INAUGURAL - DISSERTATION

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich-Mathematischen Gesamtfakultät

der

Ruprecht - Karls - Universität

Heidelberg

vorgelegt von Diplom-Biologin

Nina Röckel

aus Zweibrücken

Tag der mündlichen Prüfung: .....

ANALYSIS OF CODING MICROSATELLITE MUTATIONS IN MSI COLORECTAL CARCINOMAS AND CHARACTERIZATION OF THEIR EFFECTS ON THE CELLULAR GLYCOSYLATION MACHINERY

Gutachter: Prof. Dr. Bernhard Dobberstein

Prof. Dr. Jürgen Kopitz

Meinen Eltern

I. LIST OF FIGURES

II. LIST OF TABLES

1. Abstract

Microsatellite instability (MSI), i.e. length variations of repetitive DNA sequences (microsatellites) occurs in colorectal cancer (CRC) due to defects in the DNA mismatch repair (MMR) system and is a hallmark of tumors associated with hereditary non-polyposis colorectal cancer (HNPCC, 1 – 5% of total CRC) as well as sporadic CRC (15% of total CRC). Instability in coding region microsatellites (cMNR) of expressed genes lead to frameshift mutations and a subsequent loss of protein function or the translation of a truncated protein. Despite the obvious significance of altered glycoprotein synthesis, transport and glycosylation patterns in colorectal cancer, aberrant glycosylation and glycosylation pathways have not been investigated in MSI colorectal tumors.

Using a bioinformatics-based approach 28 cMNR harboring candidate genes were identified that encode proteins of the cellular glycosylation machinery. Coding MNR mutation analysis in MSI CRC cell lines revealed a high mutation frequency in two genes: the glycoprotein transporter gene *LMAN1/ERGIC53* (52%; 12/23) and the xylosyltransferase gene *XYLT2* (35%; 8/23) that catalyzes the first step in proteoglycan synthesis. Apart from their occurrence in cultured cell lines these genetic alterations were also found at similar frequency in MSI colorectal adenomas (*LMAN1*: 40%, 8/20; *XYLT2*: 21%, 4/19) as well as carcinomas (*LMAN1*: 46%, 78/170; *XYLT2*: 26%, 27/105). Biallelic mutations, abrogating normal protein function, were detected in MSI CRC cell lines (LMAN1, XYLT2) and in primary colorectal tumors (LMAN1).

Biallelic mutant *LMAN1* was transcribed but not translated into a stable protein in MSI CRC cell lines or *LMAN1*-mutated areas in tumors. MSI CRC cell lines with biallelic *LMAN*1 mutations behaved different to *LMAN1*-proficient CRC cell lines. *LMAN1*-deficient cell lines exhibited severely reduced transport and secretion of the LMAN1 cargo protein alpha-1-antitrypsin (A1AT), an inhibitor of angiogenesis and tumor growth but impaired secretion could be restored upon *LMAN1* re-expression. A strong correlation between lower local levels of A1AT and enhanced tumor growth has been described. Furthermore, re-expression of *LMAN1* in *LMAN1*-deficient LoVo cells led to a cell surface glycoprotein pattern different from the pattern observed on LoVo cells lacking LMAN1. Increased binding of PHA-L lectins to the cell surface suggested a glycosylation pattern known to be involved in the progression of cancer from a tumorigenic to a metastatic phenotype.

A radioactive incorporation assay with ^{3}H-xylose also revealed changes in proteoglycan synthesis of a *XYLT2*-deficient MSI CRC cell line transiently transfected with a *XYLT2*-construct.

Overall, two members of the cellular glycosylation machinery, a glycoprotein transporter and a glycosyltransferase, appear to be involved in MSI tumorigenesis. Genetic alterations in *LMAN1* and *XYLT2* mark early events already detectable in preneoplastic lesions. Furthermore, loss of LMAN1 and XYLT2 protein function changed the secretion of proteins and the cell surface glycosylation pattern. Both genes might influence MSI tumor progression by changing cell-cell communication and interactions.

1.1 Zusammenfassung

Mikrosatelliten-Instabilität (MSI), d.h. Längenvariationen in repetitiven DNA-Sequenzen (Mikrosatelliten), wird in kolorektalen Karzinomen (CRC) durch die funktionelle Inaktivierung eines zellulären DNA-Reparatursystems verursacht, dem sogenannten Mismatch-Reparatur (MMR) System. MSI ist ein Kennzeichen von MSI-Darmtumoren, die entweder sporadisch (15% aller Kolonkarzinome) oder im Rahmen des hereditären nicht-polypösen Kolonkarzinoms (hereditary non-polyposis colorectal cancer HNPCC; 15% aller Kolonkarzinome) entstehen. Instabilität in Mikrosatelliten eines kodierenden Genabschnitts (cMNR) exprimierter Gene, führen zu Basenfehlpaarungen, die anschließend einen Verlust der Proteinfunktion oder die Entstehung verkürzter Proteine (Neopeptide) zur Folge haben. Eine veränderte Glykoprotein-Synthese, veränderter Transport und ein verändertes Glykolisierungsmuster haben eine offensichtliche Bedeutung in kolorektalen Karzinomen, allerdings wurden abnormale Glykosylierung und Glykosylierungswege bisher nicht in MSI kolorektalen Tumoren untersucht.

Unter Verwendung von Datenbank-Analysen wurden 28 Kandidatengene mit einem cMNR identifiziert. cMNR-Leserastermutations-Analysen bei einer Auswahl von MSI CRC Zellinien zeigten bei zwei Genen eine hohe Mutationsfrequenz: bei dem Glykoprotein-Transporter Gen *LMAN1/ERGIC53* (52%; 12/23) und dem Xylosyltransferase Gen *XYLT2* (35%; 8/23), das den ersten Schritt der Proteoglykan-Synthese katalysiert. Neben genetischen Veränderungen in MSI CRC Zellinien wurden ähnlich hohe Mutationsfrequenzen in MSI kolorektalen Adenomen (*LMAN1*: 40%, 8/20; *XYLT2*: 21%, 4/19) und in MSI kolorektalen Karzinomen (*LMAN1*: 46%, 78/170; *XYLT2*: 26%, 27/105) beobachtet. Biallelische Mutationen, die eine veränderte Proteinfunktion zur Folge haben,

wurden in MSI CRC Zellinien (LMAN1, XYLT2) und in primären kolorektalen Tumoren (LMAN1) gefunden.

Das biallelisch mutierte *LMAN1*-Gen wurde in MSI Darmkrebs-Zellinien transkribiert, jedoch nicht in ein stabiles Protein translatiert. Auch im Gewebe konnte eine biallelische *LMAN1*-Mutation mit einem Proteinverlust assoziiert werden. MSI Darmzellinien mit biallelischen *LMAN1*-Mutationen zeigten ein verändertes Verhalten im Vergleich zu *LMAN1*-exprimierenden MSI Darmzellinien. *LMAN1*-defiziente Zellinien wiesen einen reduzierten Transport und eine reduzierte Sekretion von einem LMAN1-Transportprotein Alpha-1-Antitrypsin (A1AT) auf. A1AT ist ein Inhibitor der Angiogenese und des Tumorwachstums. Die eingeschränkte A1AT-Sekretion konnte durch Re-Expression von *LMAN1* in *LMAN1*-defizienten Zellinien wieder hergestellt werden. Es wurde gezeigt, dass ein Verlust der A1AT-Sekretion einen Wachstumsvorteil für den Tumor darstellt. Außerdem führte ein LMAN1-Verlust zu einem veränderten Glykosylierungsmuster an der Zelloberfläche von *LMAN1*-defizienten LoVo Zellen im Vergleich zu LoVo Zellen, die *LMAN1* stabil exprimierten. Die vermehrte Bindung von PHA-L Lektinen an die Zelloberfläche war verbunden mit mehr β1-6 Verzweigungen in *LMAN1*-defizienten LoVo Zellen. β1-6 Verzweigungen sind Glykosylierungsmuster, die mit der Tumorprogression und der Metastasen-Entstehung in Verbindung gebracht werden.

Der radioaktive Einbau von ^{3}H-Xylose zeigte Veränderungen in der Proteoglykan-synthese *XYLT2*-defizienter MSI-Darmzellinien, die mit einem *XYLT2*-Konstrukt transient transfiziert wurden.

Es wurden insgesamt zwei Mitglieder der zellulären Glykosylierungsmaschinerie, ein Glykoproteintransporter und eine Glykosyltransferase, identifiziert, die einen Einfluss auf die MSI Tumorgenese haben könnten. Genetische Veränderungen in *LMAN1* und *XYLT2* sind früh auftretende Ereignisse, die bereits in prä-neoplastischen Läsionen nachweisbar waren. Zudem änderte der Verlust von LMAN1 und XYLT2 die Sekretion von Proteinen und das Glykosylierungsmuster an der Zelloberfläche. Beide Gene könnten die MSI Tumorprogression durch eine veränderte Zell-Zell Kommunikation oder Interaktion beeinflussen.

2. Introduction

2.1 Colorectal Cancer

Colorectal cancer (CRC) is the fourth most common cancer in men and the third most common cancer in women. Worldwide more than one million people develop CRC per year [Parkin et al., 2005] and 11% of total cancer deaths are related to CRC. The 5-year survival rate depends on tumor stage, varying from 80 – 95% for early stage tumors and 0 – 7% for tumors in a late stage [Weitz et al., 2005]. During carcinogenesis normal colonic epithelium transforms into colon adeno-carcinomas, accompanied by a progressive accumulation of genetic alterations [Fearon et al., 1990; Lynch et al., 2002]. CRC is caused by the sequential accumulation of multiple somatic mutations affecting different cancer-related genes. Two major pathways lead to CRC, the chromosomal instability (CIN) pathway and the microsatellite instability (MSI) pathway [Lengauer et al., 1997].

2.2 Molecular Pathogenesis of MSI and CIN

Although the majority of CRCs shows CIN (85%), about 15% of CRCs are attributable to defects in the DNA mismatch repair (MMR) system [Boland et al., 1998; Hampel et al., 2005; Ionov et al., 1993]. MMR-deficient CRCs may develop sporadically or in the context of hereditary non-polyposis colorectal cancer (HNPCC) or Lynch syndrome [Lynch et al., 1999; Thibodeau et al., 1993]. DNA MMR deficiency induces a high number of mutational events characterized by small alterations at the nucleotide level like missense, nonsense and frameshift mutations [Söreide et al., 2006]. These mutations occur mainly at short repetitive DNA sequences, termed microsatellites, because these structures are particularly prone to DNA polymerase slippage during DNA replication. The resulting phenotype is termed high-level microsatellite instability (MSI-H). For molecular diagnostics of MSI-H tumors the National Cancer Institute (NCI) has recommended a panel of five microsatellite markers, known as Bethesda markers, which include two mononucleotide repeats, BAT25 and BAT26 [Boland et al., 1998], and three dinucleotide repeats, D2S123, D5S346 and D17S250 [Umar et al., 2004]. If two or

more markers are mutated, tumors are diagnosed as MSI-H. In contrast, CIN cancer is characterized by large chromosomal alterations like allelic losses, chromosomal amplifications and translocations, as well as loss of heterozygosity (LOH). Most CIN cancers are microsatellite stable (MSS). Tumors with CIN have mutations in *APC, p53* and the *RAS* gene [Samowitz et al., 2001], whereas tumors with MSI show, after the initial loss of the MMR function, mutations in specific target genes, such *as β-Catenin*, *transforming growth factor ß receptor II* (*TGFBR2*) and *activin A receptor type IIA* (*ACVR2A*) [Kim et al., 2003; Markowitz et al., 1995]. Both pathways require an early step mutation of distinct genes followed by characteristic further mutations (Fig. 2.1).

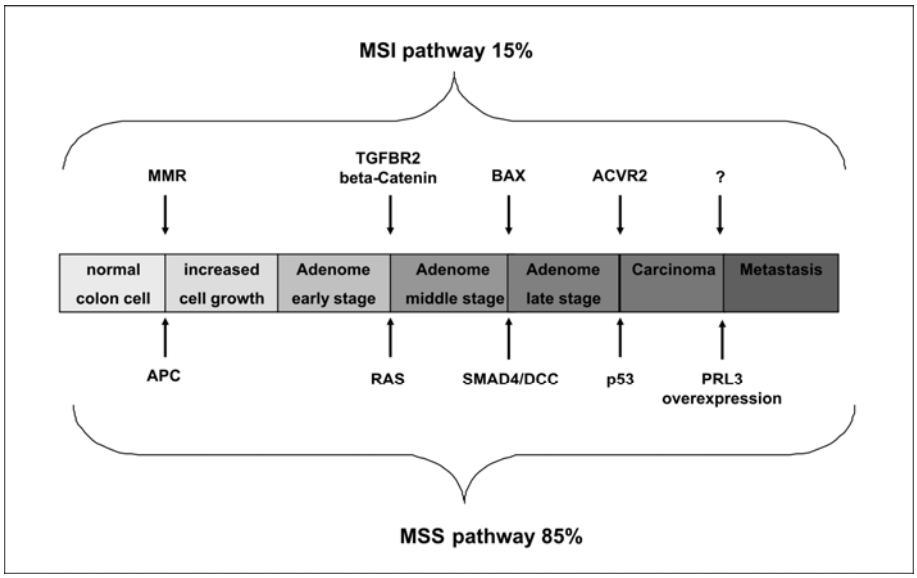


Figure 2.1 Characteristics of the two major pathways in colorectal cancer [Söreide et al., 2006]. About 85% of CRCs account for the CIN/MSS and 15% for the MSI pathway. MSS tumors show mutations in *APC, p53* and the *Ras*-genes, whereas MSI tumors have mutations in the *MMR* genes and subsequently specific target genes are mutated, such as *beta-Catenin, TGFBR2, BAX and ACVR2*.

2.3 Microsatellite Instability (MSI) in Colorectal Cancer

Sporadic as well as hereditary MSI-H cancers show functional inactivation of the MMR system. In HNPCC and Lynch syndrome tumors defects are caused by germline mutations in one of four *MMR* genes (*MLH1, MSH2, MSH6 and PMS2*) whereas in sporadic MSI-H tumors defects are attributable to epigenetic silencing of the MLH1 promoter [Boland et al., 1998; Ionov et al., 1993; Kane et al., 1997; Thibodeau et al., 1993]. In addition to MLH1 promotor methylation in sporadic MSI-H tumors, these tumors are associated with *BRAF* mutations [Rajagopalan et al., 2002]. BRAF is a serine/threonine kinase involved in the MAPK signalling pathway [McGivern et al., 2004]. The occurrence of MLH1 promoter hypermethylation and *BRAF* mutations distinguish HNPCC from sporadic MSI-H colon cancers.

2.3.1 Molecular background of MSI tumors

Microsatellites represent the most variable types of DNA sequences in the genome and are estimated to account for about 3% of the human genome. Microsatellites consist of repetitive units of different length with one to five base pairs, non-randomly distributed throughout the human genome [Ellegren, 2004; Li et al., 2004]. The maintenance of microsatellite stability is controlled by the MMR system, which is able to correct base substitutions and mismatches as well as insertion/deletion mutations. Figure 2.2 shows the role of the MMR system in maintaining the length of microsatellite sequences. Microsatellites are distributed in non-coding regions (intragenic or intergenic) or in coding regions. Intragenic regions like promoters, 3`-untranslated regions and introns can be important regulators of gene expression, while intergenic regions could have functions in chromatin organization and recombination [Shah et al., 2010]. Mutations in microsatellites located in coding regions of expressed genes might lead to a loss of protein function or the translation of truncated proteins due to frameshift mutations. Coding microsatellites consisting one type of nucleotide (in Fig. 2.2 an A8 repeat is shown) are known as mononucleotide repeats (cMNRs). Examples for genes frequently affected by cMNR mutations in MSI tumors are *TGFBR2*, *ACVR2A*, *caspase 5* (*CASP5*) or the *BCL2-associated X protein* (*BAX*) [Duval et al., 2002].

Frameshift-derived truncated proteins harbor neopeptide tails at their C-terminus that represent potentially antigenic epitopes capable to induce cellular and/or humoral immune responses [Linnebacher et al., 2001; Schwitalle et al., 2008]. It is generally assumed that coding microsatellite frameshift mutations in some of these genes provide a growth advantage to MMR-deficient cells and hence drive MSI tumorigenesis [Woerner et al., 2003].

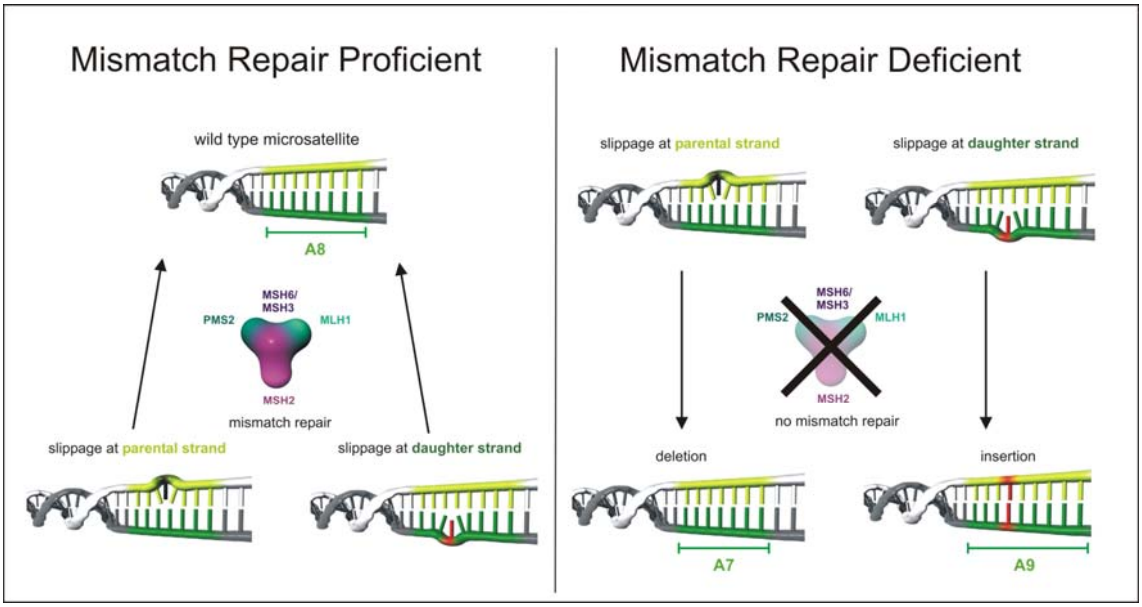


Figure 2.2 Role of the MMR system in maintaining microsatellite length [Kloor et al., 2005]. MSI is associated with frameshift mutations in repetitive sequences, which can not be repaired by the MMR-system because of a defect either in the germline in HNPCC or by epigenetic silencing (MLH1-promotor) in sporadic MSI-H colorectal tumors. In MMR proficient cells slippages of polymerase enzymes during replication can occur either on the parental or daughter strand (left panel). The MMR system recognizes slippages and they are repaired. In MMR deficient cells insertion or deletion mutations are manifested, since they are not recognized by the MMR system (right panel).

Our group has established a database of human cMNR mutations (www.seltarbase.org) and proposed a statistical model that allows the prediction of genes, that when mutated might provide a growth advantage to affected cells, including prediction of selective targets in human MSI-H tumorigenesis [Woerner et al., 2010]. The statistical model of this database is based on a sigmoid regression analysis aiming at the identification of genes involved in MSI carcinogenesis by their mutation frequency in cMNRs. There, the given cMNR length is brought into

correlation with the average mutation rate, revealing genes with increased or decreased mutation frequency. Accordingly, major efforts in MSI tumor research have been directed towards identifying these genes and analyzing their functional role.

2.3.2 Clinicopathological characteristics of MSI tumors

MSI colorectal tumors show a distinct clinical phenotype characterized by a poor differentiation [Greenson et al., 2003], strong lymphocytic infiltration [Dolcetti et al., 1999; Smyrk et al., 2001], better prognosis and a mucinous histology [Jass, 1998]. Besides, MSI tumors are more frequently found in the proximal colon, are nearly diploid and when associated with Lynch-syndrome occur at an earlier age [Umar, 2004]. In contrast, sporadic MSI CRC is most frequently found in older women [Young et al., 2001]. Furthermore MSI tumors are resistant against certain DNA-damaging chemotherapeutic drugs, such as 5-fluorouracil (5-FU) [Carethers et al., 1999]. In addition, colorectal cancers exhibiting MSI show a strong lymphocytic infiltration and a low frequency of distant metastasis [Buckowitz et al., 2005]. This might be attributable to enhanced immunogenicity conferred by MSI tumor specific neopeptides [Linnebacher et al., 2001; Schwitalle et al., 2008]. In addition to neopeptides, MSI-specific changes of glycopeptide expression on the cell surface might also contribute to the high immunogenicity in MSI tumors.

2.4 Eucaryotic glycosylation

More than 5% of all mammalian genes are involved in glycosylation, encoding for glycosyltransferases, glycoprotein transporters or glycoprotein modifiers [Apweiler et al., 1999; Hebert et al., 2005; Lowe et al., 2003]. Nine monosaccharides are commonly found in mammals and used in the enzymatic process of glycosylation [Ohtsubo et al., 2006]: D-glucose (Glc), N-Acetyl-D-glucosamine (GlcNAc), D-galactose (Gal), N-Acetyl-D-galactosamine (GalNAc), D-mannose (Man), D-xylose (Xyl), D-Glucuronic acid (GlcA), L-Fucose (Fuc) and sialic acid, e.g. N-Acetylneuraminic acid (NeuAc). Conserved biosynthetic pathways provide all nine monosaccharides from sugars ubiquitously present in the diet. These monosacchrides can be covalently attached to other molecules and then sequentially elongated and branched creating protein-associated N- and O-glycans, glycosaminoglycans (GAGs), and lipid-associated glycans such as glycosphingolipids (GSLs), as well as free glycans in the form of hyaluronan [Potapenko et al., 2010]. N-linked glycoproteins possess a glycosidic bond to asparagine residues of proteins and O-linked glycoproteins are linked to the core protein via serine or threonine residues. Glycans participate in multiple mechanisms of cellular regulation, spanning from protein quality control and intracellular trafficking to roles in extracellular compartments where cell-cell communication is modulated by adhesion, receptor activation, signal transduction and endocytosis (Fig. 2.3; [Ohtsubo et al., 2006]). Details on glycan function and the different glycan structures important for the available study are presented in sections 2.4.1 – 2.4.5.

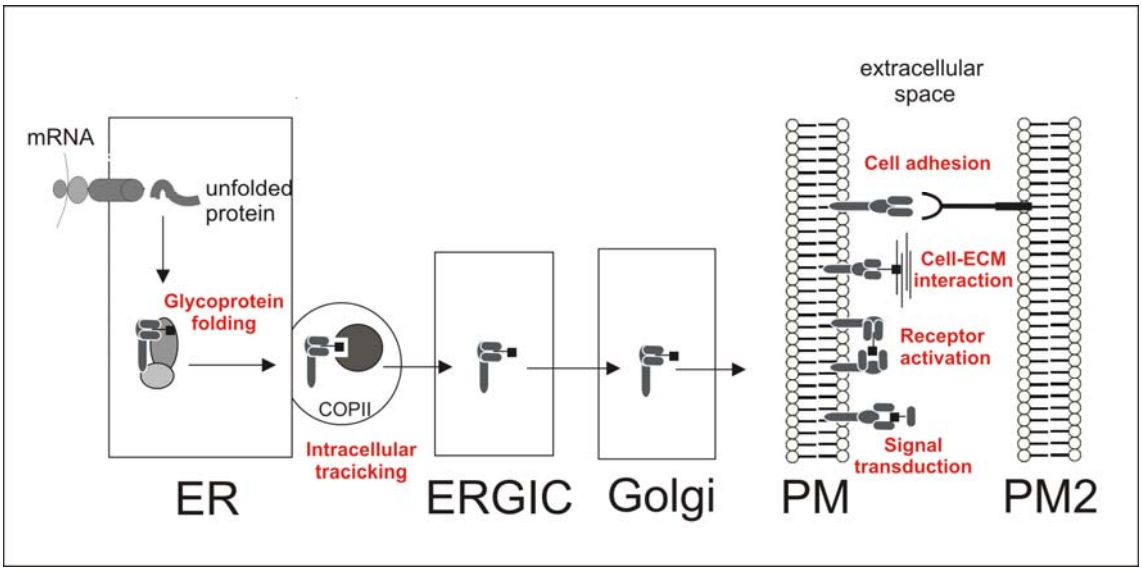


Figure 2.3 Overview of glycan functions. Glycans produced in the secretory pathway possess mutilple functions in cellular processes, including protein folding, quality control and intracellular trafficking. Cell-cell communication is modulated by adhesion, receptor activation and signal transduction. ER: endoplasmic Reticulum; ERGIC: ER-Golgi intermediate compartment; PM: plasma membrane.

2.4.1 N-glycosylation and the secretory pathway

Transfer of the Dolichol-linked precursor to nascent proteins

Oligosaccharide chains of N-gycosylated proteins are connected to the core protein via an asparagine residue in a distinct sequence motif. They are pre-assembled on a lipid-based anchor in the Endoplasmic reticulum (ER). This lipid-based anchor, the lipid-linked oligosaccharide dolicholpyrophosphate (LLO-Dol-PP) comprises a 14-residue oligosaccharide, composed of nine Man, two GlcNAc and three Glc residues. The transfer is catalyzed by an oligosaccharyltransferase (OST) linking the 14-residue oligosaccharide en bloc to an asparagine (Asn) residue in the consensus sequence Asn-X-Ser/Thr (Fig. 2.4). After the covalent attachment of the oligosaccharide from LLO-Dol-PP to asparagine residues, a series of processing reactions occur, including folding and quality control [Gabius H-J, 2009].

Refolding and quality control in the ER

The ER contains several proteins that accelerate the folding of newly synthesized proteins within the ER lumen. ER chaperones, the homologous lectins calnexin

(CX) and calreticulin (CR), bind to mono-glucosylated glycoproteins (Fig. 2.4; [Trombetta et al., 1998]). The removal of glucose residues, catalyzed by α-glucosidases (Glc I, II), is associated with protein folding mechanisms and starts immediately after transfer from the lipid-linked oligosaccharide.

