomics* 12:921-31

128.        Seko A., Nagata K., Yonezawa S. and Yamashita K. (2002) Down-regulation of Gal 3-O-sulfotransferase-2 (Gal3ST-2) expression in human colonic non-mucinous adenocarcinoma. *Jpn J Cancer Res* 93:507-15

129.        Estrella R. P., Whitelock J. M., Packer N. H. and Karlsson N. G. (2007) Graphitized carbon LC-MS characterization of the chondroitin sulfate oligosaccharides of aggrecan. *Anal Chem* 79:3597-606

130.        Roughley P. J. and White R. J. (1992) The dermatan sulfate proteoglycans of the adult human meniscus. *J Orthop Res* 10:631-7

131.        Prakobphol A., Tangemann K., Rosen S. D., Hoover C. I., Leffler H. and Fisher S. J. (1999) Separate oligosaccharide determinants mediate interactions of the low-molecular-weight salivary mucin with neutrophils and bacteria. *Biochemistry* 38:6817-25

132.        Zappone B., Ruths M., Greene G. W., Jay G. D. and Israelachvili J. N. (2007) Adsorption, lubrication, and wear of lubricin on model surfaces: polymer brush-like behavior of a glycoprotein. *Biophys J* 92:1693-708

133.        Swann D. A., Hendren R. B., Radin E. L., Sotman S. L. and Duda E. A. (1981) The lubricating activity of synovial fluid glycoproteins. *Arthritis Rheum* 24:22-30

134.    Gaisinskaya A., Ma L., Silbert G., Sorkin R., Tairy O., Goldberg R., Kampf N. and Klein J. (2012) Hydration lubrication: exploring a new paradigm. *Farad Discuss* 156:217-33

135.    Flannery C. R., Hughes C. E., Schumacher B. L., Tudor D., Aydelotte M. B., Kuettner K. E. and Caterson B. (1999) Articular cartilage superficial zone protein (SZP) is homologous to megakaryocyte stimulating factor precursor and Is a multifunctional proteoglycan with potential growth-promoting, cytoprotective, and lubricating properties in cartilage metabolism. *Biochem Biophys Res Commun* 254:535-41

136.    Schaefer D. B., Wendt D., Moretti M., Jakob M., Jay G. D., Heberer M. and Martin I. (2004) Lubricin reduces cartilage--cartilage integration. *Biorheology* 41:503-8