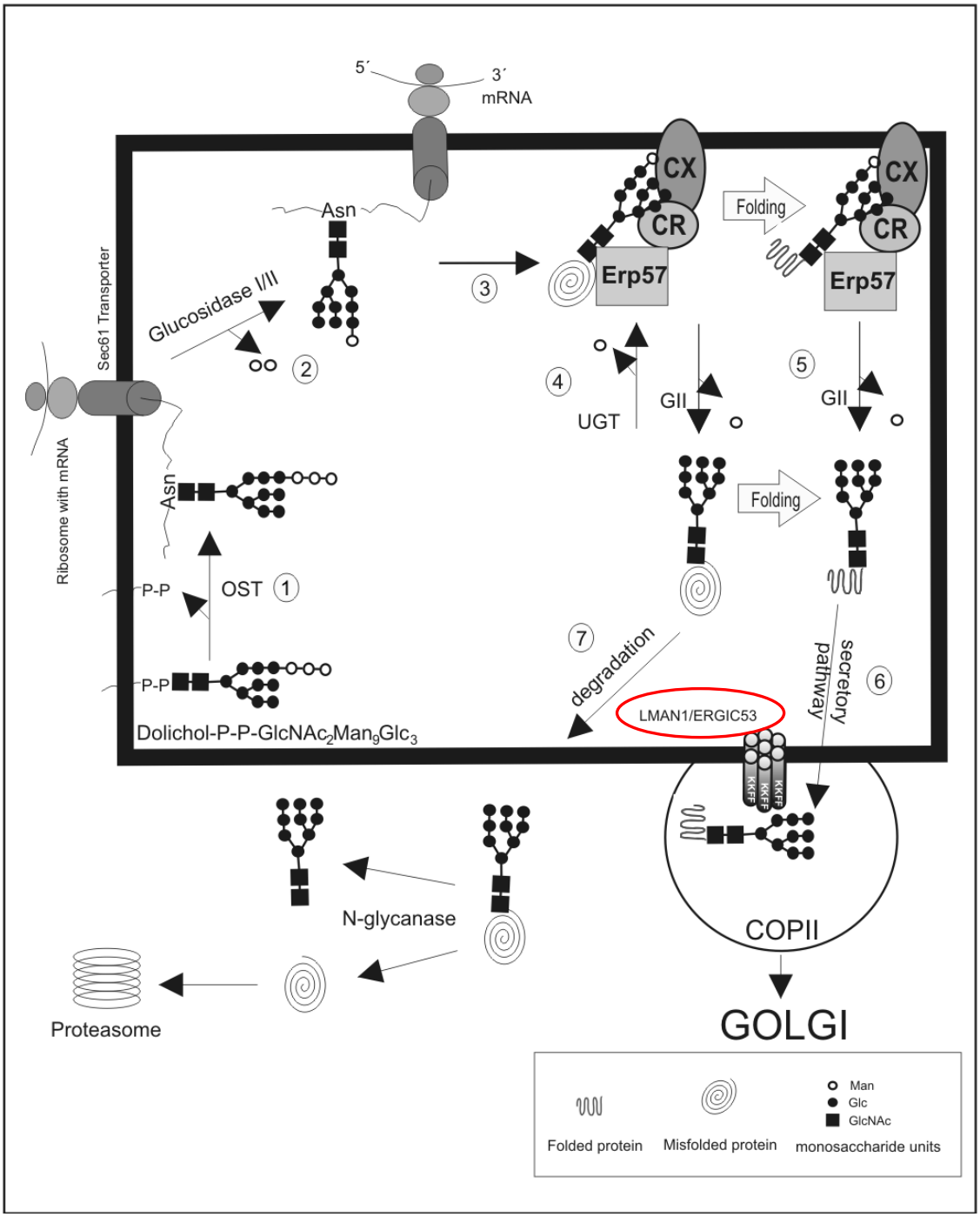


Figure 2.4 Model for the quality control of glycoprotein folding [Varki et al., 2009]. Proteins translocating over the Sec61 transporter into the ER are N-glycosylated by the oligosaccharyltransferase (OST (1)). Two glucose residues are removed by the sequential

action of α-glucosidases GI and GII to generate a mono-glucosylated N-glycoprotein (2) that is recognized by Calnexin (CX) and/or Calreticulin (CR), associated with the protein disulfide isomerase ERp57 (3). Chaperones and protein disulfide isomerase (PDI) Erp57 are part of the N-glycoprotein quality control in the ER. Misfolded glycoproteins are re-glucosylated by the UDP-dependent glucosyltransferase (UGT) (4). Glycoproteins that have acquired their native conformation get hydrolyzed by Gl II removing the remaining glucose residue and are subsequently released from the lectin anchors (5). Correctly folded glycoproteins are not recognized by UGT and are further transported over COPII vesicles through the secretory pathway, in some cases glycoproteins are actively transported by LMAN1/ERGIC53 which serves as another checkpoint control (6). Glycoproteins remaining in misfolded conformations are re-translocated to the cytoplasm, where they are de-glycosylated by N-glycanases and degraded by the proteasome (7).

CX and CR form a complex with protein disulfide isomerase Erp57, exhibiting thioredoxin activity und thus helping the bound glycoprotein to build disulfide bonds. Properly folded glycoproteins are fully de-glucosylated by glucosidase II (Glc II) and the lectins CX and CR release the glycoprotein $GlcNAc_2Man$, which is then packaged in COPII-coated vesicles and transferred to the Golgi [Liu et al., 1999]. Incompletely folded glycoproteins are re-glucosylated by UDP-dependent glycosyltranferases (UGT) and moved out of the ER lumen for proteasomal degradation. UGT is both a folding sensor and a glucosyltransferase, and it is considered to be a decision maker for ER exit of glycoproteins. Further ER-resident proteins involved in quality control are ER ManI and EDEM (ER degradation-enhancing α-mannosidase I–like protein), which trim the mannose residues as an indication for correctly folded glycoproteins [Hebert et al., 2005]. For example, glycan structures free of α-1,2-linked Man are a prerequisite for further trimming and elongation reactions in the Golgi apparatus.

Transport of glycoproteins to the Golgi

Correctly folded glycoproteins can be directly transported via lectin-mediated transport in CopII vesicles towards the Golgi apparatus or the transport is triggered by mature protein conformation-dependent mechanisms [Hauri et al., 2000; Tang et al., 2005]. The lectin-mediated transport also functions as a checkpoint for a correct folding of glycoproteins, by only binding and transporting correctly folded proteins. One prominent example for a lectin is ERGIC53/LMAN1, loading a subset of glycoproteins, e.g. coagulation factors V and VIII, cathepsins C and Z and alpha-

1-antitrypsin (A1AT), into COPII-coated vesicles leaving the ER for the ER-Golgi intermediate compartment (ERGIC).

Further Glycosylation in the Golgi

In the Golgi apparatus glycoproteins are trimmed, elongated and branched, resulting in three main types of N-glycans: oligomannose N-glycans, complex N-glycans and hybrid N-glycans, each still possessing the main core structure Asn-$GlcNAc_2Man_3$ (Fig. 2.5). Further mannosidases and several glycosyltransferases are involved in trimming, elongation and branching. In addition, nucleotide sugar transporters import the necessary monosaccharides, like different UDP-monosaccharides [Mendelsohn et al., 2007; Seelentag et al., 1998].

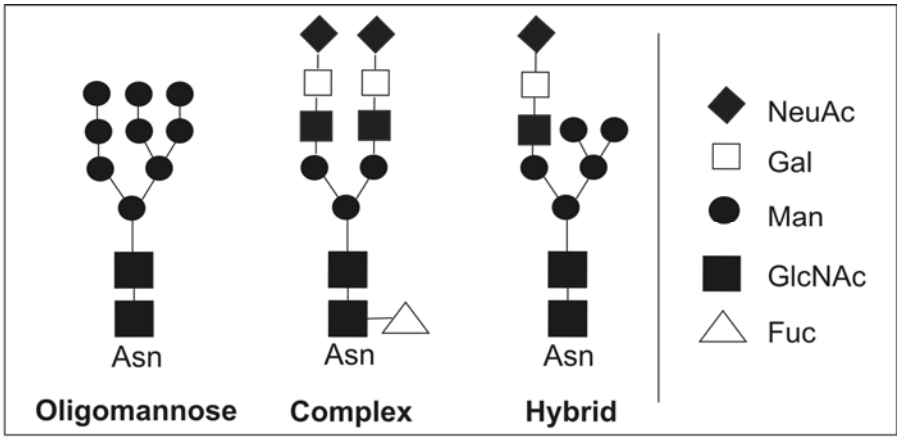


Figure 2.5 Major N-glycans present in mammals. The main N-glycans are oligomannose-, complex- and hybrid-N-glycans all possessing the common core Asn-$GlcNAc_2Man_3$.

Enzymes that catalyze the attachment of terminal sugars

As the N-glycan transits through the medial- and trans-Golgi, it becomes susceptible to glycosyltransferases increasingly localized toward the end of the assembly line and therefore controlling more distal structural modifications such as the addition of sialic acid, glucuronic acid and sulfate. Terminal sugars added by sialyltransferases, glucuronyltransferases and sulfotransferases present the first contact sites for lectins and antibodies.

2.4.2 O-glycosylation

O-glycosylated proteins are linked to the polypeptide chain via serine or threonine residues, without requiring a consensus sequence. O-glycan biosynthesis is simpler than asparagine-linked oligosaccharide generation and is facilitated in the ER and Golgi apparatus [Van den Steen P. et al., 1998]. The initiating step is the addition of a monosaccharide, mainly GalNAc from a nucleotide-activated monosaccharide (e.g. UDP-GalNAc) to a serine or threonine residue, catalyzed by site-specific glycosyltransferases (e.g. GalNAc-transferases). Seven different linkages of O-glycans to proteins are known: O-linked GalNAc (mucin type), O-linked GlcNAc, O-linked Gal (collagen), O-linked Man, O-linked Glc, O-linked Fuc and glycosaminoglycans referred to in 2.4.3.

O-GalNAc structures are also known as mucins. Mucins are defined as cell surface or secreted glycoproteins less branched than most N-glycans and commonly with bi-antennary structures. Mucins (Fig. 2.6) are an essential part of the mucosal protective barrier. Secreted mucins, like MUC2, form a negatively charged gel which interacts with the apical membrane-anchored glycocalyx. In contrast, membrane bound mucins like MUC1 and MUC4 are involved in cell signalling. *MUC2* is highly expressed in mucinous carcinomas and in early colorectal cancers [Allen et al., 1998; Lee et al., 2003] and it is used as a diagnostic marker in colorectal cancer [Lugli et al., 2007]. Eight different core structures are presently known for extension beyond the GalNAc-Ser/Thr linkage with core 1 and core 2 relatively common and core 3 and 4 showing lower occurrence, whereas the remaining structures are relatively rare. Each core determines the final structure of the extended and completed glycans formed through a common biosynthetic pathway [Brockhausen I., 2007; Corfield AP., 2007; Van den Steen P. et al., 1998]. Several glycosyltransferases are involved in the initiating and elongating steps of O-glycan biosynthesis and act in a hierarchical and tissue-specific manner [Brockhausen I., 2007]. Each glycosyltransferase possesses a substrate specificity determining the resulting O-glycan structures. Galactosyltransferases (GalT) and N-acetylglucosaminyl transferases (GlcNAcT) catalyze the core structures presented in Figure 2.6. Incompletely glycosylated cores are known as Tn- and T-antigens. Usually these

precursors are further elongated and terminally linked with fucose and sialic acid. Also Lewis histo blood group antigens are fucosylated and sialylated at the terminus of mucins (Fig. 2.8 A). Other modifications are sulfation, acetylation and methylation.

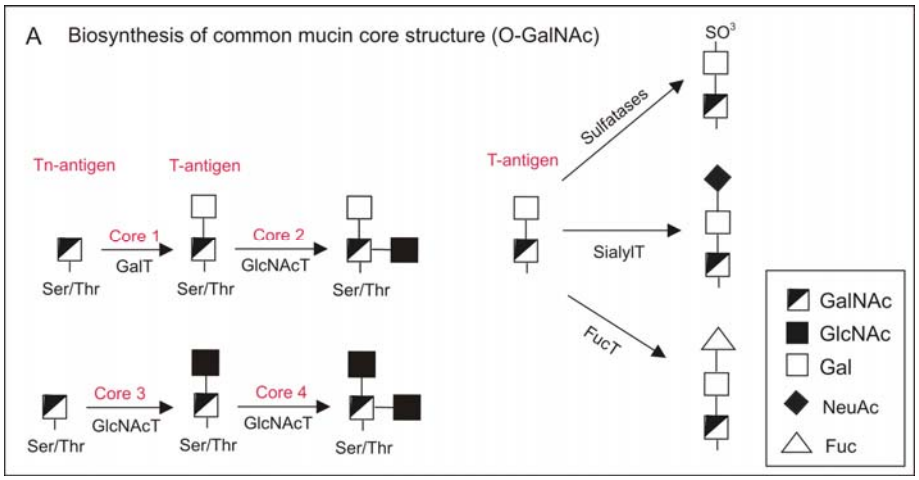


Figure 2.6 Biosynthesis and structure of O-GalNAc structures. The linkage of *N*-acetylgalactosamine to serine or threonine to form the Tn antigen, catalyzed by polypeptide-N-acetylgalactosaminyltransferases (ppGalNAcTs), is the basis for all core structures.

Non-mucin O-glycans regulate protein turn over by interacting with kinases and phosphatases [Hart et al., 2007; Yang et al., 2006], are common structures in the epidermal growth factor (EGF) domains and thereby involved in Notch signalling [Luther et al., 2009], control cell adhesion by binding to components of the extracellular matrix [Barresi et al., 2006], and can be a major part of the connective tissue in form of collagens [Liefhebber et al., 2010].

2.4.3 Glycosaminoglycans: Components of Proteoglycans

Proteoglycans (PGs) and the attached glycosaminoglycans (GAGs) are components of the extracellular matrix (ECM) and the cell surface involved in many biological processes including extracellular matrix deposition, biomechanical lubrication, cell-cell interactions, tumor cell growth, viral infections or neurite outgrowth [Iozzo, 1998].

PGs consist of a core protein and one or more covalently attached GAG chains. GAGs are linear building blocks composed of disaccharide units of an

amino sugar (GlcNAc or GalNAc) and an uronic acid (GlcA and IdoA). These disaccharide units are attached to all GAGs forming long unbranched chains of 50 to 150 units. For three GAG types the first step in PG synthesis is the assembly of a tetrasaccharide linker region (serine-Xyl-Gal-Gal-Glc), in which an UDP-xylosyltransferase (XYLT) catalyzes the initial reaction by transferring xylose from UDP-xylose to selected serine residues of the core protein. This reaction is followed by sequential action of UDP-galactosyltransferase and UDP-glucuronyltransferase in the ER/Golgi compartments. If assembly occurs on this tetrasaccharide linker, heparan sulfate (HS), dermatan sulfate (DS) or chondroitin sulfate (CS) can be synthesized. Chondroitin sulfates differ from dermatan sulfates by containing iduronic acid (IdoA), which is the post-synthetic epimerization of glucuronic acid (GlcA) at position C5. For these three GAGs two xylosyltransferases XYLT1 and XYLT2, differentially expressed in human tissues, catalyze the rate limiting step [Götting et al., 2007].

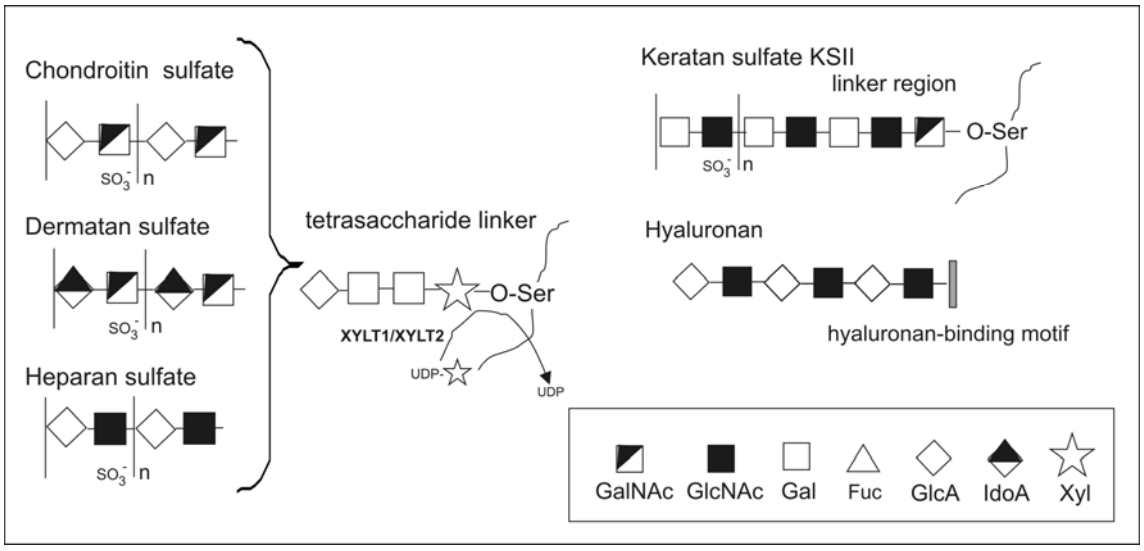


Figure 2.7 Structure of disaccharide repeat units in O-linked glycosaminoglycans (GAGs). GAGs consist of repeating disaccharide units composed of N-sulfated GalNAc or GlcNAc sugars and either an uronic acid (GlcA or IdoA) or Gal. IdoA results from postsynthetic epimerization of GlcA at C5. Hyaluronan lacks sulfate groups, but the rest of the GAGs contain sulfates at various positions. DS is distinguished from CS by the presence of iduronic acid (IdoA). The xylosyltransferases XYLT1 and XYLT2 present initial enzymes in proteoglycan synthesis of dermatan, chondroitin and heparan sulfate. Keratan sulfates are not established via a tetrasaccharide linker, lack uronic acids and instead consist of sulfated Gal and GlcNAc residues.

Another linkage without the tetrasaccharide linker is found in keratan sulfates (KS), which are characterized by a molecular heterogeneity. KS I molecules are N-linked to asparagine over GlcNAc and possess GlcNAc-Gal disaccharide units, whereas KS II molecules are O-linked to serine or threonine residues via GalNAc also possessing GlcNAc-Gal as repetitive units. Both KS types are highly sulfated. The fifth main type of GAGs is hyaluronan. Hyaluronan synthesis is catalyzed by hyaluronan synthases (HAS), each of which contain dual catalytic activities required for the transfer of GlcNAc and GlcA units from the corresponding nucleotide sugars. Figure 2.7 shows the composition of the five GAGs mentioned.

Four main PG-structures exist in mammals, presented by extracellular aggregates, interstitially localized PGs, PGs associated with the basement membrane and membrane bound PGs. Interstitially localized extracellular large aggregating (hyaluronan-binding) PGs are involved in cellular attachment, cell proliferation and differentiation by interacting with cell surfaces and ECM molecules [Day, 1999; Hascall et al., 1974; Iozzo et al., 1996; Spicer et al., 2003]. Extracellular leucine-rich PGs, named small leucine-rich proteoglycans (SLRPs), interact with collagens that form the framework of the ECM. In addition, biglycan and decorin can bind growth factors and influence their bioactivity [Bengtsson et al., 1995; Iozzo, 1999; Kinsella et al., 2004; Matsushima et al., 2000]. The function of PGs in the basement membrane include permeability control by acting as a selective barrier for macromolecules, but they also serve as an adhesive matrix for endo- and epithelial cells and they can control the invasion of cancer cells [Iozzo, 2005]. Cell-surface PGs, like the syndecan family act as co-receptors, by immobilizing ligands, presenting them to a specific receptor, and preventing their degradation. In addition they mediate the adhesion of skeleton structures and facilitate focal adhesions. Their ectodomain can be shed from the cell surface by enzymatic cleavage mediated by endoglyosidic enzymes, like HS-specific endoglycosidases. HS-degradation in mammalian cells is mediated by a single dominant endoglycosidase, named heparanase [Parish et al., 2001]. Shed syndecans have been found in serum and in body fluids during wound healing and cancer, probably altering the integrity and functional state of tissues and providing a mechanism by which cells can respond rapidly to changes in the extracellular environment [Vlodavsky et al., 2001]. Enzymatic degradation of HS is, therefore,

likely to be involved in fundamental biological phenomena, ranging from pregnancy, morphogenesis, and development to inflammation, angiogenesis, and cancer metastasis [Bernfield et al., 1992].

2.4.4 How, where and why are proteins glycosylated?

After protein synthesis, proteins are post-translationally modified in the ER and Golgi apparatus. Polypeptide chains are trimmed, serine, threonine, or tyrosin residues are phosphorylated or sulfated, and oligosaccharides can be added. Many extracellular proteins in the serum, urine, saliva and lymph as well as integral proteins and proteins on the extracellular membrane are glycosylated. Length, charge and sequence of sugar chains depend on the species, tissue, age and the condition of the organism, like pregnancy, disease or cancer. The glycosylation of proteins has various functions: glycosylation protects proteins from proteolytic digest, can change the affinity of receptor-ligand interactions and the activity of certain hormones and enzymes. It can serve as a signal in intracellular transport and the interaction of sugar chains in proteins of extracellular matrices and membranes, regulating migration and distribution of cells in the organism (Kobata et al., 1992).

A specific glycoprotein leaving the Golgi apparatus may carry different types of glycans depending on the situation and the different purposes to fulfill. The complexity of the enzymatic reactions and the topological order of the enzymes involved along the assembly line of glycoprotein synthesis determine the display of glycans on the cell surface and the sorting of glycosylated proteins. N-glycan structures are found on cell receptors for signalling and interaction processes (e.g. EGFR, FGFR, PDGFR and integrin/cadherin induced signaling), on immunological proteins (immune globulines, NK cells) and N-glycans can form a molecular lattice with galectins, opposing glycoprotein endocytosis. N-glycans are thereby involved in several cellular processes like cell adhesion, self/nonself recognition, molecular trafficking and clearance, receptor activation and endocytosis [Ohtsubo et al., 2006]. O-glycans, as well as GAGs in proteoglycans, possess physical and structural functions as protective barriers (e.g. mucins) or control cell-adhesion and cell signalling by mediating cell-cell and cell-matrix interactions. Furthermore they are components of the extracellular matrix (ECM) and the cell surface. [Hart et al.,

2007; Liefhebber et al., 2010; Luther et al., 2009; Parish et al., 2001; Spicer et al., 2003; Yang et al., 2006]. Overall, with these several functions in fundamental biological processes alterations in these structures could lead to abnormalities, which are also considered to be functional in tumorigenesis.

2.5 Glycans involved in tumorigenesis

Altered glycosylation is a universal feature of cancer cells, and certain glycan structures are well-known markers for tumor progression. Glycosylation plays a principal role in a number of cellular processes of key importance for tumorigenesis. Enzymes and transporters mediating growth receptor regulation, growth factor modulation, cell-cell adhesion, immune system modulation, cell motility or the adhesion to endothelium are of great impact for normal cell behavior [Potapenko et al., 2010]**.** If any of these important key players is mutated, normal cell behavior is changed, which may affect adhesion properties and cell-cell signalling as well as modulate the immune response.

2.5.1 N-glycans and cancer

Three main types of N-glycans, oligomannose N-glycans, complex N-glycans and hybrid N-glycans, exist which are further trimmed, elongated and branched by mannosidases and glycosyltransferases. To yield complex tri- and tetra-antennary N-glycans additional branches can be initiated by GlcNAcT-IV (*MGAT4*) and by GlcNAcT-V (*MGAT5*). GlcNAc-branched complex-type N-glycans are found among others on glycoprotein receptors. A striking example of such glycan receptor-mediated function is the binding of galectins to MGAT5-modified N-glycans at the T-cell receptor reduced T-cell activation and autoimmunity by preventing T-cell recruitment to antigen presentation sites [Demetriou et al., 2001]. The interaction between galectins and branched N-glycans produces a thermodynamically stable array of galectins and glycoproteins at the cell surface, called the galectin lattice (Fig. 2.8; [Demetriou et al., 2001]). Expression of *MGAT5* and branched N-glycans are increased in various carcinomas [Dennis et al., 1999], while galectin-3 expression is associated with tumor progression and cancer metastasis in several tumor types [Takenaka et al., 2004]. Cell lines with increased *MGAT5* expression show an increased frequency of metastasis in animal models [Goetz, 2009]. The plant lectin L-phytohemagglutinin (PHA-L) emerged as a marker for preferentially recognizing branched N-glycans bearing the β1-6 branched GlcNAcT-V product in tumor tissues [Przybylo et al., 2008].

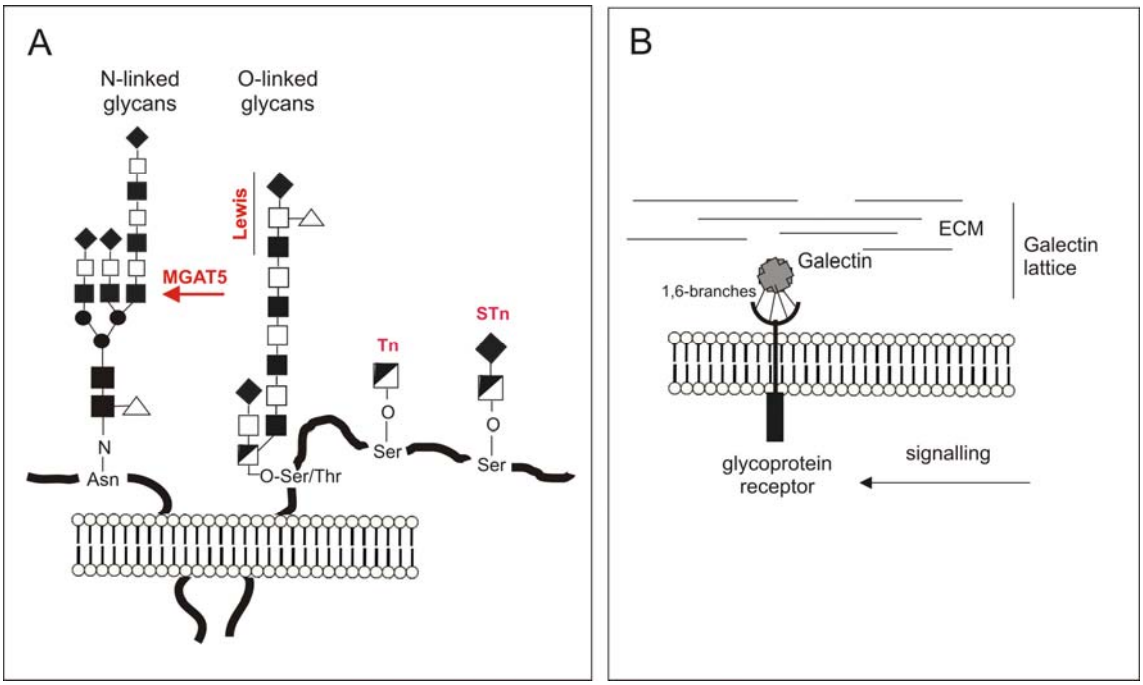


Figure 2.8 N- and O-glycans involved in tumor progression. **A.** Complex N-glycans contain two or more antennae. The O-glycans shown are of the O-GalNAc type, also known as mucins. Lewis antigens are carbohydrate determinants composed of four sugars in specific linkage to one another, and are commonly over-expressed on tumor-cell mucins. Tn and STn are tumor antigens that consist of truncated/incompleted O-linked chains. In many tumors their accumulation correlates with invasion. **B.** shows the involvement of branched complex N-glycans, generated by the enzyme MGAT5. Galectins bind branched complex N-glycans on glycoprotein receptors. Together they build the galectin lattice. Expression of *MGAT5* and branched N-glycans is increased in various carcinomas [Fuster et al., 2005; Goetz, 2009].

2.5.2 O-glycans and cancer

O-linked glycans, namely mucins, regulate adhesion and modulate immune response. The synthesis of core structures 1 – 4, and rare core structures 5 – 8, is initiated by GalNAc transferases transferring GalNAc to either a serine or a threonine of the core protein [Ten Hagen et al., 2003]. Mucins are produced by normal mammalian cells and contain a mixture of extended core 2 structures. Mucins play a role in invasion, metastasis and protection [Hollingsworth et al., 2004]. Cancer cells frequently synthesize mucins with incomplete oligosaccharide chains which are often sialylated, resulting in aberrant core 1 structures. The synthesized structures are called Tn-antigen (GalNAc-α1-O-serine/threonine) and STn-antigen (Siaα2-6GalNAc-α1-O-serine/threonine), which are markers for poorly

differentiated adenocarcinoma and their increased occurrence is associated with advanced colon cancer [Itzkowitz et al., 1990]. Changes in these mucins influence growth and survival of the cell, because potential ligands responsible for interactions between cancer cells and their environment are changed. Selectins, which mediate cell adhesion by acting as glycoprotein receptors, were reported to bind mainly carcinoma mucins [Aychek et al., 2008]. In addition, selectin binding to cell surface-bound mucins may also transmit signals into cancer cells during cancer progression, e.g. by enhancing growth factor receptor signalling [Hollingsworth et al., 2004; Kufe, 2009].

2.5.3 Proteoglycans and cancer

Heparan sulfates (HS) are a central part of the ECM and play an important role in many biological processes [Vlodavsky et al., 2001]. As mentioned in section 2.4.3, enzymatic degradation of HS in syndecan-molecules of the ECM is involved in fundamental biological processes, ranging from pregnancy, morphogenesis, and development to inflammation, angiogenesis, and cancer metastasis [Bernfield et al., 1992]. The HS-specific endoglycosidase, named heparanase, showed an increased expression in human malignancies compared with the corresponding normal tissue [McKenzie et al., 2000]. Tumor angiogenesis can be processed by the remodelling of the vascular system (Fig. 2.9). The importance of mammalian heparanase in vascular remodelling was demonstrated [Zcharia et al., 2004]. The cleavage of heparan sulfates in the ECM might provide pro-angiogenic factors, including growth factors (VEGF, FGF2, PDGF and TGFß), that in addition bind heparan sulfates, thereby promoting tumor angiogenesis [Bogenrieder et al., 2003; Jiang et al., 2003].

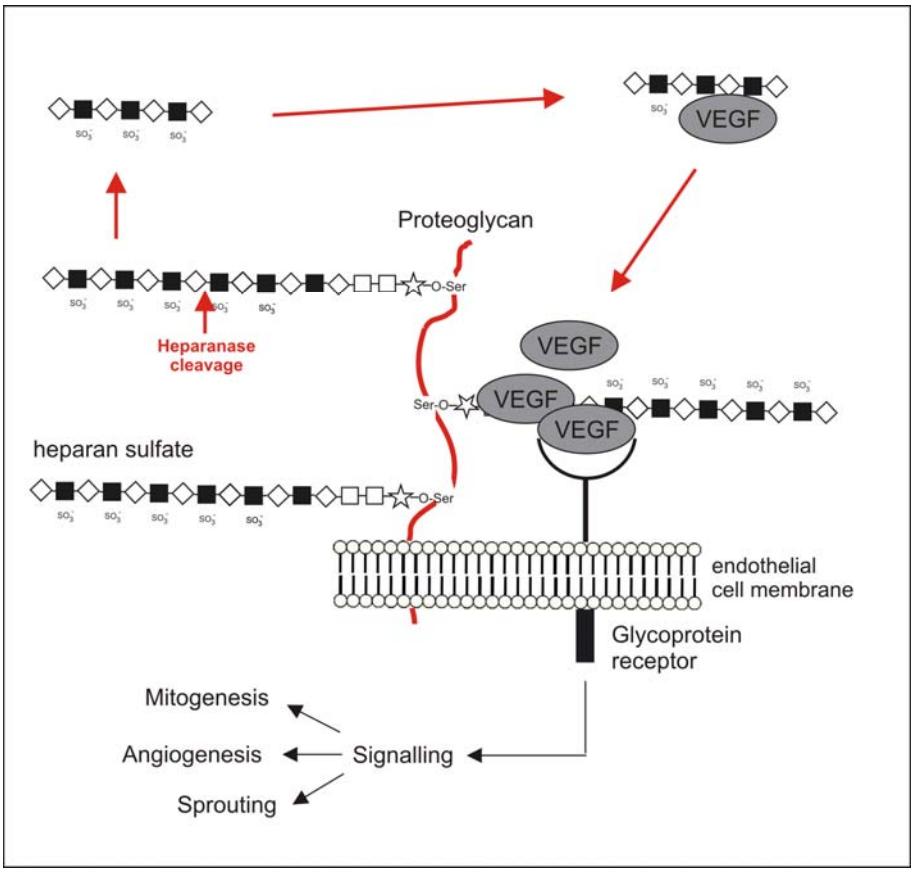


Figure 2.9 Proteoglycans involved in tumor angiogenesis. Tumor angiogenesis is required for tumor growth and the possibility of the tumor to metastasize. Heparan sulfate on the cell surface (ECM) facilitates growth-factor binding and activation of endothelia glycoprotein receptors [Fuster et al., 2005].

2.5.4 Glycans and MSI CRC

The previous sections gave an overview on the important and various functions of several glycoproteins and their involvement in tumorigenesis. In particular, glycosylation of proteins which is mediated by a framework of genes is changed in many tumor entities, affecting important cellular processes. For sporadic MSI tumors increased *MUC2* gene expression was reported and changed MUC structures may be an important step in the neoplastic molecular pathway of colon tumors [Pastrello et al., 2005]. In contrast, MUC2 loss was reported to be an appropriate prognostic marker in MSI tumor progression [Lugli et al., 2007]. Oligonucleotide microarray analysis of MSI-H and MSS colorectal carcinomas showed that expression profiles can be distinguished in both cancers [Kim et al., 2004]. For MSI-H carcinomas three

genes with an involvement in glycosylation were shown to be upregulated, *SIAT4B*, *B4GALT1* and *GALNT5*. Apart from these preliminary data on glycosylation genes in MSI CRC, aberreant glycosylation and glycosylation pathways have not been investigated in MSI-H tumors. Detailed analyses of glycosylation genes and their involvement in MSI tumorigenesis will be part of this thesis. The specific aims are discussed in the next section.

2.6 Aims of the work

The accumulation of tumor-promoting coding microsatellite mutations is considered to be the driving force in MSI colorectal carcinogenesis [Furlan et al., 2002; Woerner et al., 2001]. Frameshift mutations in coding regions lead to a loss of protein function or the mutated protein is translated with a newly synthesized carboxy-terminal neopeptide [Linnebacher et al., 2001; Schwitalle et al., 2008]. The strong immune response in MSI cancers could be explained by these tumor-specific neopeptides. However, also deregulated protein glycosylation in cancer cells might lead to the formation of highly immunogenic tumor specific glycoantigens, mainly located on the cell surface. Besides acting as tumor antigens it is likely that they are actively involved in tumor progression and metastasis. Therefore detection of defined carbohydrate and/or glycoprotein tumor antigens and the identification of abnormal glycosylation pathways in tumors is of special interest for improved diagnostic and therapeutic strategies, as well as for better understanding of MSI tumorigenesis [Dube et al., 2005; Dwek et al., 2004; Vlad et al., 2004; Patsos et al., 2009].

The present thesis aimed at the identification and characterization of cMNR harboring genes encoding proteins of the cellular glycosylation machinery with emphasis on frameshift mutations resulting in abberant cell surface presentation or secretion of glycoproteins. Building on that, these alterations in glycoprotein structures or secretion patterns should be further analyzed. The final aim was to establish a basis for the detection of MSI cell-specific glycosylation patterns that can be used for the development of novel diagnostic and therapeutic strategies for MSI tumors, and for the detection of changes relevant to MSI tumorigenesis. To reach these aims, frameshift mutation analysis and functional analysis were combined with biochemical methods. The specific aims of this study are:

I. The identification of cMNR harboring candidate genes encoding proteins of the glycosylation machinery and to determine their mutation frequency in MSI CRC cell lines, adenomas and carcinomas.
II. The analysis of protein expression of candidate genes identified in I. in MSI CRC cell lines and primary tissues.

III. The evaluation of the function of candidate genes mutationally inactivated in MSI CRC cell lines by using methods, such as metabolic labelling and cell surface staining.

IV. To perform protein expression and purification experiments to generate new tools for the identification of relevant interaction partners of candidate genes with an involvement in MSI tumorigenesis.

Overall, frameshift mutation analyses as well as expression and functional analyses should reveal novel insights into candidate genes presumably involved in MSI tumorigenesis. In addition, MSI tumor-specific glycosylation patterns potentially provide the basis for novel diagnostic and therapeutic approaches.

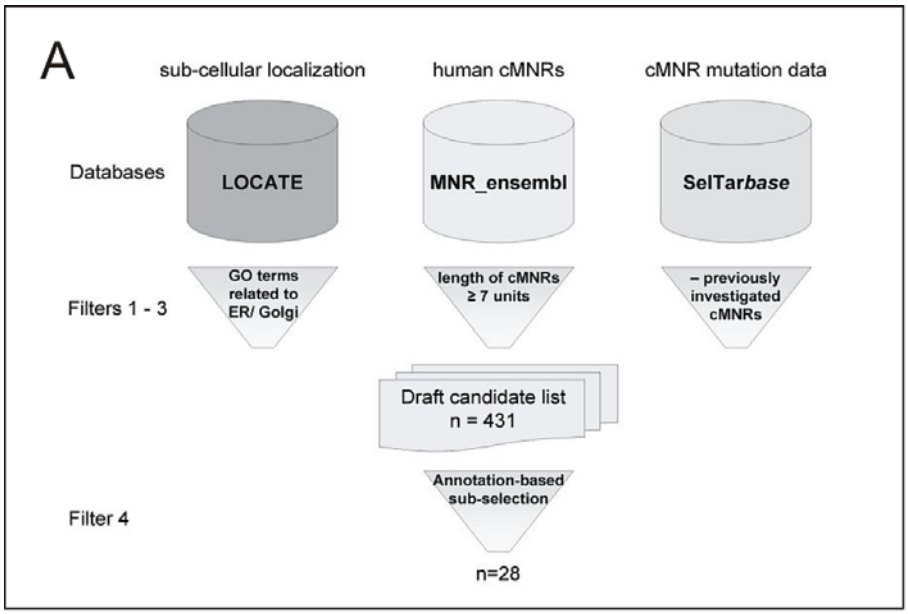
3. Results

3.1 Glycogenes involved in MSI-H colorectal tumorigenesis

Despite the obvious significance of altered glycoprotein synthesis, transport and secretion in CRC, aberrant glycosylation and glycosylation pathways have not been investigated in MSI-H colorectal tumors. In order to identify cMNR-harboring genes, encoding components of the cellular glycosylation machinery candidate genes were selected by a bioinformatics-based approach. Database analysis and specific selection criteria such as subcellular localization, cMNR length, and involvement in glycosylation were applied to pre-select candidate genes.

Potential candidates were selected from three databases: for the first database a filter for the subcellular localization (Locate database) of candidate genes was used, second the MNR_ensembl database was filtered by limiting human cMNRs to seven and more basepairs and third the SelTar*base* was filtered for candidates which were already analysed, e.g. the previously investigated genes *TGFBR2, ACVR2, AIM2* and *UPF3A*. For details on database analysis see section 6.5. These three databases revealed 431 genes and a final short list selection (Filter 4) was based on literature data showing a function of the corresponding gene product in glycan synthesis, modification or transport (Figure 3.1A). As a result, 28 candidate genes were selected for subsequent analyses.

DNA frameshift mutation profiles of these 28 candidate genes with 32 cMNRs were determined on a panel of 48 CRC cell lines (MSS, n = 25, MSI-H, n = 23; Fig. 3.2). Mutation frequencies ranged from 0 to 52%, whereas no frameshift mutations in any of these genes were detectable in MSS CRC cell lines. For *XYLT2* and *LMAN1* even biallelic mutations were found and both genes appeared as positively selected target genes according to the statistical model described in 2.3.1 (Fig. 3.1 B). To exclude that the observed frequencies of selected genes were restricted to cultured cells, mutation profiles were also determined for at least 50 MSI-H colorectal tumors (Fig. 3.3). For this analysis only those candidate genes frequently mutated (≥ 10%) in the cell culture panel were selected. An equal number of MSS colorectal tumors and blood samples from healthy donors served as controls.



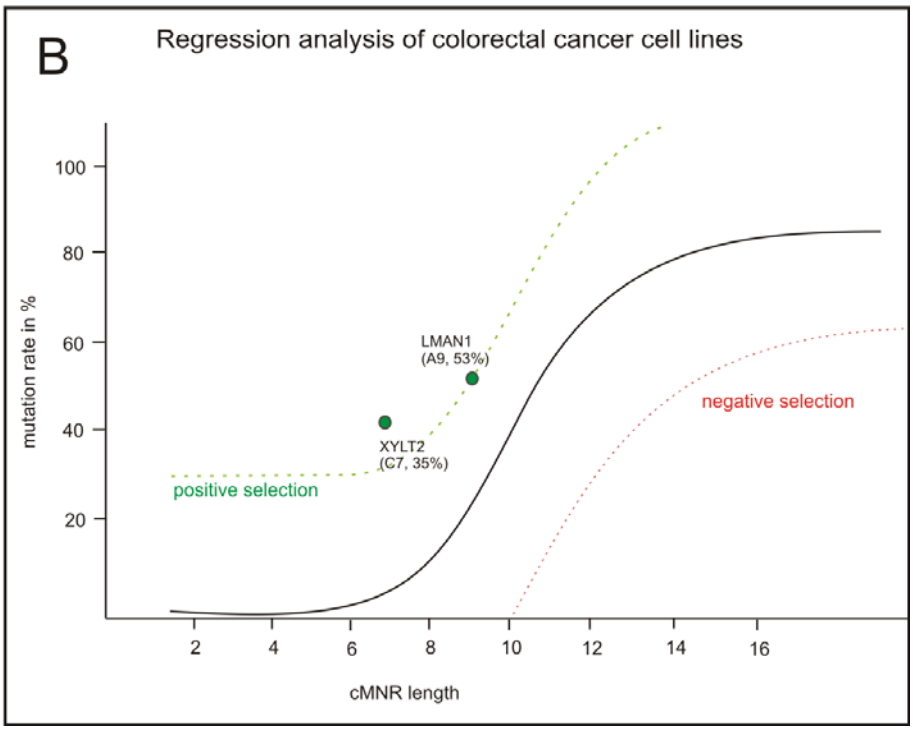


Figure 3.1 Strategy for the identification of candidate genes and the statistical model for the predicition of MSI-H target genes. **A.** Four filters for the selection of candidate genes were applied to three databases. Filters selected by location (1), cMNR-length (2) already analysed cMNRs (3), as well as by literature-based involvement in glycosylation (4). For details see methods. **B.** Regression analysis of cMNR mutations in MSI-H CRC cell lines. *LMAN1* and *XYLT2* analysed as genes involved in the glycosylation

machinery seem to be MSI-H target genes. Green dots represent putative positively selected targets.

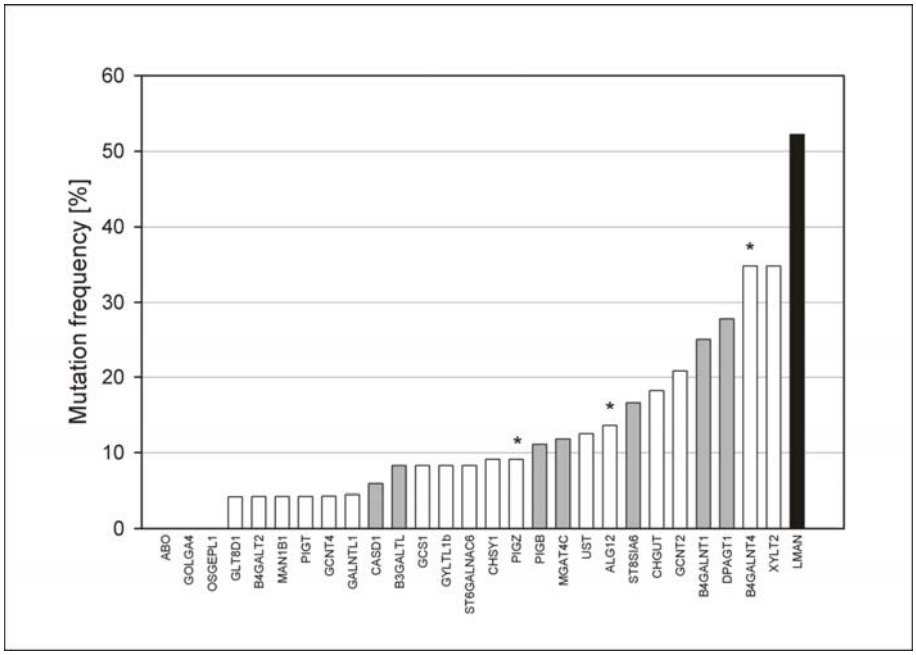


Figure 3.2 Frameshift mutation analysis of candidate glycogenes in MSI-H CRC cell lines. Different wildtype lengths of individual cMNRs are indicated by bar color (white, n = 7; gray, n = 8; black, n = 9). Overall mutation frequencies are depicted. Three genes (*) contained more than one cMNR (*B4GALNT4*, 3 G7; *ALG12*, 2 C7; *PIGZ*, 1 C7 and 1 G7).

Three genes were found to be frequently mutated (≥ 10%) in MSI-H colorectal tumors, namely *LMAN1* (46%), *XYLT2* (26%) and *B4GALNT4* (23%). This finding was corroborated by extended analysis of these 3 genes on an additional panel of MSI-H tumors (*LMAN1 n* = 120, *XYLT2 n* = 55, *B4GALNT4 n* = 43). Besides, *XYLT2* and *LMAN1* harbored biallelic mutations in two (*XYLT2*: 2/23; 8.6%) and four (*LMAN1*: 4/23; 17,4%) MSI-H CRC cell lines investigated. Biallelic mutations imply a complete knockout or a dysfunction for the protein. Because cell lines with biallelic mutations represent powerful tools for further research the two genes *LMAN1*, a glycoprotein transporter and *XYLT2* (*Xylosyltransferase 2*), an initial enzyme in proteoglycane synthesis were selected for further analysis.

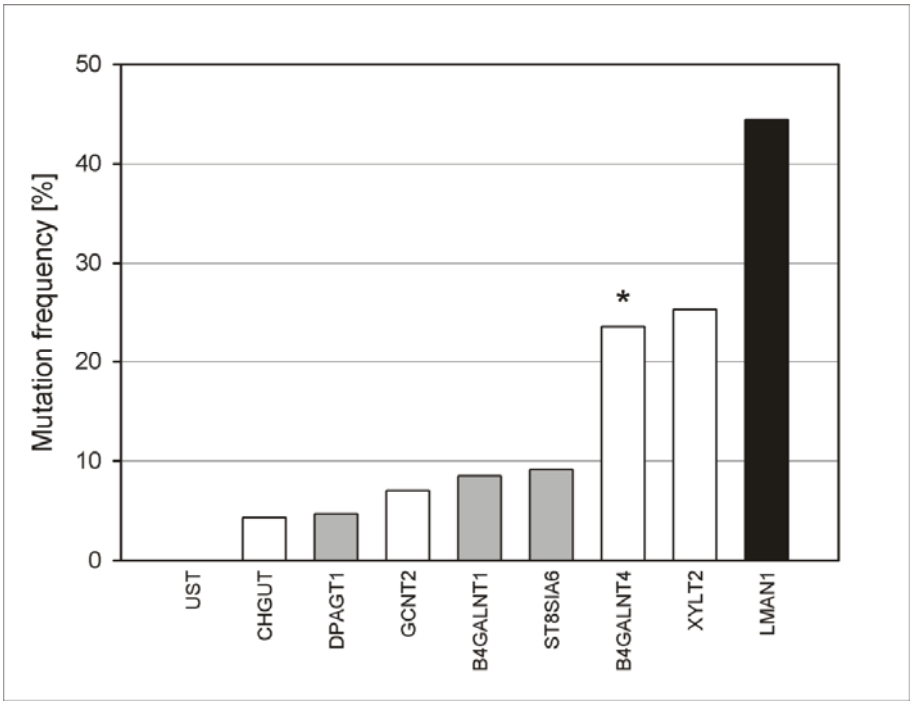


Figure 3.3 Frameshift mutation frequencies of a subset of candidate glycogenes in primary MSI-H CRCs. Gene-specific mutation data were obtained by analyzing 50 MSI-H carcinomas except for *LMAN1* (n = 170 tumors), *XYLT2* (n = 105 tumors), and *B4GALNT4* (n = 93 tumors). Bar characteristics (asterisks, shading) as described in Fig. 3.2.

To evaluate if *LMAN1* and *XYLT2* mutations were early events in MSI-H tumorigenesis a subset of 20 MSI-H colorectal adenomas was analyzed. A significant fraction of these adenomas (*LMAN1* 40%, 8 of 20; *XYLT2* 21%, 4 of 19) showed cMNR frameshift mutations. *LMAN1* gene mutation status was compared with the UICC tumor stages for MSI-H CRCs and a comparison of sporadic and HNPCC MSI-H cancers was examined, but no significant correlation to any UICC stage could be made (data not shown). LMAN1 mutations appeared to be more frequently observed in HNPCC compared to sporadic MSI-H colorectal carcinomas. However, the low number of cases which were available for these analyses revealed no significant differences.

Overall, it was determined that two genes involved in the glycosylation machinery were frequently and also biallelically mutated in MSI-H CRCs. These genes might play a role in MSI-H tumorigenesis. The following sections will present detailed analyses of *LMAN1* and *XYLT2*.

3.2 LMAN1 mutational inactivation in MSI-H colorectal tumorigenesis

3.2.1 *LMAN1* transcript is present, but LMAN1 protein is not detectable in biallelic-mutated cells

To understand the nature of the genetic defects observed in MSI-H CRC cells and tumors (3.1) *LMAN1* expression was analyzed at the transcript and protein level. For the analysis of *LMAN1* transcription cDNAs from 23 MSI-H CRC cell lines, one MSS cell line (SW948) and normal human colonic mucosa were used. *LMAN1* mRNA was found to be expressed in all samples, including cell lines harboring mono- or biallelic *LMAN1* cMNR frameshift mutations and was also found in normal colon tissue (Fig. 3.4A).

To verify the transcription data at the translational level, Western blot analysis with an LMAN1-specific antibody was performed on a panel of *LMAN1*-mutated and *LMAN1*-wild type cells. LMAN1 was detectable in all 19 heterozygously mutated cell lines and in cells expressing wildtype *LMAN1*, whereas four cell lines Colo60H, LoVo, HDC9 and Vaco 6, harboring biallelic *LMAN1* mutations, lacked LMAN1 protein expression (Fig. 3.4B). Although the LMAN1-antibody recognizes an N-terminal epitope, no truncated protein (expected size of 36 kDa) was detected in cell lines with mono- or biallelic cMNR frameshift mutations. *LMAN1* mutated transcript is predicted to be NMD-sensitive with a premature termination codon occurring more than 55 base pairs upstream of the last exon-exon border (explained in 4.2.2) and resulting in the degradation of the transcript. However, the transcript could be detected by RT-PCR in all mutated cell lines.

Expression analysis showed that biallelically *LMAN1*-mutated cell lines were able to transcribe but not to translate a mutated *LMAN1*. The identification of *LMAN1*-deficient cell lines provides a valuable tool for studying the functional consequences of LMAN1 inactivation.

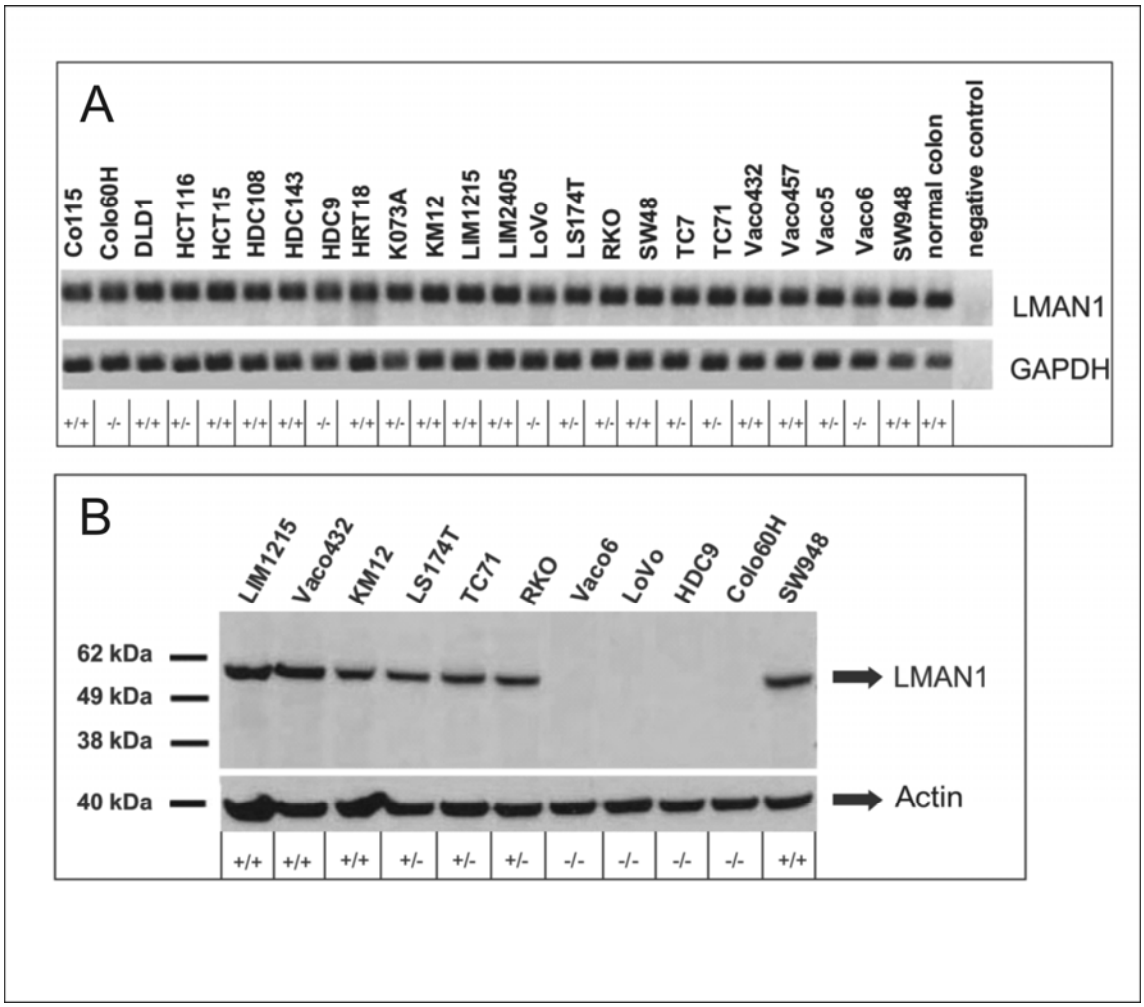


Figure 3.4 LMAN1 expression in MSI-H CRC cell lines. **A.** RT-PCR analysis of *LMAN1* expression in CRC cell lines with homozygous wild-type (+/+), homozygous mutant (−/−), or heterozygous (+/−) cMNR status. The expected PCR fragment of 122 bp could be amplified from cDNA of normal human colon, 1 MSS cell line (SW948), and all MSI-H CRC cell lines. GAPDH was amplified as a cDNA quality control. **B.** A single band of 53 kDa corresponding to the LMAN1 protein was detected in MSI-H and MSS CRC cell lines harboring homozygous wild-type or heterozygous cMNR in MSI-H and MSS cell lines. MSI-H CRC cell lines with biallelic cMNR frameshift mutations lacked LMAN1 protein expression. No truncated LMAN1 protein (expected size, ~36 kDa) was visible in heterozygously or homozygously mutated cell lines. Molecular weight marker bands are depicted. Actin served as loading control.

3.2.2 Complete or partial loss of LMAN1 protein expression in MSI-H CRC tumors

In cultured cells, LMAN1 protein loss was observed as a consequence of biallelical *LMAN1* mutations (see Fig. 3.4B). In order to investigate the expression of *LMAN1* in primary tissues we performed immunohistochemical analysis on a set of 50 MSI-H colorectal carcinomas.

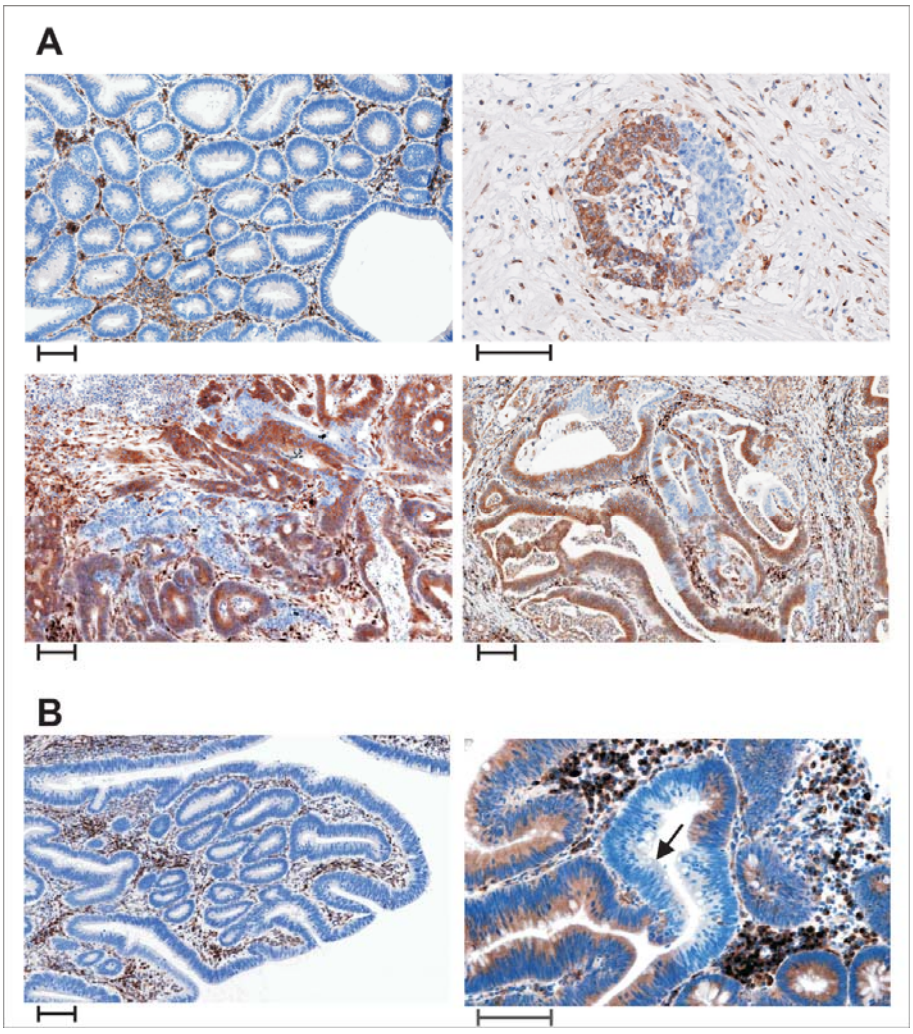


Figure 3.5 Immunohistochemical detection of LMAN1 protein in MSI-H colorectal tumors. Representative immunohistochemical staining patterns of MSI-H colorectal tumors using an LMAN1-specific antibody. LMAN1 was visualized by using 3,3`-beta-diaminobenzidine (brown). Tumor sections were counterstained with hematoxylin (blue). In all tumor sections, positive LMAN1 staining of stromal cells and tumor-infiltrating lymphocytes served as control. Scale bars: 100 µm in all panels. **A.** Four carcinomas that display

either complete or partial loss of LMAN1 expression, sometimes comprising areas with a heterogeneous speckled pattern. **B.** In two MSI-H colorectal adenomas, complete loss (left) as well as focally restricted loss of LMAN1 expression was observed (right). Arrow, crypt with mixed staining.

A significant fraction of these carcinomas showed either local (38%; 19/50) or complete loss (6%; 3/50) of *LMAN1* expression. Representative examples of this heterogeneous staining pattern are presented in Figure 3.5. Moreover, when a small number of MSI-H adenomas were subjected to immunohistochemical staining, a similar pattern (Fig. 3.5B) was observed. Regional loss of LMAN1 expression might be attributable to biallelic mutational inactivation of the *LMAN1* gene in specific tumor areas. In order to address this issue, DNA from areas with or without LMAN1 expression was isolated by regional microdissection on selected MSI-H carcinomas (n = 4) and MSI-H adenomas (n = 3).

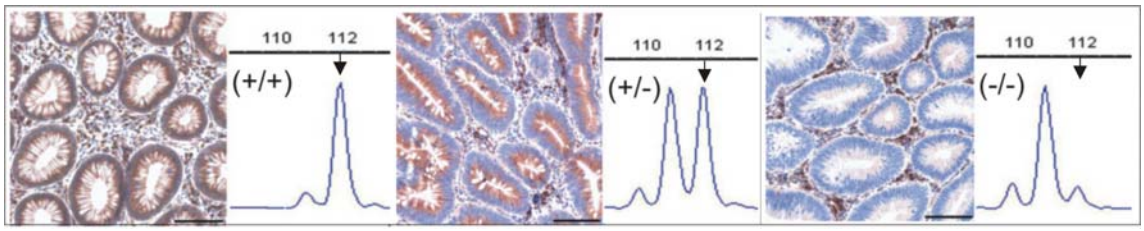


Figure 3.6 Molecular analysis of intratumoral LMAN1-deficient areas. Regional microdissection of a colorectal carcinoma. Non-malignant colorectal epithelium strongly expressing LMAN1 showed only *LMAN1* cMNR wild-type allele peaks (+/+) in electropherograms of fragment length analysis. Tumor areas that have retained LMAN1 staining but with reduced staining intensity compared with non-tumorous stromal cells exhibited an additional peak in frameshift analysis corresponding to a heterozygous state wildtype and one bp deletion at the *LMAN1* coding microsatellite (+/−). Tumor areas completely lacking LMAN1 staining only displayed homozygous mutant allele and peak pattern (−/−). Arrow indicates wild type *LMAN1* and should relieve the shifts of one allele (+/−) or both alleles (−/−), respectively.

DNA frameshift mutation analysis revealed biallelic cMNR frameshift mutations in the *LMAN1* gene in tumor areas and concomitant loss of LMAN1 protein expression (Fig 3.6 right panel). Areas with decreased LMAN1 expression compared to normal tissue (Fig 3.6 left panel) showed heterozygous mutations for *LMAN1* (Fig 3.6 middle panel). Correlation between protein expression and frameshift mutation pattern was also observed in adenomas (data not shown).

From these results it can be concluded that primary MSI-H colorectal tumors frequently show heterogeneous LMAN1 expression abnormalities caused by regional biallelic cMNR frameshift mutations in MSI-H tumor cells. In addition, the occurrence of *LMAN1* mutations in preneoplastic lesions marks an early event during MSI tumorigenesis.

3.3 *LMAN1* expression changes the phenotype in MSI-H CRC cells

It is reasonable to assume that inactivating mutations or altered expression of *LMAN1* cause changes in the transport of glycosylated proteins that follow the secretory pathway. Changes in the glycoprotein export might change the cell's phenotype. In order to identify phenotypical changes in MSI-H CRC cells *LMAN1* full-length cDNA was expressed in *LMAN1*-deficient MSI-H CRC cells. Stable clones were checked for cell viability and LMAN1 expression and subsequently analyzed for phenotypic changes. Glycoprotein profiling on the cell surface was performed by Lectin-FACS analysis.

3.3.1 Stable LMAN1 expression does not change cell viability and growth

For stable expression of LMAN1 a *LMAN1* clone was established by transfecting *LMAN1* cDNA into LoVo cells. LoVo cells show a biallelic mutation in the cMNR of the *LMAN1* gene (A9 → A8). The expression of LMAN1 was checked by Western blot analysis. Compared to an MSI-H CRC cell line endogenously expressing LMAN1 (Co115), the expression intensity was weaker in one *LMAN1*-transfected LoVo clone (Fig. 3.7), while all other clones checked after freeze/thaw cycles lacked LMAN1 expression. This *LMAN1*-transfected LoVo clone permanently expressed the *LMAN1* gene.

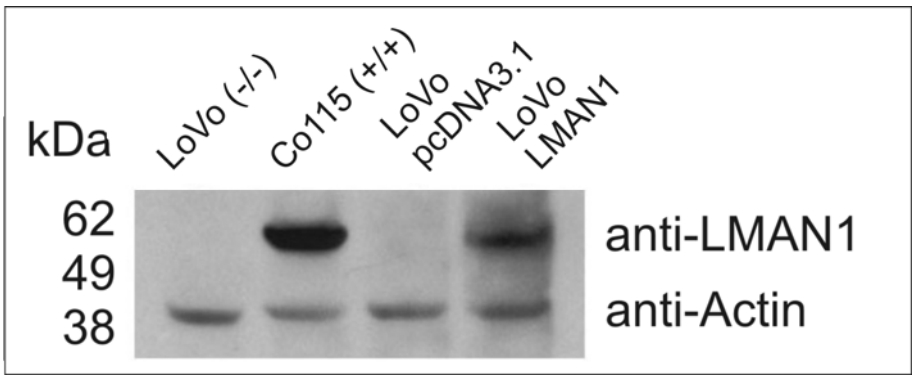


Figure 3.7 Western blot analysis of stably transfected LoVo cells constitutively expressing LMAN1 protein. A single band of 53 kDa

corresponding to the LMAN1 protein was detected in a MSI-H CRC cell line endogenously expressing *LMAN1* (Co115) and in LoVo cells constitutively expressing *LMAN1*. No protein was detected in LoVo cells deficient for *LMAN1* and in mock-transfected LoVo cells expressing the empty vector control (pcDNA3.1). Molecular weight marker bands are depicted. Probing with an anti-actin antibody served as loading control.

LoVo cells stably transfected with *LMAN1* (*LMAN1*-transfected LoVo cells) were characterized by cell viability and compared to LoVo cells stably transfected with the empty vector pcDNA3.1 (mock-transfected LoVo cells). *LMAN1*-transfected LoVo cells did not differ from mock-transfected LoVo cells in a proliferation assay (Fig. 3.8), thus no difference in cell viability was observed. In addition, *LMAN1*-transfected LoVo cells showed no growth advantage or disadvantage compared to the parental line (LoVo) or to mock-transfected LoVo cells.

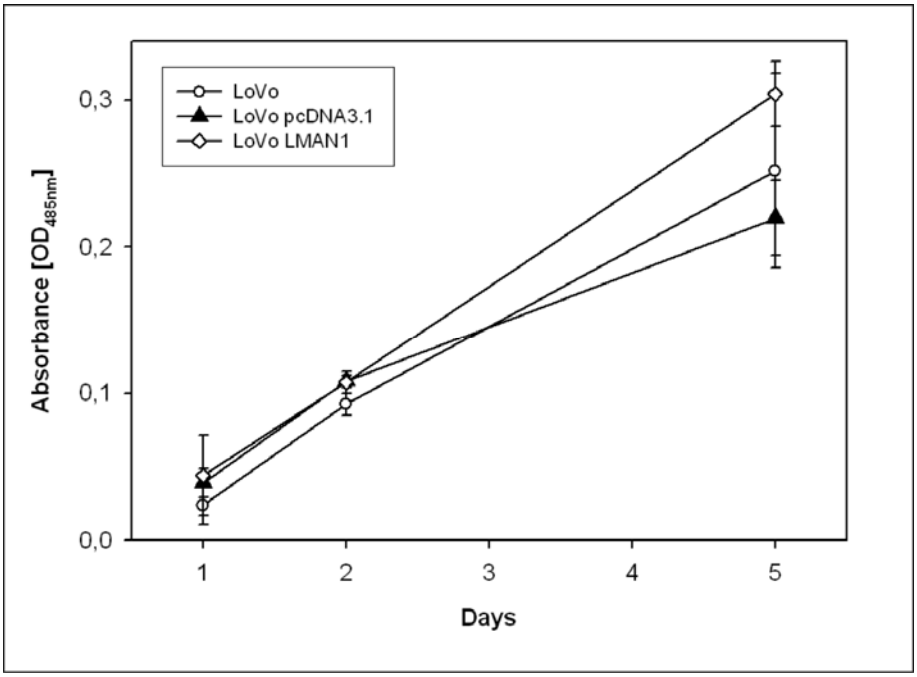


Figure 3.8 Cell Proliferation Assay of stably transfected LoVo cells. 3000 LoVo-, mock-transfected and *LMAN1*-transfected LoVo-cells were sowed per well of a 96-well plate and MTS/formazan solution was added after 1, 2 and 5 days. The MTS tetrazolium compound is bio-reduced by cells into a colored formazan product that is soluble in tissue culture medium. Absorbance was measured at 485 nm. Formazan is directly proportional to the number of living cells in culture. Values represent triplets for each cell line and day. Error bars represent standard deviations.

3.3.2 *LMAN1* expression changes the glycan profile on the cell surface

LMAN1 is a mannose-specific lectin and reconstituted *LMAN1*-expression could change the cell surface pattern by the transport of specific LMAN1-cargo proteins that function as glycoproteins on the cell surface. To compare the cell surface pattern in *LMAN1*-deficient and *LMAN1*-reconstituted LoVo cells, Lectin-FACS analysis was performed. The experiment was established using *LMAN1*-transfected LoVo cells or mock-transfected LoVo cells (3.3.1). A selection of biotinylated plant lectins, specific for distinct building blocks of human glycans was used (Tab. 3.1) with ConA, LEA, PSA and PHA-E specifically detecting mannose-harboring structures (Tab. 3.2 in bold).

Table 3.1 Lectin panel for glycan profiling of cell surfaces and effect of LMAN1 expression

Plant name	Acronym	LMAN1 effect in %	Oligosaccharide
Canavalia ensiformis	**ConA**	no change	GlcNAcβ2Manα6(GlcNAcβ2Manα3)Manβ4GlcNAc
Lycopersicon esculentum	**LEA**	no change	Core and stem regions of high-mannose-type N-glycan
Pisum sativum	**PSA**	+ 14,6 (p-value:0,281)	N-Glycan binding enhanced by core fucosylation
Phaseolus vulgaris erythroagglutinin	**PHA-E**	no change	Bisected complex-type Nglycans: Galβ4GlcNAcβ2 Manα6 (GlcNAcβ2Manα3) (GlcNAcβ4)Manβ4GlcNAc
Phaseolus vulgaris leukoagglutinin	PHA-L	− 22,2 (p-value:0,057)	Tetra- and triantennary N-glycans with β6-branching
Maackia amurensis I	MAA-I	no change	Neu5Acα3Galβ4GlcNAc/Glc, 3 -sulfation tolerated
Artocarpus integrifolia	JAC	+ 28,5 (p-value:0,125)	Galβ3GalNAcα, α3-sialylation tolerated
Triticum vulgare	WGA	+ 19,9 (p-value:0,267)	(GlcNAc)n, Galβ4GlcNAcβ6Gal
Sambucus nigra	SNA	no change	Neu5Acα6Gal/GalNAc, clustered Tn antigen
Viscum album	VAA	no change	Galβ2(3)Gal, Galα3(4)Gal, Galβ3(4)GlcNAc without/with α2,6-sialylation, Fucα2Gal
Arachis hypogaea	PNA	no change	Galβ3GalNAcα

Lectin panel used to identify phenotypic changes on the cell surface of mock-transfected and *LMAN1*-transfected LoVo cells. Plant name, acronym and oligosaccharide specificity for each

lectin are given. The column "*LMAN1* effect" shows the lectin binding change upon *LMAN1* expression. Mean fluorescence intensities of *LMAN1*-transfected LoVo cells were compared to mock-transfected LoVo cells. (+) for increased and (–) for decreased binding. Values presented are the mean of the difference in mean fluorescence of four independent experiments. P-values are indicated.

Both, the percentage of positive cells and mean fluorescence intensity were measured for comparison. For *LMAN1*-transfected LoVo cells a significant decrease in cell surface PHA-L lectin binding was observed. Changes in lectin binding were also detected for PSA, Jacalin and WGA, but this trend remained statistically insignificant. There were no changes on the cell surface for any other lectins measured (Table 3.1). Figure 3.9 represents the cell surface profiling of the lectins PHA-L and jacalin (JAC). Interestingly, changes on the cell surface were observed upon *LMAN1* reconstitution, but no mannose-specific lectin showed changes upon *LMAN1* re-expression in stable LoVo cells. With PHA-L detecting tetra- and tri-antennary N-glycans and jacalin specific for Galβ3GalNAcα-structures (galactosyl (ß-(1,3) N-acetylgalactosamine), a N-glycosidically lectin and an O-glycosidically lectin showed the highest effects after *LMAN1* reconstitution.

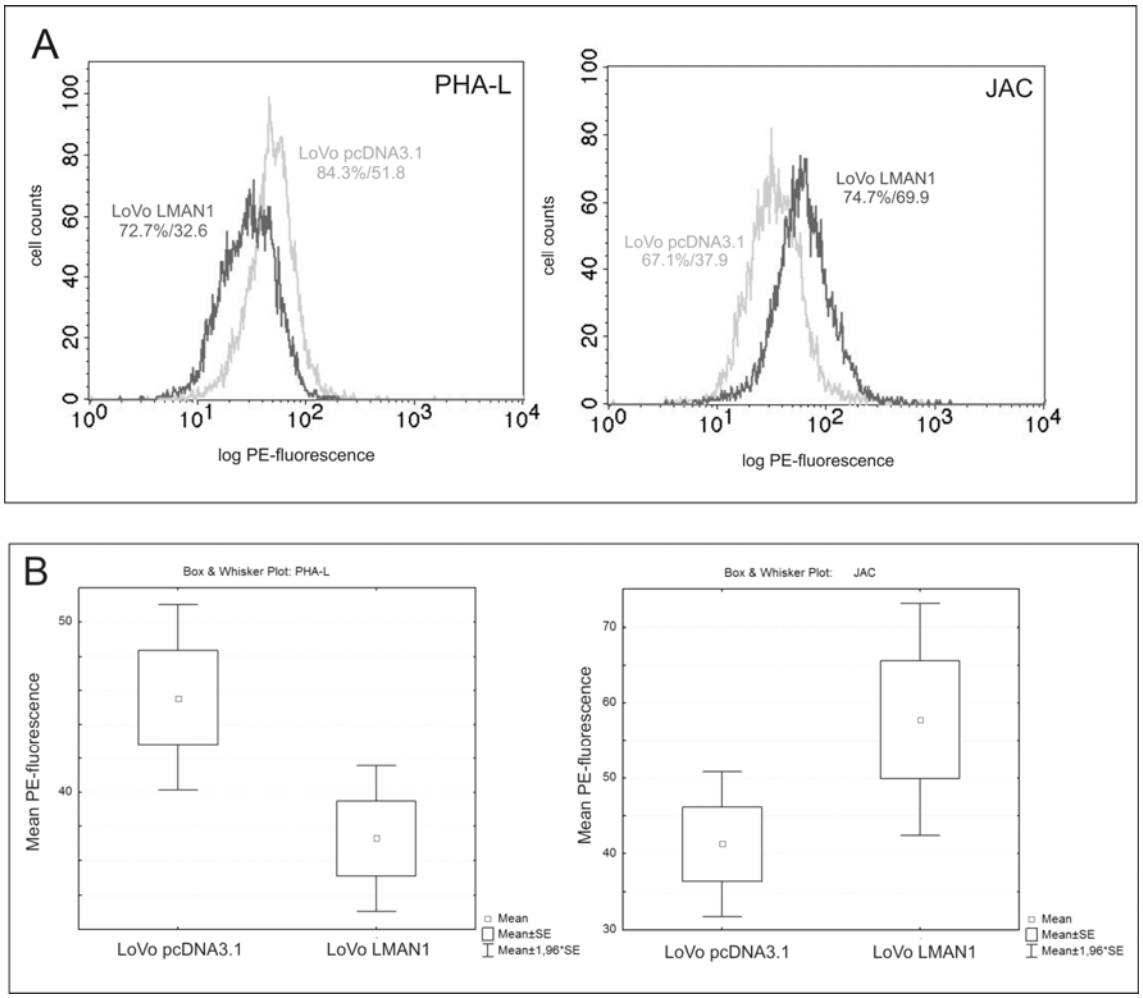


Figure 3.9 FACS histograms and Box & Whisker Plots for lectins PHA-L and JAC. **A.** Semilogarithmic representation of fluorescent surface staining of LoVo cells. Cells were treated with biotinylated lectins and subsequently incubated with Streptavidin-Phycoerithrin (PE) as fluorescent dye. Histograms represent the fluorescence intensity for both mock-transfected LoVo cells (LoVo pcDNA3.1; light grey) and *LMAN1*-transfected LoVo cells (LoVo LMAN1; dark grey). The values for the percentage of positive cells and the mean are given. JAC lectin binding was increased upon *LMAN1* expression whereas PHA-L abundance on the cell surface decreased upon *LMAN1* expression. Experiments represent data from one out of four independent measurements. **B.** Box & Whisker Plots representing the PHA-L and JAC mean fluorescence values of four independent experiments. Both lectins showed effects upon LMAN1 reconstitution. Mean fluorescence, standard deviation and standard error are presented within the Box & Whisker Plot.

3.4 Analyses of LMAN1 cargo proteins relevant to MSI-H CRC

As a transporter for mannose-harboring glycoproteins, LMAN1 plays an important role in the secretory pathway enabling the transport from the ER to the Golgi via the ERGIC (ER-Golgi Intermediate Compartment). Several LMAN1 cargo proteins are known, including the serine protease inhibitor alpha-1-antitrypsin (A1AT), the lysosomal glycoproteins cathepsin C and Z as well as the blood coagulation factors V and VIII [Nyfeler et al., 2008; Appenzeller et al., 1999; Moussalli et al., 1999; Vollenweider et al., 1998]. To get more insight into LMAN1`s role as a cargo protein receptor, different strategies for the identification and characterization of LMAN1 cargo proteins were performed.

3.4.1 Alpha-1-antitrypsin secretion is impaired in *LMAN1* deficient cells

A1AT is an inhibitor of several serine proteases involved in inhibiting angiogenesis and tumor growth [Huang et al., 2004]. A1AT also exhibits important regulatory and anti-inflammatory roles [Kalsheker, 2009]. Using an A1AT-specific ELISA, A1AT secretion in *LMAN1*-deficient and -proficient MSI-H CRC cells as well as *LMAN1*-transfected cells and mock-transfected cells was analyzed. Significantly decreased levels of A1AT were measured in a conditioned medium of *LMAN1* deficient cells (~ 5 µg/L) compared to a conditioned medium of *LMAN1* proficient cells (~ 15 µg/L, Fig.3.10 A). To test whether *LMAN1* re-expression could overcome this secretion defect, *LMAN1*-deficient Colo60H cells were transiently transfected with *LMAN1*-cDNA. A 10-fold increase in A1AT concentration in the medium of Colo60H overexpressing *LMAN1* compared with the empty vector control was observed (Fig. 3.10B). RT-PCR analysis revealed *A1AT* transcription in all cell lines and thus demonstrated that *LMAN1* re-expression had no effect on the *A1AT* transcript levels (Fig.3.11).

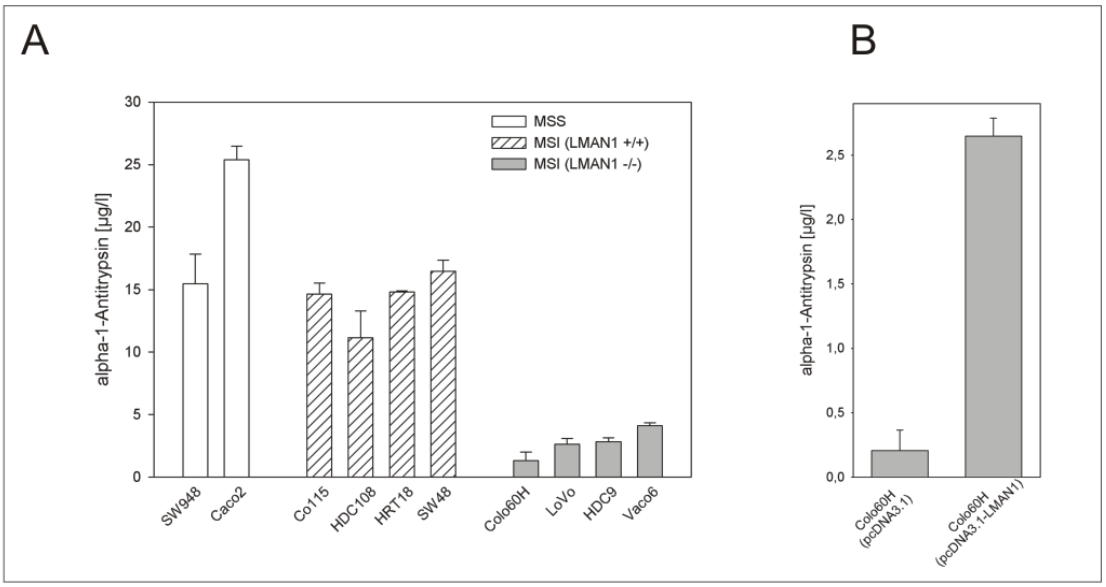


Figure 3.10 A1AT secretion in LMAN1-proficient and LMAN1-deficient cell lines. **A**. A1AT concentration in conditioned medium was measured by an A1AT specific ELISA. MSI-H cell lines harboring homozygous mutations for *LMAN1* (−/−) showed 3-fold lower levels of secreted A1AT than MSS and MSI-H cell lines with wild-type *LMAN1* cMNR alleles (+/+). The relative A1AT concentration was adjusted to the DNA amount in 80% confluent cells. Lactate dehydrogenase measurement excluded distortion by dead cells. Columns: mean of three independent experiments; bars: standard deviation. **B.** Reconstitution of A1AT secretion upon *LMAN1*-cDNA transfection into a *LMAN1*-deficient cell line. Colo60H cells were transiently transfected with an *LMAN1*-cDNA expression vector (pcDNA3.1-*LMAN1*) or empty control vector (pcDNA3.1). Forty-eight hours after transfection, there was a significant increase of A1AT concentration in conditioned medium compared with the same cell line transfected with the empty vector.

LMAN1-deficient cell lines showed a decreased A1AT secretion and this secretion defect could be abolished after transient expression of *LMAN1*. For A1AT a promotion of tumor progression was reported if less A1AT was secreted [Huang et al., 2004]. *LMAN1* mutated cells showed a decreased secretion of A1AT therefore suggesting a potential role of LMAN1-mediated A1AT secretion in MSI tumorigenesis.

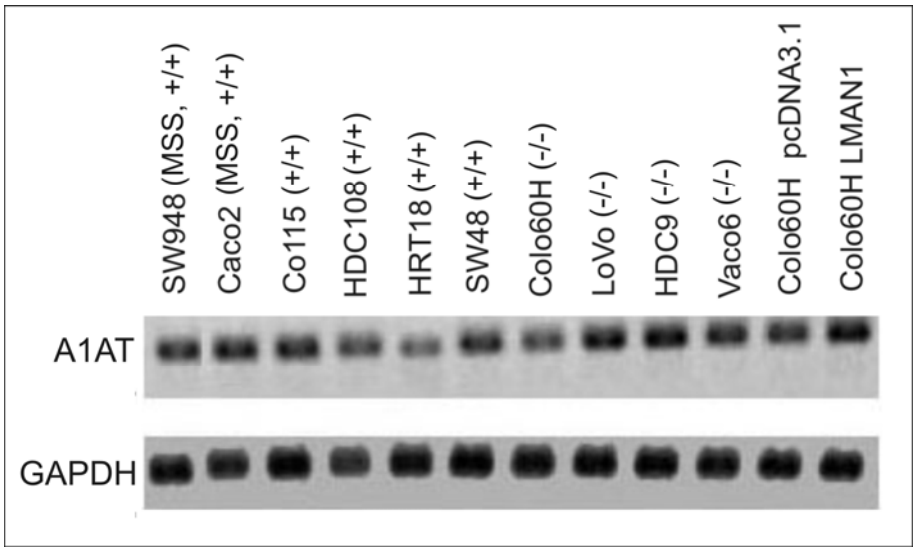


Figure 3.11 RT-PCR analysis of A1AT. RT-PCR analysis of *A1AT* expression in CRC cell lines used for A1AT ELISA. The expected *A1AT* PCR fragment of 484 bp could be amplified from cDNAs of 2 MSS cell lines (SW948 and Caco2), and MSI-H CRC cell lines mutated (−/−) and not mutated (+/+) for *LMAN1*, as well as transiently transfected Colo60H cells. As a control for cDNA, *GAPDH* was amplified.

3.4.2 Protein retention in the ER/Golgi fraction changes with *LMAN1* expression

A lack of LMAN1 transport function could affect several yet unknown glycoproteins. Since LMAN1 presence or absence changed the A1AT secretion it was determined whether *LMAN1* expression or loss of expression had effects on further glycoproteins which are transported through the secretory pathway. First indications of the overall protein pattern were determined by a pulse chase experiment with *LMAN1*-transfected LoVo cells (section 3.3.1). *LMAN1*-transfected LoVo cells and parental LoVo cells were pulsed with ^{35}S-methionine medium for 30 min and chased for 30 min. In 30 min pulse time the cells incorporated ^{35}S-methionine into current newly assembled proteins. During the 30 min chase period cargo proteins are expected to retain in the ER/Golgi fraction in *LMAN1*-deficient LoVo cells compared to *LMAN1* re-expressing cells. Subsequently, proteins with incorporated ^{35}S-methionine were detected and evaluated using the proteomweaver software. 164 protein spot matches were detected between parental LoVo cells and *LMAN1*-transfected LoVo cells.

Among these, five spots showed differences in the ER/Golgi fraction of parental LoVo cells compared to *LMAN1*-transfected LoVo cells (Fig. 3.12).

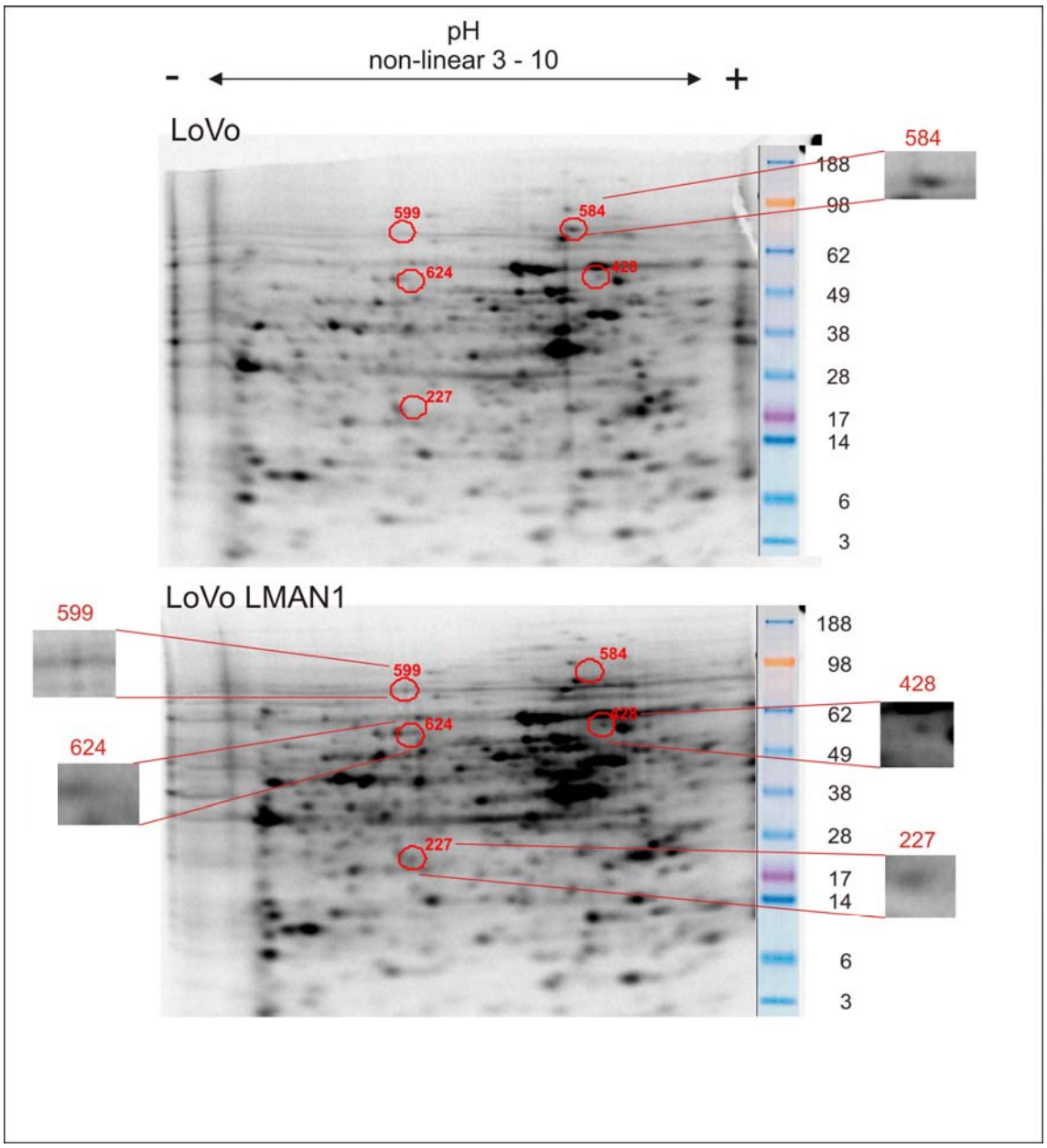


Figure 3.12 Pulse Chase Experiment with parental LoVo cells and LMAN1-transfected LoVo cells. Proteins of the ER/Golgi fraction were seperated by molecular weight and isoelectric point using 2D-gel electrophoresis and ^{35}S-methionine labelled protein spots were detected by a phosphoimager and subsequent analyzed by the Proteomweaver Sofware. The upper panel shows the ER/Golgi proteins that incorporated ^{35}S-methionine within the LoVo cells and the lower panel represents proteins with incorporated ^{35}S-methionine within *LMAN1*-transfected LoVo cells. Spot 584 disappeared in *LMAN1*-transfected LoVo cells, but was visible in parental LoVo cells all other spots appeared in *LMAN1*-transfected LoVo cells and were not

detectable in parental LoVo cells. Overall the protein assembly rate was higher in *LMAN1*-transfected LoVo cells. Detected spots are enlarged.

One spot was present in parental LoVo cells, but absent in the *LMAN1*-transfected LoVo cells (spot 584; Fig. 3.12) whereas four spots appeared only when *LMAN1* was re-expressed (spots 227, 428, 599 and 624; Fig 3.12). This data is in contrast to the expectation that cargo proteins are expected to retain in the ER/Golgi fraction. The apparent size and isoelectric point for each spot detected are summarized in table 3.2. Overall, the protein expression for some proteins appeared to be higher in *LMAN1*-transfected LoVo cells compared to parental LoVo cells, with more incorporated ^{35}S-methionine detectable.

Using an initial pulse-chase experiment we obtained preliminary evidence for potential differences in the protein pattern in a steady-state situation between parental LoVo cells and *LMAN1*-transfected LoVo cells.

Table 3.2 Spot analysis by the Proteomweaver Software

Spot number	Apparent size in kDa	Approximate isoelectric point	LMAN1 expression
599	60-90	6	+
624	40-60	6,5	+
227	20-30	6	+
428	50-80	8	+
584	60-90	7	−

Spots 599, 624, 227 and 428 appeared after *LMAN1* expression, whereas spot 584 disappeared after *LMAN1* expression. Spot 624 could correspond to LMAN1 itself with a molecular weight of 53 kDa and a isoelectric point of 6.7 and would serve as an internal control.

3.4.3 The LMAN1 carbohydrate recognition domain shows high affinity for LMAN1 substrates

Previous published studies have analyzed LMAN1 function by using different constructs, including or lacking all sub-structures [Appenzeller et al., 1999; Appenzeller-Herzog et al., 2004; Kamiya et al., 2008]. It is currently known that LMAN1 binding is dependent on pH and on Ca^{2+} concentration. Most studies used D-mannose to analyze LMAN1 binding properties. In this context it is important to know which conditions are required for binding to its specific substrates. In addition to

mannose it was decided to gain insight into LMAN1`s binding behavior using the complete cargo molecule A1AT. The aim was the purification of a functional LMAN1 protein with the carbohydrate recognition domain (CRD) as a powerful tool in search of specific yet unknown LMAN1 targets.

Two strategies were pursued to isolate functional LMAN1 protein: the purification of full length native protein expressed in mammalian cells and purification of bacterially expressed LMAN1-CRD protein followed by a refolding step. The CRD-domain including the oligomerization-domain was cloned into the pET44a(+) vector for bacterial expression and full length *LMAN1* cDNA was cloned into the mammalian expression vector pcDNA3.1 (fig. 6.1). After a successful functional binding to mannose sepharose the functional binding to A1AT sepharose was verified.

Full-length LMAN1 has a weak affinity to mannose sepharose

First, full length *LMAN1* was expressed in mammalian cells to simulate almost identical conditions to the *in vivo* situation. It was considered that in mammalian cells correct folding conditions and important interaction partners exist. Full length *LMAN1* cDNA was cloned into the mammalian expression vector pcDNA3.1 containing the SV40 ori and transfected into the human embryonic kidney cell line 293T, that contains the SV40 large T-antigen allowing the episomal replication of the plasmid for high-level gene expression. The *LMAN1*-pcDNA3.1 construct (Fig. 6.1) did not contain a pre-purification tag so that protein lysate was directly purified via a mannose sepharose column verifying functional binding of full length LMAN1. LMAN1 protein only showed weak binding to mannose sepharose column and was washed away through all wash fractions and could not be specifically eluted by mannose and EGTA (Fig. 3.13). Binding to A1AT-sepharose revealed similar results (Data not shown).

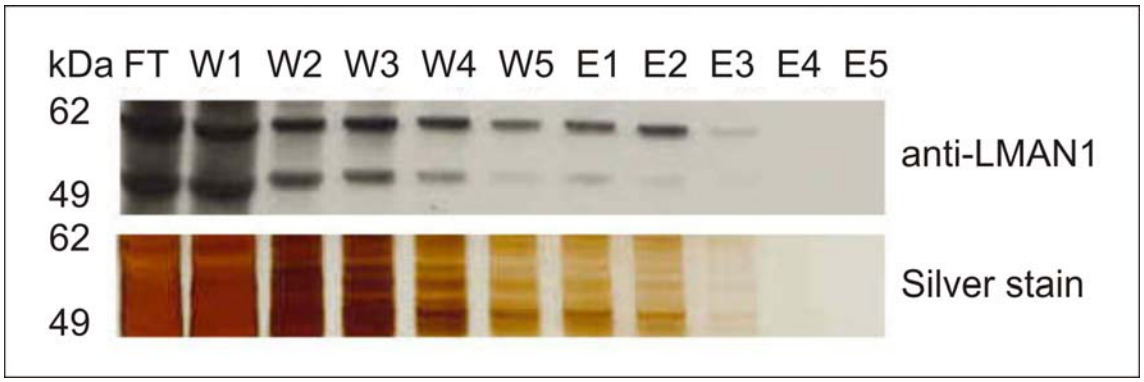


Figure 3.13 Full length LMAN1 purification using mannose-sepharose. Western blot analysis and silver staining of protein fractions for verification of LMAN1 specificity and purity. The expected size of full length LMAN1 is 53 kDa. Some LMAN1 protein bound to mannose sepharose, but was already washed away during washing steps and could not be specifically eluted with mannose and EGTA. Flow Through (FT), wash (W1-W5) and elution fractions (E1-E5). Marker bands are depicted.

LMAN1-CRD shows specific binding to selected substrates

LMAN1-CRD was cloned into the pET44(+)a vector and expressed in the bacterial strain BL21. After denaturing cell lysis, protein lysate was refolded by rapid dilution into 92 different buffer compositions (iFOLD System1). The binding capacity of refolded LMAN1-CRD protein was measured by a radioactive binding assay with iodinated alpha-1-antitrypsin (^{125}I-A1AT). The radioactive binding assay is described in section 6.3.2. For eight out of 92 buffers analyzed the refolding conditions showed high affinity of LMAN1-CRD to the substrate A1AT (Tab. 3.3). Buffer condition 2 was chosen, containing $CaCl_2$, which is an essential cofactor of LMAN1 for optimal binding to mannose-harboring glycoproteins. Large scale purification of LMAN1-CRD was performed under these buffer conditions using either a mannose sepharose column or an A1AT sepharose column. Protein purification was first performed using a mannose sepharose column. Highest amounts of protein were detected in flow through (FT) and wash fractions (W1 – W5), a slightly smaller protein was eluted in fraction E2, probably a partially degraded LMAN1-CRD protein. LMAN1-CRD showed a higher binding affinity to an A1AT sepharose column compared to the mannose sepharose column and was specifically eluted in fraction five (E5) with the expected size of 38 kDa (Fig. 3.14). The purified LMAN1-CRD showed a high binding capacity to A1AT sepharose column.

Table 3.3 Radioactive labelling using ^{125}I-A1AT

Buffer condition	Functional binding (DPM-decay per minute)	Control binding (DPM-decay per minute)
1	15.304	1.236
2	14.232	1.343
3	11.005	459
4	16.731	657
5	15.287	1.500
6	18.704	840
7	17.888	1.929
8	23.916	415

LMAN1-CRD was analysed using ^{125}I-A1AT and Ni^{2+}-sepharose. LMAN1-CRD was attached to Ni^{2+} via the N-terminal histidine tag of the pET44-CRD construct and incubated with ^{125}I-A1AT. Unbound ^{125}I-A1AT was washed away, followed by elution with 1M imidazole. Eluted material was analysed using a liquid scintillation counter (LSC) by measuring decays per minute (DPM). In the control labelling experiment labelled LMAN1-CRD was directly incubated with 1M imidazole so that imidazole served as a competitor to refolded LMAN1-CRD preventing histidine-tagged LMAN1-CRD to bind on the Ni^{2+}-sepharose. For buffer composition see table 6.1.

The two purification strategies to receive functional LMAN1 binding showed that LMAN1-CRD had a higher affinity to A1AT sepharose column compared to mannose sepharose column and that full length LMAN1 showed no appropriate binding. Taken together, our results demonstrate that LMAN1-CRD was successfully and specifically purified using the approach described above. In addition, the A1AT sepharose column was established as a highly specific tool for LMAN1 purification, superior to mannose sepharose columns.

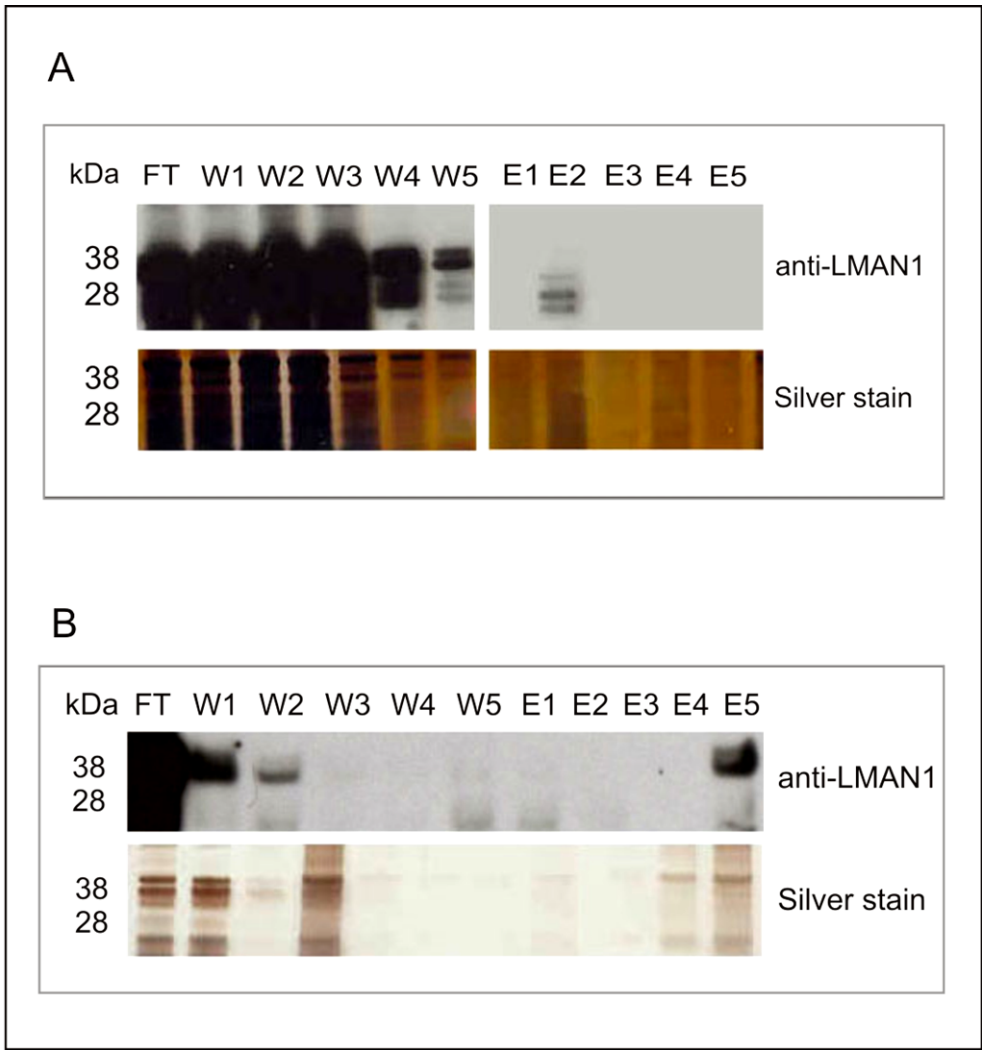


Figure 3.14 LMAN1-CRD purification using mannose- and A1AT-sepharose. Western Blot analysis and silver staining of protein fractions for verification of LMAN1 protein binding and purity, the predicted size of LMAN1-CRD is 38 kDa. **A.** LMAN1 CRD did not specifically bind to mannose (FT) and was washed away (W1-W5). A smaller band was detected in fraction E2, corresponding to degraded protein. **B.** LMAN1 CRD was able to bind to A1AT sepharose, was weakly washed away and specifically eluted with buffer containing mannose and EGTA (E5). Flow Through (FT), wash (W1-W5) and elution fractions (E1-E5). Marker bands are depicted.

3.5 XYLT2 loss of function mutation in MSI-H colorectal tumorigenesis

In addition to LMAN1 exhibiting transporter function, the enzyme, termed XYLT2 (xylosyltransferase 2), was highly mutated in MSI-H CRC cell lines (35%; 8/23). XYLT2 catalyzes the initial step in proteoglycan synthesis (Fig. 3.15). To gain a more detailed insight into the role of XYLT2 in MSI-H colorectal carcinogenesis the *XYLT2* gene was further analyzed by expression profiling and functional analysis using ^{3}H-xylose for an incorporation assay.

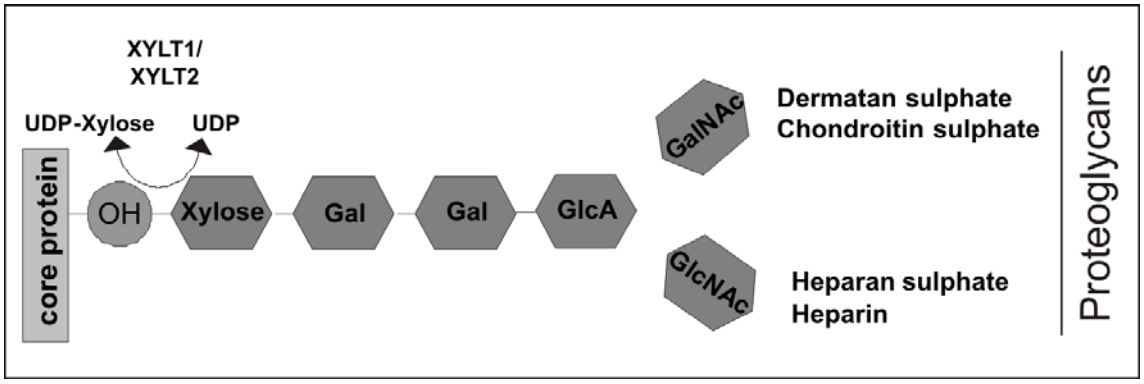


Figure 3.15 XYLT enzyme activity during proteoglycan synthesis [Götting et al., 2007]. The two XYLT enzymes catalyze the first step in proteoglycan synthesis by linking a serine hydroxyl group of the core protein with the first sugar xylose.

3.5.1 Variation of *XYLT2* and *XYLT1* expression within MSI-H CRC cell lines

Since mammalian cells can express two XYLT isoforms, expression analysis was performed for both gene isoforms, *XYLT1* and *XYLT2*. Oligonucleotides spanning at least two exons of the gene were designed and for the *XYLT2* gene the PCR product included the C7 cMNR. *XYLT* mRNA levels were investigated on 23 MSI-H CRC cells, one MSS cell line (SW948), normal human colon mucosa and the hepatocellular liver cancer cell line HepG2. For HepG2 cells a moderate *XYLT2* expression was reported [Roch et al., 2010]. *XYLT2* transcript was found to be expressed in all tested cells, including cell lines harboring mono- or biallelic *XYLT2* cMNR frameshift mutations and was also detected in normal colon tissue and HepG2

cells (Fig. 3.16A). *XYLT1* mRNA was also found to be expressed in all cell lines and normal colon tissue but not in the hepatocellular cell line HepG2 (Fig. 3.16B).

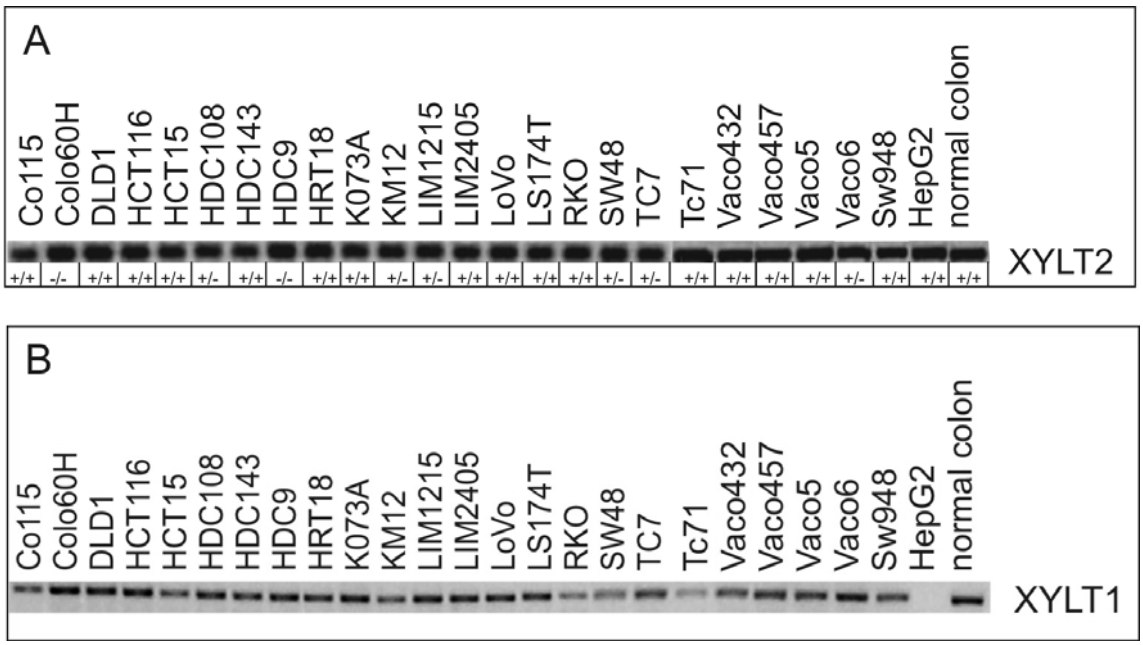


Figure 3.16 XYLT1 and XYLT2 expression in CRC cell lines and control cells. **A.** RT-PCR analysis of *XYLT2* expression in CRC cell lines with homozygously wild-type (+/+), homozygously mutant (−/−), or heterozygously mutant (+/−) *XYLT2* C7 cMNR. The expected *XYLT2* PCR fragment of 298 bp could be amplified from cDNA of normal human colon, a MSS cell line (SW948), and all MSI-H CRC cell lines. As a positive control for *XYLT2* expression HepG2 cells were used. **B.** RT-PCR analysis of *XYLT1* expression in CRC cell lines and HepG2 cells. The expected *XYLT1* PCR fragment of 250 bp could be amplified from cDNA of normal human colon, a MSS cell line (SW948), and all MSI-H CRC cell lines but not from the hepatocellular cancer cell line HepG2.

3.5.2 *XYLT2* reconstitution changes xylose incorporation in *XYLT2*-deficient MSI-H CRC cell lines

We hypothesized that *XYLT2*-deficient MSI-H CRC cell lines might have a lower xylose assembly rate compared to *XYLT2*-proficient cells, because XYLT2 catalyzes the first step in proteoglycan synthesis and links xylose to the core protein. For functional analysis of XYLT2 a ^{3}H-xylose incorporation assay was chosen. In case of a deficiency for XYLT no xylose assembly should occur (see Fig. 3.15). Due to the detection of *XYLT1* expression in CRC cells (Fig. 3.16B), a direct comparison of *XYLT2*-deficient and -proficient MSI-H CRC cell lines was not suitable, because

effects could not be attributed to XYLT2 alone. HepG2 cells only expressing *XYLT2* showed a high xylose assembly. MSI-H CRC cell lines harboring different *XYLT2* mutations, varied in their xylose assembly (Fig. 3.17) as was to be expected.

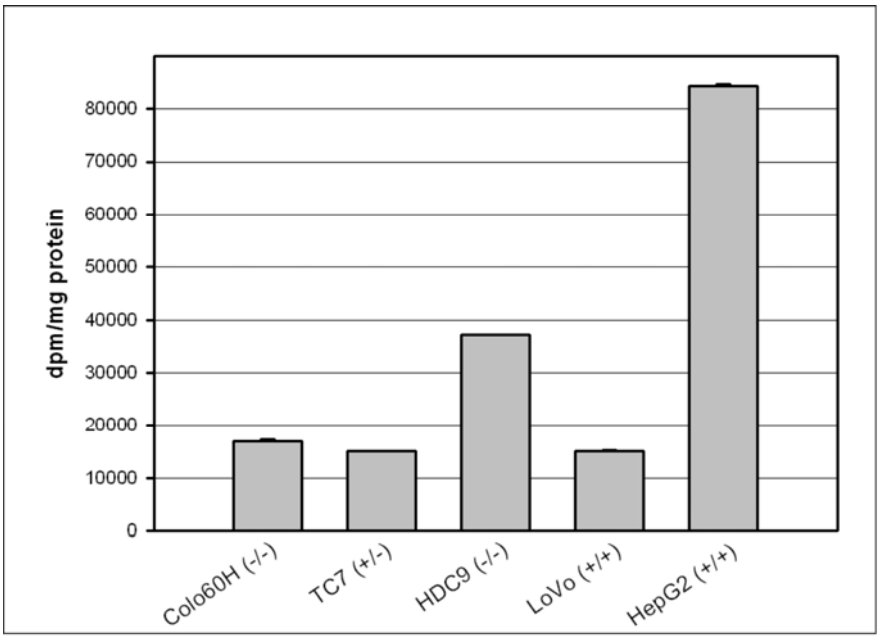


Figure 3.17 Metabolic ^{3}H-xylose labelling in MSI-H CRC cell lines. ^{3}H-Xylose assembly was performed using CRC cell lines with homozygous wild-type (+/+), homozygous mutant (−/−), or heterozygous (+/−) cMNR status for *XYLT2* and as a positive control the hepatocellular cell line HepG2 was used. Radioactive xylose amount was measured with a liquid scintillation counter in decays per minute (dpm) and protein amount was charged against dpm. HepG2 cells only expressing *XYLT2* showed a high xylose assembly. CRC cell lines harboring different *XYLT2* mutations, varied in their xylose assembly. Two independent experiments are shown. Error bars represent standard errors.

In addition to the effect that XYLT1 might take over the function of the XYLT2 protein, the genetic heterogeneity among the analyzed cancer cells account for the difference in xylose assembly rates. For that reason, one biallelically *XYLT2*-mutated MSI-H CRC cell line was transiently transfected with either *XYLT2*-cDNA or the empty vector control (pcDNA3.1). There was a minimal, but not significant increase, in xylose assembly in HDC9 cells upon *XYLT2* expression (Fig. 3.18) compared to mock-transfected HDC9 cells, which might result from low transfection efficiency. An involvement of XYLT2 in xylose assembly would probably be measurable if more

cells express *XYLT2*. Results require validation in stably transfected HDC9 cells expressing *XYLT2*, which are currently established in our laboratory.

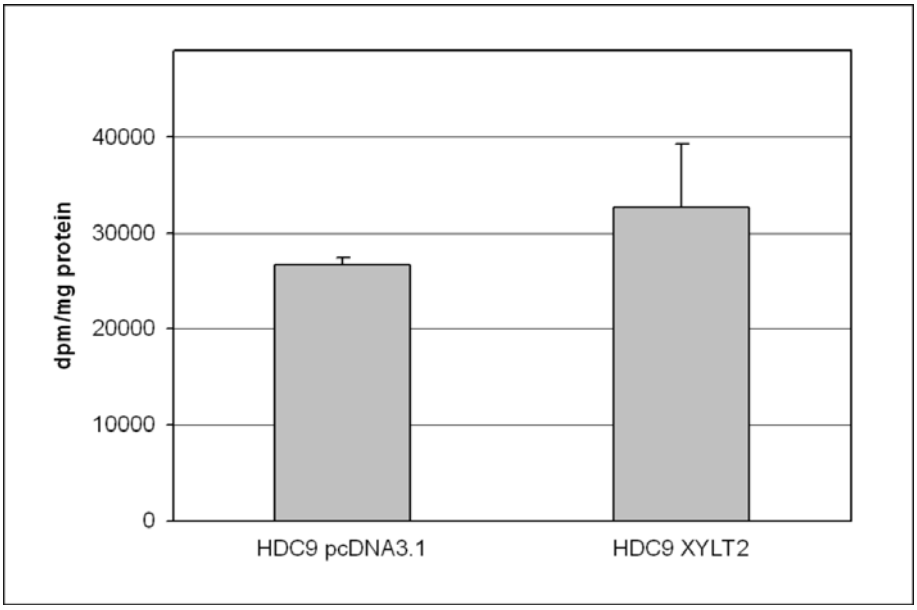


Figure 3.18 Metabolic ^{3}H-xylose labelling in HDC9 cells upon XYLT2 re-expression. Cells were transiently transfected with either XYLT2 pcDNA3.1 or the empty vector. Afterwards cells were fed with ^{3}H-xylose containing medium and harvested after 48 hours. Radioactive xylose amount was measured with a liquid scintillation counter in decays per minute (dpm) and protein amount was charted against dpm. Upon *XYLT2*-transfected HDC9 cells showed a minimal increase in xylose assembly compared to mock-transfected HDC9 cells. Results present two independent experiments. Error bars represent standard errors.

Expression analysis for *XYLT2* revealed similar results to *LMAN1* expression profiles. Biallelically *XYLT2*-mutated transcripts were still detectable although *XYLT2*-mutant transcript is predicted to be NMD sensitive with a PTC more than 55 base pairs upstream of the last exon-exon border. A changed enzyme function with regard to XYLT2 loss of function could not be confirmed yet.

4. Discussion

4.1 Glycogenes show high mutation frequencies in MSI-H CRC

The major goals of this work were the identification and characterization of coding mononucleotide repeat (cMNR) harboring genes encoding proteins of the glycosylation machinery, and the characterization of the resulting effects on cell surface glycosylation. Database analysis revealed 28 cMNR-harboring genes encoding components of the cellular glycosylation machinery. The three glycosylation genes which were already brought into correlation with MSI CRC, *B4GALT1*, *GALNT5* and *SIAT4B* [Kim et al., 2004], were not selected by this database preselection, because all three coding regions did not possess a cMNR of more than seven base pairs. Here, only new genes with an involvement in glycosylation were identified.

Two genes involved in the glycosylation machinery, *LMAN1* and *XYLT2,* showed mutation frequencies above the predicted frequency for the length of their cMNRs in MSI-H CRC cell lines, colorectal adenomas and carcinomas, suggesting a functional role as mutational targets in MSI tumorigenesis. In addition to the most frequently mutated genes, candidate genes mutated at lower frequency might however also be relevant for MSI tumor development.

4.2 Genetic alterations in LMAN1 might play a role in MSI colorectal tumorigenesis

4.2.1 LMAN1 transporter function in the cell

LMAN1 (also known as ERGIC-53: ER/Golgi intermediate compartment 53 kDa) is a membrane protein that carries secretory proteins in COPII vesicles from the ER to the Golgi apparatus [Stephens et al., 2001]. LMAN1 binds as a mannose-specific lectin to its substrates by a luminal carbohydrate recognition domain (CRD) in the neutral and Ca^{2+}-rich ER and releases its cargo in the ERGIC because of a more acidic environment and a lower Ca^{2+} concentration [Hauri et al., 2000]. LMAN1 acts as a transport receptor for the serine protease inhibitor A1AT [Nyfeler et al., 2008] and is involved in the ER-export of the lysosomal glycoproteins cathepsin C and Z as well as the blood coagulation factors V and VIII [Appenzeller et al., 1999; Moussalli et al., 1999; Nyfeler et al., 2008; Vollenweider et al., 1998]. Genetic alterations in *LMAN1* are known to be associated with a human disease termed combined deficiency of coagulation factors V and VIII (F5F8D), but connections to cancer have not been reported so far [Spreafico et al., 2008].

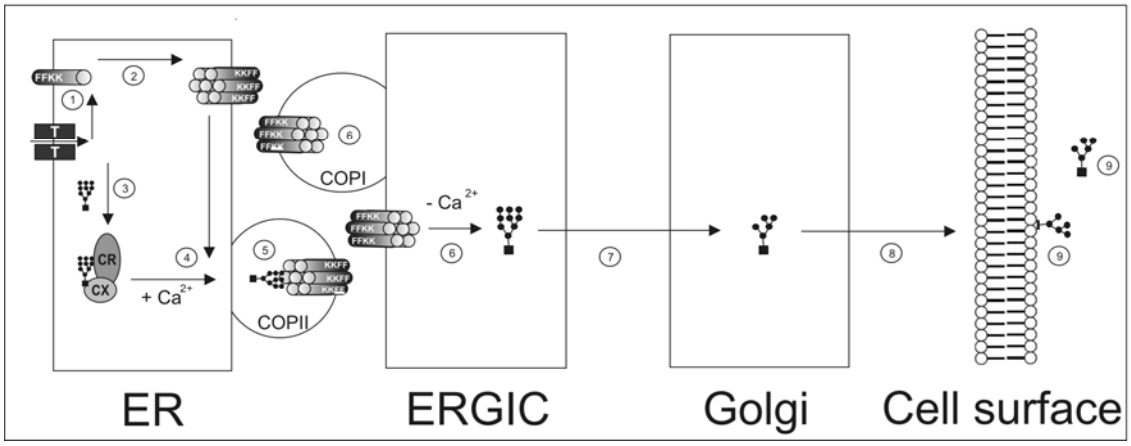


Figure 4.1 LMAN1 transporter function in the cell [Hauri et al., 2000]. LMAN1 is a Type I mannose-binding lectin, which binds to mannose (Man9) bearing glycoproteins in the ER after these glycoproteins have undergone folding and quality control (3). LMAN1 monomers are translocated into the ER-lumen (1) and form disulfide-linked homo-hexamers (2). LMAN1 hexamers bind to Man9-N-glycans of correctly folded cargo glycoproteins in the presence of Ca^{2+} (4). It cycles between ER and ERGIC, transporting glycoproteins to the ERGIC (5). In the ERGIC glycoproteins are released due to a more acidic, Ca^{2+}-low environment (6). The

glycoprotein is further transported through the secretory pathway (8) and either secreted or is associated with the cell surface (9). The C-terminal FF-motif of LMAN1 is required as the ER exit signal and free LMAN1 is recruited back to the ER via direct interaction of di-lysine (KK) with CopI vesicles.

Mutations of *LMAN1* associated with this autosomal recessive bleeding disorder (F5F8D) were found throughout the whole gene sequence, including all exons except exon 4. Nonsense, frameshift, splicing defect or missens mutations were involved. In contrast, *LMAN1* mutations that occurred in MSI-H tumors represent cMNR frameshift mutations only in exon 8 and interfere with normal LMAN1 function probably by disrupting the control region of the oligomerization domain. Two conserved cysteine residues as well as four α-helices located in the oligomerization domain are required for the formation of homohexamers [Neve et al., 2005; Nufer et al., 2003]. It was reported that LMAN1 only exits the ER in a hexameric form suggesting that only this form is functional for transporting cargo proteins [Neve et al., 2005]. Frameshift mutations in the *LMAN1* cMNR would lead to a loss of both cysteine residues and two α-helices, thus preventing oligomerization of LMAN1. The lack of oligomerization leads to misfolding and to retention in the ER by chaperones [Hurtley et al., 1989], so that a transport of cargo proteins would not be able anymore. This work provides the first evidence for LMAN1 being involved in human neoplastic disease apart from F5F8D. Several lines of evidence support the involvement of *LMAN1* in MSI tumorigenesis.

4.2.2 LMAN1 involvement in MSI-H tumorigenesis

First, a high mutation frequency of 52% was observed in MSI-H CRC cell lines with biallelic mutations in 17% of the cell lines. The *LMAN1* mutation frequency was significantly higher than the predicted frequency for the length of its cMNR of nine nucleotides (A9). Accordingly, *LMAN1* mutations most likely are subject to positive selection pressure and hence might confer a growth advantage to affected MSI tumor cells [Woerner et al., 2010]. Second, a significant fraction of LMAN1 mutations in MSI-H CRC cell lines and primary tumors affected both alleles. In general, such biallelic mutational inactivation usually points to a tumor suppressor function of the

affected gene. To understand the nature of the genetic defects observed in MSI-H CRC cells, *LMAN1* expression was analyzed at the transcript and protein level. Expression analysis showed that *LMAN1* mRNA was expressed in all MSI-H CRC cell lines, including those with mono- and biallelic mutations. *LMAN1*-mutant transcripts harbor a premature termination codon, thus being predicted to be NMD-sensitive. Messenger RNA transcripts that carry premature termination codons (PTCs) are recognized by the nonsense-mediated mRNA decay (NMD) system when the PTC appears 50 – 55 base pairs upstream of the last exon-exon border [El-Bchiri et al., 2005; El-Bchiri et al., 2008; Ionov et al., 2004]. NMD-sensitive transcripts should be recognized and degraded by the nonsense-mediated mRNA decay (NMD) system [You et al., 2007]. However, *LMAN1* transcripts were easily detectable by RT-PCR analysis in cell lines with biallelic frameshift mutations and hence seemed NMD resistant. Altered mRNA sequences that are not affected by NMD can lead to the generation of truncated proteins, carrying a carboxy-terminal neopeptide sequence due to shifts in the reading frame of the encoded amino acid sequence. These neopeptides appeared to be highly immunogenic [Linnebacher et al., 2001; Schwitalle et al., 2008], thus possibly explaining the high lymphocytic infiltration observed in MSI CRCs. To observe whether a truncated LMAN1 protein is expressed in *LMAN1*-mutated cell lines, Western blot analysis with an appropriate LMAN1 antibody detecting an amino-terminal peptide was performed. Cell lines harboring wildtype *LMAN1* or heterozygously mutated *LMAN1* showed a protein band at the expected size of 53 kDa, but no truncated protein band of 36 kDa could be detected in heterozygous or homozygous *LMAN1* cell lines. This suggests that mutant *LMAN1* transcript could represent a NMD-escape transcript, affected by nonsense-mediated translational repression [You et al., 2007] or the truncated *LMAN1*-mutant protein is recognized by ER-associated degradation (ERAD) enzymes and efficiently degraded by proteasomal degradation.

Immunohistochemical analyses on a set of 50 MSI-H colorectal carcinomas showed local loss in 38% of all carcinomas and complete loss of LMAN1 expression in 6% of all carcinomas. In addition biallelic mutations, associated with complete loss of LMAN1 protein expression, occurred frequently in MSI-H colorectal carcinomas and adenomas. This was proven by regional microdissection on a few MSI-H carcinomas and MSI-H adenomas. DNA frameshift mutation analysis revealed

biallelic cMNR frameshift mutations in the *LMAN1* gene and loss of LMAN1 protein expression in these tumor areas. Interestingly, some LMAN1-stained tumor slides gave the impression that *LMAN1*-deficient tumor regions appeared to show a higher invasiveness compared to areas expressing *LMAN1*. But due to the low number of tumors that showed this effect no final conclusion could be drawn. In order to address this issue more tumor slides should be analyzed and the metastatic behavior of *LMAN1*-deficient cell lines could be examined. Furthermore, *LMAN1* mutations were observed through all tumor stages (UICC stages, data not shown) analyzed and occurred already in adenomas. Interestingly, *LMAN1* mutations seemed to occur more frequently in hereditary MSI-H CRCs (HNPCC) than in sporadic MSI-H CRCs suggesting that *LMAN1* mutations appear in patients with germline mutations in the MMR-system and are less frequent observed after epigenetic silencing of the MLH1 promotor accompanied by *BRAF* mutations in sporadic MSI tumors.

Impaired A1AT secretion in LMAN1-deficient MSI-H cancer cell lines

The third and most striking evidence for an involvement of LMAN1 in MSI tumorigenesis relates to the observation that *LMAN1*-deficient cell lines showed impaired A1AT secretion. Alpha-1-antitrypsin (A1AT), a serine protease inhibitor, was reported to be involved in tumorigenesis [Huang et al., 2004]. Its expression is found in many different tissues including the human colon [Carlson et al., 1988; Geboes et al., 1983]. A1AT exerts anti-proteolytic function and confers angio-inhibitory activity [Huang et al., 2004]. Comparative cDNA microarray expression analyses of human normal and tumor tissues revealed that lower levels of *A1AT* transcripts correlated with larger tumor size [Huang et al., 2004] and higher levels of non-circulating A1AT within tumors tend to be associated with a better prognosis [Allgayer et al., 1998]. Interestingly, it was hypothesized that A1AT may be protective against MSI-H-CRC development, because increased A1AT levels in postmenopausal women receiving estrogen were associated with a reduced incidence for MSI-H CRC [Yang et al., 2000]. Work by Huang and colleagues also shows significant variations in local A1AT levels in human tumor tissues [Huang et al., 2004]. The observation of intratumoral heterogeneity of local LMAN1 deficiency might provide a molecular explanation for this finding. The A1AT-ELISA results showed significantly lower levels of A1AT secreted into the cell culture medium in *LMAN1*-deficient cell lines, reaching about

20% of the level secreted by the *LMAN1*-proficient control cell line Caco2. These data on A1AT secretion correspond well with the results of Nyfeler and colleagues on mouse embryonic fibroblasts (MEF) derived from LMAN1 (−/−) knockout mice [Nyfeler et al., 2008]. LMAN1-deficient MEFs secreted only approximately 25% of A1AT in a given time period compared with LMAN1 (+/+) MEFs. Furthermore, the re-expression of *LMAN1* in a *LMAN1*-deficient cell line abolished the secretion defect, thus the A1AT secretion decrease could be attributed to the loss of LMAN1 transporter function.

A1AT has an important role in inhibiting several serine proteases with essential roles in processes like blood coagulation, apoptosis, cell differentiation, ECM degradation and inflammation. One prominent protease which is inhbited by A1AT is the enzyme neutrophil elastase (NE). NE is expressed in neutrophils which are the first immune cells arriving at a site of inflammation [Heutinck et al., 2010]. In addition to their function in promoting bacterial killing, NEs are involved in degrading ECM compounds and in migration [Korkmaz et al., 2008; Rao et al., 2004]. NE was reported to hydrolyze ECM compounds such as collagens, fibronectins and proteoglycans [Mainardi et al., 1980; McDonald et al., 1980; McGowan et al., 1989]. Breast tumors with high levels of free NE were associated with a poor prognosis [Akizuki et al., 2007]. A loss of A1AT as the inhibitor of NE could contribute to that effects. Proteases such as NE may play a pathological role in facilitating cancer cell invasion and metastasis by dissolution of the tumor matrix. Other proteases like proprotein convertases (PCs) were reported to play an important role in tumor cell metastasis in human colorectal tumor cells. Overexpression of another serine protease inhibitor termed α-1-antitrypsin Portland (α1-PDX) significantly reduced the ability of colorectal tumor cells to form liver metastasis [Scamuffa et al., 2008]. Thus inhibition of proteases deregulating the environment of tumor cells is an important task and loss of this function generates an advantage for the tumor cell. If α1-PDX presents also a cargo protein of LMAN1 has to be proven.

With the A1AT secretion defects in *LMAN1*-deficient cell lines and the restoration of this defect by transient *LMAN1*-transfection a direct correlation between A1AT and LMAN1 as a specific transporter of A1AT could be made. The loss of LMAN1 transporter function might decrease A1AT levels in the tumor tissue

thus driving tumor progression or even lead the formation of metastasis by enhanced levels of proteases changing the ECM of the tumor [Roeckel et al., 2009].

4.2.3 The screening for new LMAN1 cargo proteins

Apart from A1AT secretion defects, alternative mechanisms could also contribute to a tumor-promoting effect of *LMAN1* mutations. *LMAN1* mutations might confer more indirect effects via impaired transport of known or yet unknown cargo proteins. For example, *LMAN1* deficiency might promote MSI colorectal tumorigenesis by affecting the secretion of client proteins such as cathepsin C and Z or by interfering with interacting proteins such as MCFD2, FGFR3 and SUMF1 [Fraldi et al., 2008; Lievens et al., 2008; Nyfeler et al., 2006]. Moreover, it is reasonable to assume that additional yet unknown LMAN1 cargo proteins exist, which might be affected by a loss of LMAN1 function.

To identify new LMAN1 cargo and binding proteins, two techniques were established: (i) a pulse-chase experiment to identify the general distribution of proteins in the ER of *LMAN1*-proficient *versus LMAN1*-deficient MSI-H tumor cells and (ii) the use of purified LMAN1 protein as a bait to capture interacting proteins. With the pulse chase experiment the retention of proteins upon *LMAN1* re-expression in a *LMAN1*-deficient background was observed using ^{35}S-methionine. By comparison of isoelectric points and molecular weights already known cargo proteins were excluded. With the functional purification of the LMAN1 protein an important tool for the identification of new LMAN1 cargo proteins was generated. An agarose-immobilized LMAN1 protein should subsequently be used for the fishing of new LMAN1 cargo proteins in the medium of MSI-H colorectal cancer cells.

Protein retention upon LMAN1 expression

The aim of the pulse chase experiment was to identify new LMAN1 cargo proteins retained in the ER/Golgi fraction in *LMAN1*-deficient LoVo cells due to a loss of LMAN1 transporter function. 164 proteins with incorporated ^{35}S-methionine were detected in the ER/Golgi fraction. Only five protein spots showed differences in *LMAN1*-deficient LoVo cells compared to *LMAN1*-transfected LoVo cells. A possible reason for the low number of differences detected, might be the cells`s ability to compensate some, but not all, of the effects caused by *LMAN1* deficiency by utilizing other transport pathways. It is known that LMAN1 can interact with additional proteins representing a checkpoint for quality control of glycoproteins [Fraldi et al., 2008;

Lievens et al., 2008; Nyfeler et al., 2006]. The additional function of LMAN1 as a checkpoint together with its interaction partners could also work with other lectins, suggesting alternative pathways playing similar roles like LMAN1 complexes and taking over their function. An alternative explanation might be that retained proteins were detected by a certain quality control system in *LMAN1*-deficient cells and immediately degraded, because normal transport conditions were interrupted.

However another explanation appears most likely. If one recalls that approximately 3000 – 4000 proteins are secreted and just as many pass the secretory pathway as membrane bound proteins [Bonin-Debs et al., 2004], it is obvious that the employed experimental strategy most likely detected only the most abundant proteins. Although only one potential cargo protein was detected in *LMAN1*-deficient cells, other LMAN1 cargo proteins might be present which were below the detection limit of this method. If LMAN1 cargo proteins remain in the ER/Golgi fraction but can not be detected because they are masked by high abundance proteins, variation and optimization of the kinetic parameters would be required. In particular, short pulse times in the range of minutes and extended chase times between two and 24 hours should be considered, thus removing the label from all proteins transported through the secretory pathway. Accordingly, retention of low abundant proteins in the ER/Golgi due to *LMAN1* deficiency should be increased. To further reduce the sample complexity and enrich potential low abundant LMAN1 cargo proteins, pre-fractionations could be applied before 2D-electrophoresis e.g. by liquid phase isoelectric focussing or chromatofocussing. Since LMAN1 is a mannose-specific lectin, an even more specific enrichment of possible cargo proteins would be affinity chromatography via mannose-specific lectins.

The pulse chase experiment in this study provides a first clue about additional LMAN1 cargo proteins and hence also might contribute to understand the role of LMAN1 in MSI tumorigenesis. For the identification of new LMAN1 cargo proteins an alternative labelling method with the incorporation of ^{13}C-carbon-labelled amino acids could be considered. By this labelling method a mass spectrometric distinction between proteins with incorporated ^{13}C and non-labelled proteins is possible. However, the identification of ^{13}C-labelled proteins in stained 2D-gel spots will be hindered by the fact that not only the ^{13}C-labelled proteins but all proteins including all ER/Golgi-resident will appear as prominent spots on the gel. In order to address

this issue another pre-fractionation step after ER/Golgi fractionation might be sufficient e.g. a subsequent fishing with the LMAN1-CRD.

The LMAN1 carbohydrate recognition domain could provide a powerful tool for cargo protein detection

An alternative experimental strategy was to apply LMAN1`s lectin function for the identification of new LMAN1 cargo proteins. To this end functional *LMAN1* constructs were required.

For functional analysis two different *LMAN1* constructs were applied, the full length LMAN1 protein and its carbohydrate recognition (CRD)-domain. The binding of these constructs to two affinity matrices, mannose sepharose and A1AT sepharose revealed that the *LMAN1*-CRD construct had the highest binding capacity to an A1AT sepharose column. The binding of full length LMAN1 to the mannose sepharose column was less efficient. The results of full length LMAN1 were not in line with LMAN1 binding assays of Appenzeller-Herzog and colleagues [Appenzeller-Herzog et al., 2004], although cell lysis and purification steps were performed identically. Appenzeller-Herzog purified the full length Myc-tagged LMAN1 via a mannose sepharose column and detected the LMAN1 construct indirectly by using a Myc-antibody. One possible explanation for the inefficient binding of full length LMAN1 in the experiment established in this work, might be insufficient folding upon synthesis in 293T cells. Unphysiologically high amounts of LMAN1 protein synthesized in these cells might lead to the formation of aggregates and thus incorrect folding of monomeric LMAN1 polypeptides. Similarly, impaired multimerization could contribute to inefficient binding, since for optimal LMAN1 function hexamer formation is essential [Neve et al., 2005]. Alternatively, a co-factor or interaction partner of LMAN1, required for efficient binding, was missing from or was not expressed in 293T so that LMAN1 was not able to bind tightly to the mannose sepharose column.

LMAN1-CRD was purified by a denaturing inclusion body lysis step and subsequent refolding was established using the iFOLD™ protein refolding system 1. The *LMAN1*-CRD construct lacked the transmembrane and cytosolic domains, so that bacterial expression and protein refolding as well as purification steps were easier to handle. In addition, it is known that only the oligomerization domain is

sufficient for a functional hexamerization of LMAN1 protein [Neve et al., 2005]. In complete, 92 buffer conditions were tested and correct LMAN1-CRD folding was controlled using a radioactive binding assay with ^{125}I-A1AT. LMAN1-CRD binding to ^{125}I-A1AT was efficient for eight buffers. Buffer condition 2 was chosen since this condition showed a low dpm-value for the control and besides this buffer contained the cofactor Ca^{2+} which is essential for optimal LMAN1 binding [Zhang et al., 2009]. The *LMAN1*-CRD construct lacking the transmembrane as well as the cytosolic domain showed specific binding to the A1AT sepharose column, but binding to the mannose sepharose column was not sufficient.

The low affinity of LMAN1 to mannose might be explained by missing cofactors or modifications such as phosphorylation important for LMAN1 function. It is known that LMAN1 can interact with additional proteins representing a checkpoint for quality control of glycoproteins [Fraldi et al., 2008; Lievens et al., 2008; Nyfeler et al., 2006]. Under certain circumstances LMAN1 needs interaction partners for the correct transport of cargo proteins as it was shown for the coagulation factors V and VIII [Nyfeler et al., 2006], where the MCFD2-LMAN1 complex forms a specific cargo receptor for the ER-to-Golgi transport of these proteins. Furthermore LMAN1 was reported to function as a checkpoint for FGFR3 maturation, allowing only functional receptors to reach the cell surface [Lievens et al., 2008]. There are two further publications where an LMAN1 checkpoint function in quality control was demonstrated [Anelli et al., 2007; Fraldi et al., 2008]. In both cases LMAN1 interacts with another protein, termed Erp44.

Although it was reported that transmebrane and cytosolic domains are not sufficient for the functional binding of LMAN1, these domains might play a role in correct conformation of LMAN1 in addition to the efficient cofactors.

Furthermore LMAN1 might only detect mannose in a specific conformation order on the cargo protein. This would explain why LMAN1 binding to A1AT sepharose column was more successfull. For A1AT binding best results were obtained at a pH of 7.4 at 4°C and Ca^{2+} concentrations of 1 mM. LMAN1-CRD protein was specifically eluted with mannose and EGTA, but also high protein amounts were found in the flow through fraction.

In the present study, optimized conditions for a functional purification of LMAN1-CRD have been established and can now be directly applied to large scale

functional LMAN1 purification. Future work will concentrate on the establishment of experiments for the fishing of LMAN1 substrates in cell extracts and conditioned medium with an agarose-immobilized LMAN1-CRD.

The LMAN1 pulse chase experiment as well as the purification of LMAN1-CRD revealed on one hand potential new LMAN1 cargo proteins and on the other hand an appropriate tool for the specific identification of new cargo proteins. Additional cargo proteins could explain the role of LMAN1 in MSI tumorigenesis.

4.2.4 *LMAN1* re-expression changes the cell surface glycoprotein pattern

Although only secreted proteins are known as LMAN1 client proteins, one might speculate that membrane-bound glycoproteins may also be transported by this mechanism. Support for this hypothesis was gained from Lectin-FACS analysis that showed changes in the cell surface glycoprotein pattern occuring in *LMAN1*-deficient LoVo cells upon reconstituted *LMAN1* expression. In particular, the presence of core substitutions and branching in N-glycans and the sialylation status of N- and O-glycans were profiled with a lectin panel. Mock-transfected *LMAN1*-deficient LoVo cells showed a higher binding of PHA-L lectins to the cell surface compared to *LMAN1*-transfected LoVo cells, whereas jacalin (JAC) lectin binding was reduced.

PHA-L is an N-glycosylation-specific lectin able to agglutinate lymphocytes in mitogenic active cells and has been termed the “L” subunit, for leuco-agglutinin instead of the PHA-E lectin agglutinating erythrocytes [Leavitt et al., 1977]. As mentioned in section 2.5.1, PHA-L is as a marker for N-glycans bearing the β1-6 branch. The synthesis of the β1-6 branch is catalyzed by the enzyme GlcNAcT-V (MGAT5), which was shown to be hyperactive in the progression of cancer from a tumorigenic to a metastatic phenotype [Demetriou et al., 1995; Dennis et al., 1999; Przybylo et al., 2008]. LMAN1 loss might influence tumor progression, but how LMAN1 loss can contribute to a higher β1-6 branching is difficult to reconcile. The disturbance of the branching relation on the cell surface might in general influence the balance. Increased β1-6 branching could lead to a decrease of other branches for cell-cell communication and thus influene cell behavior. Perhaps LMAN1 loss leads to an uncontrolled transport of overexpressed *MGAT5* resulting in upregulated β1-6

branching on the cell surface. Seelentag and colleauges observed an increase in β1-6 branches in metastatic colon carcinoma cells [Seelentag et al., 1998]. For *LMAN1*-deficient MSI-H CRC cells a test for metastatic behavior, as well as the overexpression of *MGAT5* should be performed to understand the correlation of *LMAN1* deficiency and increased β1-6 branching. In addition, it is well documented that effects of β1-6 branches in glycoproteins may trigger cell growth, motility, cell adhesion and differentiation [Mendelsohn et al., 2007]. These N-glycan branches generate ligands for lattice-forming lectins, like galectins, that regulate surface levels of glycoproteins including epidermal growth factor (EGF) and transforming growth factor-beta (TGFß) receptors [Lau et al., 2007]. Influences on signalling in *LMAN1*-deficient cells were not addressed yet, but warrant further investigation.

JAC is an O-glycoside-specific lectin binding to oligosaccharide structures containing GalNAc, also called Tn-antigens. JAC will also bind to these structures in a mono- or di-sialylated form thus allowing an earlier recognition of changed O-glycan structures. Surprisingly, PNA another lectin detecting higher sialylated Tn-antigens failed to show a differential staining pattern to the cell surface upon *LMAN1* re-expression. The Tn-antigen was reported to be present on most human cancers [Springer, 1984]. Mock-transfected LoVo cells bound less JAC lectins compared to *LMAN1*-transfected LoVo cells, thus *LMAN1*-deficient cells did not possess a great portion of Tn-antigens on the cell surface. As described in section 2.5.2 Tn-antigens are incompletely glycosylated mucins containing less total carbohydrate, and frequent sialylated core 1 structures. Possibly enzymes necessary for the elongation of O-glycans were no longer transported by LMAN1, thus leaving the O-glycan incompletely glycosylated. Conversely, glycoproteins transported by LMAN1 can be O-glycosylated in the Golgi apparatus in addition to their N-glycosylation on serine or threonine residues of the same protein, so that the higher binding of JAC to the cell surface of *LMAN1*-transfected cells could be explained.

PSA and WGA showed also less binding to glycan structures on the cell surface of *LMAN1*-deficient LoVo cells compared to mock-transfected LoVo cells. PSA has specificity towards α-linked mannose-containing oligosaccharides with core fucose and WGA shows specificity towards N-acetylglucosamine structures [Monsigny et al., 1980; Tateno et al., 2009]. Core fucosylation of N-glycans is a common modification on membrane proteins in humans. Fucosylated N-glycans

exhibit a smaller degree of conformational flexibility and lower flexibility increases the lifespan of some serum proteins. Fucosylated N-glycans are involved in a variety of biological and pathological processes [Ma et al., 2006].

Overall, loss of *LMAN1* expression led to a cell surface glycoprotein pattern different from the pattern observed on LoVo cells re-expressing *LMAN1*. More PHA-L lectins bound on the cell surface suggested more β1-6 branching in *LMAN1*-deficient cells, a glycosylation pattern known to be involved in the progression of cancer from a tumorigenic to a metastatic phenotype [Demetriou et al., 1995; Dennis et al., 1999; Przybylo et al., 2008]. *LMAN1*-deficient MSI-H CRC cells will be tested for an involvement in metastatic behavior in future experiments.

4.3 Consequences of *XYLT2* mutations in MSI-H colorectal tumors

XYLT2 catalyzes the transfer of xylose from UDP-xylose to selected serine residues of a core protein representing the initial enzyme in proteoglycan biosynthesis (Fig. 4.2). Proteoglycans are present in the extracellular matrix and on the cell surface, playing a key role in cellular signalling, as storage depots for growth factors and cytokines, and interaction with a broad variety of counterpart molecules [Iozzo, 1998]. Two xylosyltranferase isoforms (*XYLT1* and *XYLT2*) are found that are differentially expressed in cell types and tissues [Götting et al., 2007] They serve as biochemical markers in systemic sclerosis [Götting et al., 1999] and are associated with several other diseases like Pseudoxanthoma elasticum, diabetic nephropathy and osteoarthritis [Götting et al., 2007]. Changed expression in glycosyltransferases like *XYLT2* will probably change the glycosylation pattern on the cell surface. Cell surface changes will possibly affect cell-cell interaction or cell-matrix interactions. In general, the glycan profile of tumor cells shows quantitative differences with increased and decreased expression compared to normal cells [Gabius H-J, 2009]. Mucin-type O-glycans, Lewis antigens and Tn-antigens are already used as tumor markers for different tumor entities, like colorectal cancer, breast cancer, pancreatic and prostate cancer [Brockhausen, 2006]. This study reported for the first time *XYLT2* mutations in MSI-H colorectal tumors. The *XYLT2* C7 cMNR showed the second highest mutation frequencies with 35% on MSI-H CRC cell lines, 26% on MSI-H colorectal carcinomas and 21% on MSI-H colorectal adenomas. Furthermore, biallelic *XYLT2* mutations were observed in two cell lines and *XYLT2* was predicted as a positive selected gene, thus suggesting *XYLT2* as a tumor-promoting gene with a presumably tumor suppressor function. Expression analysis of both *XYLT* genes showed that *XYLT1* transcript was detected in all MSI-H CRC cell lines, as well as the control cell lines SW948 and normal colon mucosa but not in the XYLT2 positive control cell line HepG2. *XYLT1* presents the second isoform of *XYLT* genes in mammals, but without a cMNR in its coding region. *XYLT2* mRNA was transcribed in all MSI-H CRC cell lines, including HDC9 and Colo60H cells, that are biallelically mutated for *XYLT2*.

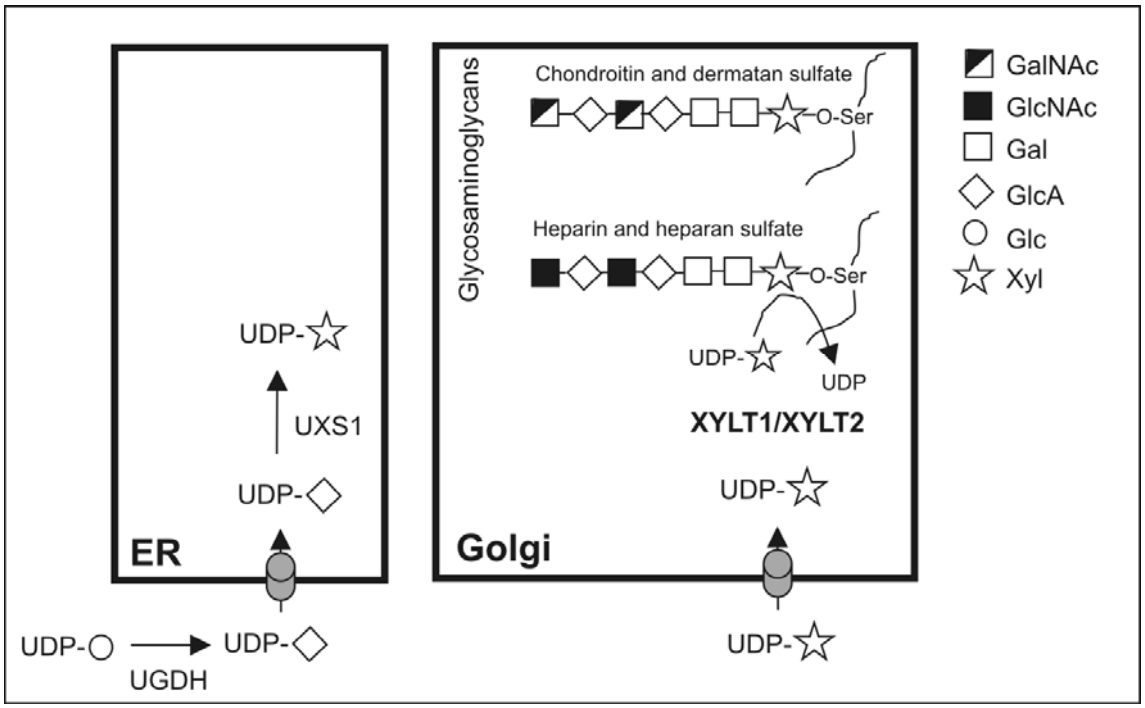


Figure 4.2 XYLT metabolism in mammals (modified from [Bakker et al., 2009]). XYLT enzymes 1 and 2 catalyze in a tissue specific manner the intial reation in glycosaminoglycan (GAG) biosynthesis. UDP xylose is synthesized in two steps by the enzymes UGDH (UDP-Glycose dehydrogenase), forming UDP-GlcA, and UXS (UDP-Xylose synthase). One transporter is capable of transporting UDP-GlcA over the ER-membrane, whereas the other transports UDP-xylose over the Golgi membrane. After the tetrasaccharide linker either GalNAc-GlcA or GlcNAc-GlcA disaccharide units are attached to form chondroitin and dermatan sulfate or heparin and heparan sulfate.

Mutated *XYLT2* transcript should be a subject of nonsense-mediated mRNA decay (NMD; [You et al., 2007]. The interesting question was whether a truncated XYLT2 protein is expressed in biallelically mutated cells or if the transcript is subjected to nonsense-mediated translational repression [You et al., 2007] or the truncated protein is degraded in the proteasome. So far no appropriate antibody specifically detecting the XYLT2 protein is available.

^{3}H-xylose incorporation experiments have demonstrated no differences in MSI-H CRC cell lines and in comparison a high ^{3}H-xylose incorporation rate in control cells HepG2. For HepG2 cells, only expressing the *XYLT2* isoform, ^{3}H-xylose incorporation could be attributed to XYLT2 enzyme activity alone. *XYLT1* expression taking over the function of XYLT2, as well as the heterogeneity of examined cancer cell lines might contribute to the contradictory effects in different *XYLT2* mutated

MSI-H CRC cell lines. To investigate solely XYLT2 function ^{3}H-xylose incorporation could be examined in cell lines deficient for both isoforms or *XYLT1* expression could be silenced by *XYLT1* siRNA. For the mutant chinese ovary hamster cell line CHO pgsA-745 no *XYLT1* and *XYLT2* transcripts were detected which was accompanied by lost chondroitin and heparan sulfate structures on the cell surface [Cuellar et al., 2007; Esko et al., 1985]. For further analysis in MSI-H colorectal cancer cell lines a *XYLT2*-deficient cell line (HDC9) was used for overexpressing *XYLT2*. Changed ^{3}H-xylose incorporation in these cells could be directly linked to higher levels of *XYLT2* expression. To prove that decreased or increased *XYLT2* expression influences the ^{3}H-xylose incorporation rate, HDC9 cells were transiently transfected with either the empty vector control pcDNA3.1 (mock-transfected HDC9 cells) or *XYLT2*-cDNA (*XYLT2*-transfected HDC9 cells). The effects of ^{3}H-xylose incorporation in mock-transfected HDC9 cells and *XYLT2*-transfected HDC9 cells showed only a slight increase in ^{3}H-xylose in HDC9 cells re-expressing *XYLT2*. These low effects in ^{3}H-xylose incorporation were probably caused by low transfection efficiency which was less than 10% for HDC9 cells. The effects of *XYLT2* in ^{3}H-xylose incorporation, could be more reliably evaluated if more cells expressed *XYLT2.* This could be accomplished in stably transfected HDC9 cells expressing *XYLT2*, which are currently established in our laboratory. In addition to a total ^{3}H-xylose metabolic labelling the incorporation rate into individual mucin proteins should be analyzed in future experiments.

Taken together, the results for *XYLT2* on MSI-H CRC cell lines and tumors showed that *XYLT2* is mutated frequently and biallelically and mutations occur in preneoplatic lesion thus providing first evidence for an involvement in MSI tumorigenesis. To further validate these results additional data on XYLT2 protein expression need to be acquired in MSI-H CRC cell lines and primary tissues. Unfortunately, suitable XYLT2-specific antibodies are not available. Alternatively, XYLT2 enzyme activity or loss of activity could be examined in MSI-H colorectal tumor tissues known to carry biallelic mutations.

How might XYLT2 loss of function contribute to MSI-H tumorigenesis? In contrast to almost all other glycosyltransferases the xylosyltransferases are secreted into the extracellular space, although xylosylation of the core protein occurs in the Golgi apparatus [Götting et al., 1999; Götting et al., 2000]. It was shown that

xylosyltransferases are secreted into the extracellular space together with large chondroitin sulfate proteoglycans [Götting et al., 2002], probably controlling proteoglycan synthesis. If the mutated XYLT2 protein is also secreted, this XYLT2 neopeptide would be a suitable marker for MSI tumorigenesis. Functional evidence for XYLT contributing to tumorigenesis and especially of XYLT2 to MSI tumorigenesis was not determined yet. However, as an initial enzyme in proteoglycan synthesis XYLT2 loss probably changes cell surface structures due to a loss of chondroitin, dermatan and heparan sulfates. Significant changes in proteoglycan content have been reported in the stroma surrounding tumors, and it has been suggested that these alterations can support tumor growth as well as progression and invasion [Iozzo et al., 1993; Lesley et al., 1997]. The proteoglycan content in *XYLT2*-deficient cells and tumors has to be proven.

The current data for the first time provide a preliminary evidence for potential association of *XYLT2* mutations with MSI tumorigenesis. Finding evidence for changed XYLT2 activity in MSI-H colorectal tumors would be of great interest, since XYLT2 could be a promising marker for MSI-H tumorigenesis, as measuring the XYLT2 activity in patient`s bloods or other body fluids is efficient.

4.4 Candidate genes with lower mutation frequency or no mutations in their cMNRs

In addition to *LMAN1* and *XYLT2*, cMNR-harboring glycosylation genes with lower or no mutation frequency were observed in DNA frameshift mutation analysis. These genes will be shortly discussed here.

B4GALNT4 (*beta-1,4-N-acetyl-galactosaminyl-transferase 4*), the second gene mutated in more than 30% of MSI-H CRC cell lines, might be an interesting candidate for further analysis. *B4GALNT4* harbors three cMNRS (3 x G7) in its coding region, so mutations may result in three different neopeptides, if mutated proteins are expressed. However, no biallelic mutations were observed for *B4GALNT4* making a role as a classical tumor suppressor less likely for this gene. As a galactosaminyl-transferase B4GALNT4 catalyzes the linkage of N,N-diacetyl-lactose-diamine (LacdiNAc) which is found on glycoproteins and glycohormones of several tissues [Gotoh et al., 2004]. It was reported that *B4GALNT4* showed an

elevated expression in prostate cancer, together with fucosyltransferase I (FUT1) which is responsible for the high amount of PSA (prostate-specific antigen) in prostate cancer [Fukushima et al., 2010]. Expression analyses of *B4GALNT4* in MSI-H CRC have yet not been investigated.

Furthermore glycogenes not mutated in any MSI-H CRC cell line might play such an important role that mutations are suppressed in these genes. Genes as *ABO* (*alpha 1-3-N-acetylgalactosaminyltransferase*), *GOLGA4* (*Golgin A 4*) or *OSGEPL1* (*O-sialoglycoprotein endopeptidase-like 1*) might have important functions so that mutations in these genes are not compatible with cell survival and/or proliferation. *ABO* encodes proteins determining the individual blood groups dependent on the expressed splice variant, with three main variant alleles A, B and 0 [Novaretti et al., 2008]. *AB0* is known to be involved in pancreatic carcinogenesis and frequently shows alterations in tumors, known as Lewis antigens [Brockhausen, 1999; Wolpin et al., 2010]. AB0 blood type could be incorporated into predictive models for pancreatic cancer, because there was a tendency that blood types A, AB, and B were at greater risk for developing pancreatic cancer [Wolpin et al., 2010]. For *GOLGA4* a potential role in dendrocyte differentiation and/or maturation was reported [Cowan et al., 2002]. For *OSGEPL1* only expression analyses made by the human protein atlas consortium (www.proteinatlas.org) are available showing a strong staining in a few CRCs.

This thesis investigated cMNR harboring genes with high mutation frequencies in MSI-H colorectal cancer. The analysis of genes showing no mutation frequency might also be interesting to understand the biology of MSI tumors.

4.5 Conclusions and Perspectives

During the course of this thesis a few frequently mutated genes involved in glycoprotein synthesis and modification were identified in MSI CRC. Two candidates with significantly increased mutation rates were identified (*LMAN1*, *XYLT*2). For these genes insights into potential mechanisms for tumor progression were evaluated, namely biallelic mutations in both genes led to changed glycosylation patterns on the cell surface. The examined cell surface glycoprotein pattern in *LMAN1*-deficient cells revealed structures known to be involved in the progression of

cancer from a tumorigenic to a metastatic phenotype. *LMAN1* and *XYLT2* mutations occurred frequently, early and biallelically in MSI-H CRC tumors. Expression analysis identified *LMAN1*- and *XYLT2*-deficient cell lines providing valuable tools for studying the functional consequences of protein inactivation. The results in primary MSI-H tumor tissues frequently show heterogeneous LMAN1 expression abnormalities caused by regional biallelic cMNR frameshift mutations in tumor cells. *LMAN1* mutations led to decreased secretion of a known tumor growth affecting serine protease inhibitor, A1AT. The loss of LMAN1 transporter function might also decrease A1AT levels in tumor tissues and thus drive tumor progression. Hence, it would be interesting to verify this assumption by determining the A1AT immunohistochemical staining pattern in LMAN1 negative areas. The use of a pulse chase experiment showed that the method in general provides a good platform for the identification of new LMAN1 cargo proteins but several optimization steps have to be performed to detect the probably low abundant proteins. Optimized conditions for a functional purification of LMAN1-CRD could be established during this thesis, however the amounts of protein have to be scaled up. An agarose-immobilized LMAN1-CRD domain allows direct fishing of new LMAN1 cargo proteins. Additional cargo proteins with an involvement in tumorigenesis like A1AT would support, strengthen and/or even explain the role of LMAN1 in influencing MSI tumor progression.

By analyzing the transporter gene *LMAN1* and the xylosyltransferase gene *XYLT2* this thesis identified two members of the cellular glycosylation machinery as novel MSI target genes. At least for the *LMAN1* gene loss of function leads to severe consequences i.e. secretion defects and an altered cell surface glycosylation pattern. How these alterations might affect cell-communication, signalling mechanisms and in general the biology of MSI tumors and warrants further investigation.

Moreover, altered glycosylation pattern are expected to result in altered glycopeptides that could provide an additional source of MSI tumor-specific antigenic peptides complementary to the well-known frameshift neopeptides potentially useful for therapeutic application (vaccination).

Finally, uncovering novel LMAN1-interacting proteins might enhance the understanding of the specificity of the secretory pathway, and in addition, give a more detailed insight into the molecular pathogenesis of MSI tumors.

5. Materials

5.1 Instruments

Instrument (Specification)	Supplier
Agarosegel Chamber Sub Cell GT	Biorad (München, Germany)
ÄktaPurifier FPLC System	GE Healthcare (München, Germany)
Analytical scale (BP 210 D)	Sartorius (Göttingen, Germany)
Camera (Electrophoresis Docu System 120)	Kodak (Stuttgart, Germany)
Centrifuge (5810R)	Eppendorf (Hamburg, Germany)
Centrifuge (Biofuge 13)	Heraeus Holding GmbH (Hanau, Germany)
Centrifuge (Microcentrifuge 1-14)	Sigma Laborzentrifugen (Osterode, Germany)
Centrifuge (Sigma 3MK)	Sigma Laborzentrifugen (Osterode, Germany)
Centrifuge (Varifuge 3.0R)	Heraeus Holding GmbH (Hanau, Germany)
Centrifuge Heraeus (Modell T 110 L)	Heraeus Holding GmbH (Hanau, Germany)
Digital pH Meter pH 525	WTW (Weilheim, Germany)
Electrophoresis chamber (Sub Cell GT)	Biorad (München, Germany)
Electroporator for cells (Nucleofector 1)	Lonza Biosystems (Basel, Switzerland)
ELISA-Reader (GENios)	GENios Tecan (Crailsheim, Germany)
FACS (FACSCalibur)	Becton Dickinson (Franklin Lakes, USA)
Genetic analyzer (ABI 3100)	Applied Biosystems (Darmstadt, Germany)
Incubator (BD6220)	Fisher Scientific (Loughborough, UK)
Liquid Scintillation Counter (TRI-CARB 2900)	Perkin Elmer (Boston, USA)
Microscope (DMBRE)	Leica (Bensheim, Germany)
Microscope (Leica DMIL)	Leica (Bensheim, Germany)
Microscope CK 40	Olympus Europa Holding GmbH (Hamburg)
NuPAGE MES Running buffer	Invitrogen (Karlsruhe, Germany)
NuPAGE X-Cell Sure Lock	Invitrogen (Karlsruhe, Germany)
NuPAGE ZOOM IPG Runner Cassette	Invitrogen (Karlsruhe, Germany)
NuPAGE Protein Electrophoresis System	Invitrogen (Karlsruhe, Germany)
PCR system (GeneAmp 22400)	Perkin Elmer (Waltham, USA)
PH meter (Calimatic 761)	Knick (Berlin, Germany)
Phosphorimager (FLA-3000)	Fujifilm (Düsseldorf, Germany)
Photometer (Ultrospec 3300)	Amersham Pharmacia (Cambridge, UK)
Pipet aid (Pipetman)	Gilson (Limburg-Offheim)
Pipettes (2 µl; 10 µl; 20 µl; 200 µl; 1000µl)	Gilson (Limburg-Offheim)
Power supply (Consort E835)	Peqlab (Erlangen, Germany)
Power supply (Power Pac 300)	Biorad (Munchen)

Robo Cycler Gradient96	Stratagene (Boeblingen, Germany)
Shaker (Certomat H)	Sartorius (Göttingen, Germany)
Sonopuls Homogenizer (Bandelin-Sonopuls)	Bandelin electronic GmbH (Berlin, Germany)
Thermomixer (5436)	Eppendorf (Hamburg, Germany)
Ultracentrifuge (TLA-100.2 rotor)	Beckmann Coulter (Krefeld, Germany)
Ultra-Low Temperatur Freezer MDF-U53V	Sanyo (München, Germany)
UV Transilluminator	Konrad Benda Laborgeräte (Wiesloch, Germany)
Vortex (MS1 Minishaker)	IKA (Staufen, Germany)
Waterbath (Grant SUB6)	Grant (Cambridge, UK)
Waterbath (SW 20)	Julabo Labortechnik (Seelbach, Germany)

5.2 Consumables, reagents and chemicals

Item	**Supplier**
^{125}I-iodine (100 mCi/ml)	Hartmann Analytic (Braunschweig, Germany)
^{35}S-methionine (10 mCi/ml)	Hartmann Analytic (Braunschweig, Germany)
37% formaldehyde	Carl Roth (Karsruhe, Germany)
[1-3 ^{3}H]-D--xylose (10 mCi/ml)	Hartmann Analytic (Braunschweig, Germany)
Acetic acid	Serva (Heidelberg, Germany)
Alpha1-Antitrypsin	Biopur (Bubendorf, Germany)
Alpha1-Antitrypsin-Agarose	Biopur (Bubendorf, Germany)
Ampicillin sodium salt	Sigma-Aldrich (Taufkirchen, Germany)
Bacto-Agar	Fluka Chemie GmbH (Buchs, Switzerland)
beta-cyclodextrin	Promega (Madison, USA)
BigDyeTerminator v1.1 Sequencing Kit	Applied Biosystems (Darmstadt, Germany)
Boric acid	Merck (Darmstadt, Germany)
Bovine Serum Albumin (BSA)	Sigma-Aldrich (Taufkirchen, Germany)
$CaCl_2$	Merck (Darmstadt, Germany)
Citric acid	Merck (Darmstadt, Germany)
Diaminobenzidine (Liquid DAB+ substrate)	DAKO (Hamburg, Germany)
Dimethyl sulfoxide (DMSO)	Merck (Darmstadt, Germany)
Disodium hydrogen phosphate (Na_2HPO_4)	VWR International GmbH (Bruchsal, Germany)
DMEM / Ham's F-12 with L-Glutamine	PAA (Cölbe, Germany)
dNTP-Mix	Invitrogen (Karlsruhe, Germany)
Dulbecco's PBS (1x) without Ca & Mg	PAA (Cölbe, Germany)
Ethanole absolute	Merck (Darmstadt, Germany)
Ethidiumbromide	Sigma-Aldrich (Taufkirchen, Germany)
Ethylene glycol tetraacetic acid (EGTA)	Sigma-Aldrich (Taufkirchen, Germany)
Ethylenediaminetetraacetic acid (EDTA)	Merck (Darmstadt, Germany)

Fetal Calf Serum (FCS)	Invitrogen (Karlsruhe, Germany)
Folin-Ciocalteu phenol reagent	Merck (Darmstadt, Germany)
FuGENE HD Transfection Reagent	Roche (Mannheim, Germany)
G-418 Sulphate	PAA (Cölbe, Germany)
Gel-Dry Drying Solution	Invitrogen (Karlsruhe, Germany)
Glycerol 86%	Carl Roth GmbH (Karlsruhe, Germany)
Guanidine-hydrochloride	Novagen/Merck (Darmstadt, Germany)
HEPES	Invitrogen (Karlsruhe, Germany)
HIDI-formamide	Applied Biosystems (Darmstadt, Germany)
Horse serum	Vector Laboratories (Burligame, USA)
Isopropyl ß-D-1-thiogalactopyranoside (IPTG)	Sigma-Aldrich (Taufkirchen, Germany)
Kodak BioMax films	Sigma-Aldrich (Taufkirchen, Germany)
L-Arginin-hydrochloride	Novagen/Merck (Darmstadt, Germany)
Lipofectamine 2000 Reagent	Invitrogen (Karlsruhe, Germany)
Mannose sepharose 4B	GALAB (Geessthacht, Germany)
Methanol	Merck (Darmstadt, Germany)
$MgCl_2$	Merck (Darmstadt, Germany)
Milk powder	Carl Roth GmbH (Karlsruhe, Germany)
NaOH	AppliChem (Darmstadt, Germany)
Ni^{2+}-sepharose (HisTrap FF)	GE Healthcare (Uppsala, Sweden)
N-Laurosylsarcosine	Merck (Darmstadt, Germany)
NuPAGE 4 – 12% Bis-Tris Mini Gel	Invitrogen (Karlsruhe, Germany)
Oligo(dT) primers	Invitrogen (Karlsruhe, Germany)
PEG-6000	Sigma-Aldrich (Taufkirchen, Germany)
Penicillin / Streptomycin (100x)	PAA (Cölbe, Germany)
Pierce Iodination Beads	Thermo Scientific (Rockford, USA)
Potassium chloride	J.T. Baker (Deventer, Holland)
Potassium dihydrogenphosphate	Gerbu Biochemicals (Gaisberg, Germany)
Potassium sodium tartrate ($KNaC_4H_4O_6$)	Merck (Darmstadt, Germany)
RPMI 1640 with L-glutamine	PAA (Cölbe, Germany)
RPMI-1640 without methionine and L-glutamine	Sigma-Aldrich (Taufkirchen, Germany)
Silver nitrate ($AgNO_3$)	Carl Roth GmbH (Karlsruhe, Germany)
SOC-medium	Invitrogen (Karlsruhe, Germany)
Sodium acetate ($C_2H_3NaO_2$)	J.T. Baker (Deventer, Holland)
Sodium bisulfite ($Na_2S_2O_3$)	Merck (Darmstadt, Germany)
Sodium carbonate (Na_2CO_3)	J.T. Baker (Deventer, Holland)
Sodium Chloride (NaCl)	AppliChem (Darmstadt, Germany)
Trichloracetic acid (TCA)	Carl Roth GmbH (Karlsruhe, Germany)
Tris (2-carboxyethyl) Phosphine HCl (TCEP)	Novagen/Merck (Darmstadt, Germany)
Tris Base	Carl Roth GmbH (Karlsruhe, Germany)

Triton-X-100	Sigma-Aldrich (Taufkirchen, Germany)
Trypsin-EDTA (1x) 0,05 % / 0,02 % in D-PBS	PAA (Cölbe, Germany)
Trypton/Pepton aus Casein	Carl Roth GmbH (Karlsruhe, Germany)
Tween-20	Sigma-Aldrich (Taufkirchen, Germany)
Ultima Gold LSC Cocktail	Perkin Elmer (Boston, USA)
UltraPureTM Agarose	Invitrogen (Karlsruhe, Germany)
Western Lightning Chemiluminescence Reagent Plus	Perkin Elmer (Boston, USA)
Yeast extract	Carl Roth GmbH (Karlsruhe, Germany)

5.3 Commercially available kits

Item	**Supplier**
A1AT-ELISA	ImmunDiagnostik (Bensheim Germany)
Biorad Protein Assay	BioRad (München, Germany)
CellTiter 96® AQ_{ueous} One Solution Cell Proliferation Assay (MTS)	Promega (Mannheim, Germany)
DNeasy Blood & Tissue Handbook	Qiagen (Hilden, Germany)
Endo-free Plasmid Maxi Kit	Qiagen (Hilden, Germany)
Endoplasmic Reticulum Isolation Kit	Sigma-Aldrich (Taufkirchen, Germany)
High Pure PCR Product Purification Kit	Roche (Mannheim, Germany)
iFold Protein Refolding System 1	Novagen/Merck (Darmstadt, Germany)
JetQuick Gel Extraction Kit	Genomed (Löhne, Germany)
NucleoSpin Plasmid Kit	Machery Nagel (Düren, Germany)
QIAfilter Plasmid Purification	Qiagen (Hilden, Germany)
QIAquick Gel Extraction Kit	Qiagen (Hilden, Germany)
RNeasy blood and tissue kit	Qiagen (Hilden, Germany)
Vectastain Elite ABC kit	Vector (Burligame, USA)

5.4 Enzymes, bacteria, antibodies, markers and vectors

Item	**Supplier**
Enzymes	
Calf Intestinal Phosphatase (CIP)	Roche (Mannheim, Germany)
T4-Ligase	Roche (Mannheim, Germany)
Taq-Polymerase	Invitrogen (Karlsruhe, Germany)
Phusion High Fidelity Taq-Polymerase	New Englan Biolabs (Frankfurt, Germany)
SuperscriptII Reverse Transcriptase	Invitrogen (Karlsruhe, Germany)
KpnI	Promega (Madison, USA)
NotI	Promega (Madison, USA)

Enterokinase	Novagen/Merck (Darmstadt, Germany)
Bacteria	
DH5 alpha	Invitrogen (Karlsruhe, Germany)
BL21 (DE3)	Invitrogen (Karlsruhe, Germany)
Antibodies (clone, dilution, specification)	
LMAN1 (polyclonal, 1:1000, none)	Sigma-Aldrich (Taufkirchen, Germany)
Actin (monoclonal C4, 1:10.000, none)	MP Biomedicals (Solon, USA)
Anti-rabbit IgG (1:2500, HRP Conjugate)	Promega (Madison, USA)
Anti-mouse IgG (1:5000, HRP conjugate)	GE Healthcare (Munich, Germany)
Anti-mouse/anti-rabbit antibody (1:50,biotinylated)	Vector (Burligame, USA)
Streptavidin-conjugated fluorescence	
Streptavidin／R-phycoerythrin	Sigma-Aldrich (Taufkirchen, Germany)
protein molecular weight markers	
See blue plus pre-stained protein standard	Invitrogen (Karlsruhe, Germany)
Mark12 unstained standard	Invitrogen (Karlsruhe, Germany)
DNA markers	
1 kb DNA ladder	Invitrogen (Karlsruhe, Germany)
100 bp DNA ladder	Invitrogen (Karlsruhe, Germany)
ROX size standard	Applied Biosystems (Darmstadt, Germany)
Vectors	
pET44a(+)	Novagen/Merck (Darmstadt, Germany)
pcDNA3.1	Invitrogen (Karlsruhe, Germany)

5.5 Oligonucleotides

Table 5.5.1 Oligoneucleotides designed for Frameshift mutation analysis

Gene name (Ensembl)	**Sequence (5`- 3`)**
Abo_a	GGACGAGGGCGATTTCTACT
Abo_sf	TCAGGTGGCTCTCGTCGT
Alg12_c7a_a	TCTTGGGGCACCTAAGACAG
Alg12_c7a_sf	TGGGATGCCTGTGAATACAA
Alg12_c7b_a	TCCCATTTCAACTACCCAGG
Alg12_c7b_sf	GTCGGAGCCACAGCAGTC
B3galtl_a	TACATAACTGGCAGACCCCC

B3galtl_sf	GGAATTGTATTCGTCATCCAGAG
B4galnt1_a	TGTGTAGCTGGGCTCTGATG
B4galnt1_sf	AGCTCGGACTTGTTTCTGGA
B4galnt4_g7a_a	GTGAAGTTCAGTGGCAGCCT
B4galnt4_g7a_sf	GTGTGTGTGACCTCCTCCCT
B4galnt4_g7b_a	CCCCCTCTCACTGAGGAAA
B4galnt4_g7b_sf	GTCAGCCTTGAAGATGGACC
B4galnt4_g7c_a	ATAAATCCCCCTTCCCCATT
B4galnt4_g7c_sf	GAGGAGTTCCGAGACCAGTG
B4galt2_a	CAGAGAAGGAGCACAGCCTT
B4galt2_sf	CCTGACCTGCTGGATCTGTT
Casd1_a	GCAACTACAAAGAAATTTCACACG
Casd1_sf	TCCCCTTGGTCACTGTATGG
Chgut_a	AGTCAGCGTTGTAGAAGCAGC
Chgut_sf	CCCTGTCACCACAGAGATCA
Chsy1_a	CTCTTTCTTGGGCAGACAGG
Chsy1_sf	ACCATTCTCCGAAGCACCT
Dpagt1_a	GACCAAGGTTTGCAGAGGAA
Dpagt1_sf	AGAGCAATCCCAAAGTGGTG
Galntl1_a	CCTTAGAGCAGGTCCTCACG
Galntl1_sf	GGGAAAGTATGATGCCCAGA
Gcnt2_a	CCACATAGGCAGAGCCAAAG
Gcnt2_sf	TCATGCAATTGGACGGACTA
Gcnt4_a	TGAACTTAGACGGGTGCCTT
Gcnt4_sf	CAAAATAAGCACTGCCAACAAA
Gcs1_a	TGTCTTCTGGACCTCCAACC
Gcs1_sf	CTCTGTCCTCCCACTTCAGG
Glt8d1_a	CTTTGTCCCAAATGCTCTCC
Glt8d1_sf	AGCGAGTGTTGTGCTGAATG
Golg4_a	TTTTTCGGATAGGCTGATGC
Golga4_sf	TGTCACATTGATGAAAGAAGAGC
Gyltl1b_a	CACAAGGGTGACATCGTGAG
Gyltl1b_sf	AGACCCCTGCTTTGAGTTCC
Lman1_a	GGAGGAATTTGAGCACTTTCA
Lman1_sf	CACCCATGTCAGCTTTGCTA
Man1b1_a	AACTGCTTGAATGCGAGTCC
Man1b1_sf	CAACCTGTTTGAGAGCACGA
Osgepl1_a	GCTAATCTTGACTAAGACTGCAGGA
Osgepl1_sf	TGAAGAAATAGTGTTCCAGGATGA
Pigb_a	AAAATCATTTGCCTCTGTTCTAAAG
Pigb_sf	GAGTTTGTCTCTGATGATTGATCG
Pigz_c7_a	AGCACACCCACCCACTACAC
Pigz_c7_sf	ATGTCCACCACCTCCACTG
Pigz_g7_a	CAGTGGAGGTGGTGGACAT
Pigz_g7_sf	CAGGGGTTACCACAAAGAGG
St6galnac6_a	GTCTCATTCGTGGTTGAGCA
St6galnac6_sf	GAGGGCTCACCTGCAGTAGT
St8sia6_a	TTTTTAATTTGGCAGATATGGGA

St8sia6_sf	AGGAAAATGCTGGCAGAAGA
Ust_a	CTCGGTTTCCAGTGGAGTTG
Ust_sf	ACCAGCGGATGAGATACGAG
Xylt2_a	CATGACATCACTGAGCCCAC
Xylt2_sf	TCGAGTCGACTGTGAACCAG

a: antisense sf: sense Fluorescein labelled; Gene names as annotated on the ENSEMBL homepage (www.ensembl.org/index.html).

Table 5.5.2 Oligonucleotides designed for cloning and RT-PCR

Cloning	**Sequence (5` - 3`)**	**Usage**
LMAN1-CRD_s	PO_4-GGACGGCGTGGGAGGA	cloning in pET44a(+)
LMAN1-CRD_a	PO_4-TCAGTGGACCGTAGACAAACATG	
LMAN1c_KpnI_s	GTGGTACCAAGATGGCGGGATCCAGGC	cloning in pcDNA3.1
LMAN1c_NotI_a	CCTGGCGGCCGCAGGAAAATGGTAGTCAAAAG	
RT-PCR		
LMAN1c_sf	GCTCTCGATCTCCTACACTCTCA	LMAN1 RT-PCR: size: 122 bp
LMAN1c_a	GGAGGAATTTGAGCACTTTCA	
A1AT_s	CGAAGAGGCCAAGAAACAGA	A1AT RT-PCR Ex2-Ex4 size: 484 bp
A1AT_a	GACCTTAGTGATGCCCAGTTG	
XYLT2c_s	TCGAGTCGACTGTGAACCAG	XYLT2 RT-PCR size: 298 bp
XYLT2c_a	AGGTAGCCCTGGAAATGGTC	
XYLT1c_Ex6	ATGCAAGGTTCATTCGGAAG	XYLT1 RT-PCR size: 250 bp
XYLT1c_Ex8	CTTCCATACGGTCCTGGAGA	
GAPDH_s	CCACCCAGAAGACTGTGGAT	GAPDH RT-PCR size: 119 bp
GAPDH_a	TTCAGCTCAGGGATGACCTT	

a: antisense; sf: sense Fluorescein labelled; c: coding region; PO_4: 5`-phosphorylation for blunt end ligation; Ex: represents the oligonucleotide location regarding the exons.

5.6 Buffers

Standard buffers used and not specifically mentioned in the method part (section 6).

LB-medium
10 g Trypton 5 g Yeast 10 g NaCl Ad 1000 ml H_2O dest, adjust pH to 7.0, autoclave
LB-Agar with ampicillin
LB-medium 2% Bacto-agar add 100 mg ampicillin per liter (100 µg/ml)
RIPA buffer
50 mM Tris base, adjust pH to 7.4 with HCl 150 mM NaCl 1% Triton-X-100 1% Sodium desoxycholate 0.1 mM $CaCl_2$ 0.1% SDS 0.01 mM $MgCl_2$
TBE electrophoresis buffer (10x)
108 g Tris base 55 g Boric acid 40 ml 500 mM EDTA Ad 1000 ml H_2O dest, adjust pH to 8.0
TBS (1x)
50 mM Tris base 150 mM Sodium chloride 1 M Hydrochloric acid (~ 9.5 ml) Ad 1000 ml H_2O dest, adjust pH to 7.5 with HCl.

Lowry Buffer

Reagent A

0.02% Cupper sulfate ($CuSO_{4)}$

4% Sodiume carbonate ($Na_2CO_{3)}$

0.8% Sodium hydroxide (NaOH)

0.04% Potassium sodiume tartrate ($KNaC_4H_4O_6$ •4 H_2O)

Reagent B

Folin-Ciocalteu phenol reagent 5x diluted in H_2O dest.

Citrate Buffer (10 x)

21.0 g Citric acid

Ad 1000 ml H_2O dest, adjust pH to 6.0 by the usage of NaOH platelets, store at 4°C.

Potassium phosphate buffer

50 mM K_2HPO_4

50 mM KH_2PO_4

Titration to pH 7.4.

HBS-Buffer (10 x)

140 mM NaCl

50 mM HEPES

750 µM Na_2HPO_4

Adjust pH to 7.12.

PBS (10 x)

80 g Sodium chloride

2 g Potassium chloride

14.4 g Disodium hydrogenphosphate

2.4 g Potassium dihydrogenphosphate

Ad 1000 ml H_2O dest, adjust pH to 7.2.

6. METHODS

6.1 Molecular Biology methods

6.1.1 Isolation of genomic DNA and RNA

Genomic DNA was isolated from cancer cell lines and paraffin-embedded archival specimens after manual microdissection of tumor and corresponding normal tissues using the DNeasy Blood & Tissue kit. RNA was isolated using the RNeasy Mini kit.

6.1.2 Oligonucleotide design

Oligonucleotides for frameshift mutation analysis, sequencing analysis and cloning PCRs were designed by the usage of Primer 3.0 software. Oligonucleotides were purchased from Thermo Scientific in HPLC-purified quality. Oligo (dT) primers were used as suggested by the supplier.

6.1.3 Standard Polymerase Chain Reaction (PCR)

For a standard PCR approach taq DNA polymerase with appropriate buffers was used. For cloning PCRs the Phusion® High Fidelity taq polymerase was selected which exhibits 3` − 5` exonuclease activity for proof-reading (6.1.7). PCRs were performed in RoboCycler® Gradient 96 Temperature Cyclers from Stratagene.

Cycling conditions for Taq DNA polymerase

Step	Number of cycles	Temperature	Time
Initial denaturation	1	94°C	4 min
Cyclic amplification (1)	35	94°C	30 sec
(2)		58°C – 62°C	45 sec
(3)		72°C	1 min
Final elongation	1	72°C	5 – 10 min

(1) denaturation, (2) annealing, (3) elongation

Standard PCR pipetting scheme

Component	Final Concentration	Volume
Reaction buffer (10x)	1 x	2.5 µl
$MgCl_2$ (50 mM)	1.5 mM	0.75 µl
dNTP-Mix (2 mM)	200 µM	2.5 µl
Primer_sense (5µM)	0.5 µM	2.5 µl
Primer_antisense (5µM)	0.5 µM	2.5 µl
Taq DNA polymerase 5 U/µl		0.125 µl
H_2O HPLC-grade	-	13.125 µl
Template DNA*	10 - 500 ng	1.0 µl
Total volume	-	25 µl
* 10-20 ng for plasmid-DNA and 200-500 ng for genomic DNA		

6.1.4 Sequencing

PCR products, like frameshift mutation fragments or cloned plasmids were sequenced using the BigDye Terminator v1.1 sequencing kit.

The following reaction mix was used

	Concentration	Volume
Template[1]	-	10.0 µl
BigDye Ready Reaction Mix	-	8.0 µl
Sequencing Primer	5 µM	2.0 µl
[1]Purified PCR product or plasmid (approx. 350 ng)		

The sequencing reaction was performed under identical conditions for all primers and templates used. In order to purify the obtained sequencing products, the products were precipitated with 10 µl 3 M sodium acetate (pH 4.5) and 250 µl of 100% EtOH and centrifuged at 13.000 rpm for 15 min at room temperature. The supernatant was discarded and the products were washed with 250 µl of 70% EtOH and centrifuged at 13.000 rpm for 5 min at room temperature. The supernatant was again discarded and the pellets were dried for 4 min in a vacuum centrifuge to remove traces of EtOH. DNA pellets were afterwards resolved in 12 µl Hi-Di formamide and analyzed on an

ABI3100 genetic analyzer. The obtained results were evaluated with Sequencing Analysis software (Applied Biosystems, version 3.7).

6.1.5 cMNR Frameshift mutation analysis

Frameshift mutation analyses were performed on genomic DNA of cell, tumor and blood samples derived as mentioned in chapter 6.1.1. For frameshift mutation analysis on tumor genomic DNAs (gDNAs) oligonucleotides were designed to obtain short amplicons of about 100 bp, thus allowing robust amplification from archival tissues. Sense primers were Fluorescein labelled (sf; Table 5.5.1). PCR was performed as described in 6.1.2. PCR fragments were analysed in 2% agarose gels and subsequently on an ABI3100 Genetic Analyzer. Size, height and profile of microsatellite peaks were analyzed using Sequencing analysis software (version 3.7). Coding microsatellite instability was scored if smaller or larger-sized amplimeres were detected in tumor DNA compared to DNA from normal tissue. Frameshift mutations were confirmed by DNA sequence analysis.

6.1.6 Reverse transcription PCR

One microgram of total cellular RNA was reverse transcribed using Superscript II reverse transcriptase according to the manufacturer's instructions. Primers for the *LMAN1* gene were designed to flank the A9 cMNR and spanned exon 8 and exon 9 to exclude amplification from genomic DNA. Primers for the *XYLT2* gene were designed to flank the C7 cMNR. PCR products were visualized on ethidium bromide-stained 2% agarose gels. All cell lines were examined by frameshift mutation analysis on cDNA for comparison with genomic DNA mutation data. Mutations were confirmed by DNA sequence analysis. As a control for loading and integrity of mRNA, reverse transcription-PCR (RT-PCR) analysis of *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) was performed. Table 5.5.2 presents the primers designed for RT-PCRs of *LMAN1*, *XYLT2*, *XYLT1*, *A1AT* and *GAPDH*.

6.1.7 Restriction digest

Restriction enzymes cut double stranded or single stranded DNA at specific restriction sites. The activity of restriction enzymes depends on their sequence specificity and cofactor requirements. Many techniques such as cloning, mutational analysis, probe preparation and linearization before stable transfection require the use of restriction digests. One unit of restriction enzyme will completely digest 1 µg of substrate DNA in a volume of 50 µl in 60 min, for most enzymes at 37°C in a water bath. Restriction enzymes were received from New England Biolabs, Roche or Promega (see section 5).

Below a typical restriction digest is depicted

Restriction Enzyme	1 Unit
DNA	1 µg
10 x Restriction buffer	5 µl (1x)
BSA	100 µg/ml (1x) if sufficient
Total Reaction	50 µl
Incubation Time	1 hour
Incubation Temperature	enzyme-dependent

6.1.8 Cloning of DNA fragments

Preperation of vectors and inserts

During cloning experiments a DNA fragment (e.g. gene of interest GOI) is inserted into a suitable vector for downstream protein expression of the GOI. The coding sequence of *LMAN1* was PCR-amplified from cDNA with primers introducing KpnI and NotI restriction sites for cloning full length *LMAN1* into the pcDNA3.1-vector. Wild type cDNA of *LMAN1* was amplified according to the Phusion PCR program using 2 µl of cDNA derived from the *LMAN1* wild type cell line SW948 as a template. *LMAN1*-CRD was PCR-amplified from *LMAN1*-pcDNA3.1 vector with 5`-phosphorylated primers for blunt end cloning into the pET44a(+) vector. Both constructs are outlined in Figure 6.1, presenting important amino acid (aa) residues for the transport function

of LMAN1 and the location site of the cMNR. For the amplification of *LMAN1*-CRD, 2 ng of plasmid DNA were used. The *XYLT2*-pcDNA3.1 vector was a kind gift of Dr. Hans Bakker (Medizinische Hochschule Hannover). The obtained PCR products were visually controlled on 1 – 2 % agarose gels and purified by gel extraction using the JETQuick Gel Extraction Spin Kit. The purified PCR product of full length *LMAN1* was digested with KpnI and NotI (see 6.1.7), creating “sticky ends” to ensure the appropriate orientation of the insert in the vector after ligation.

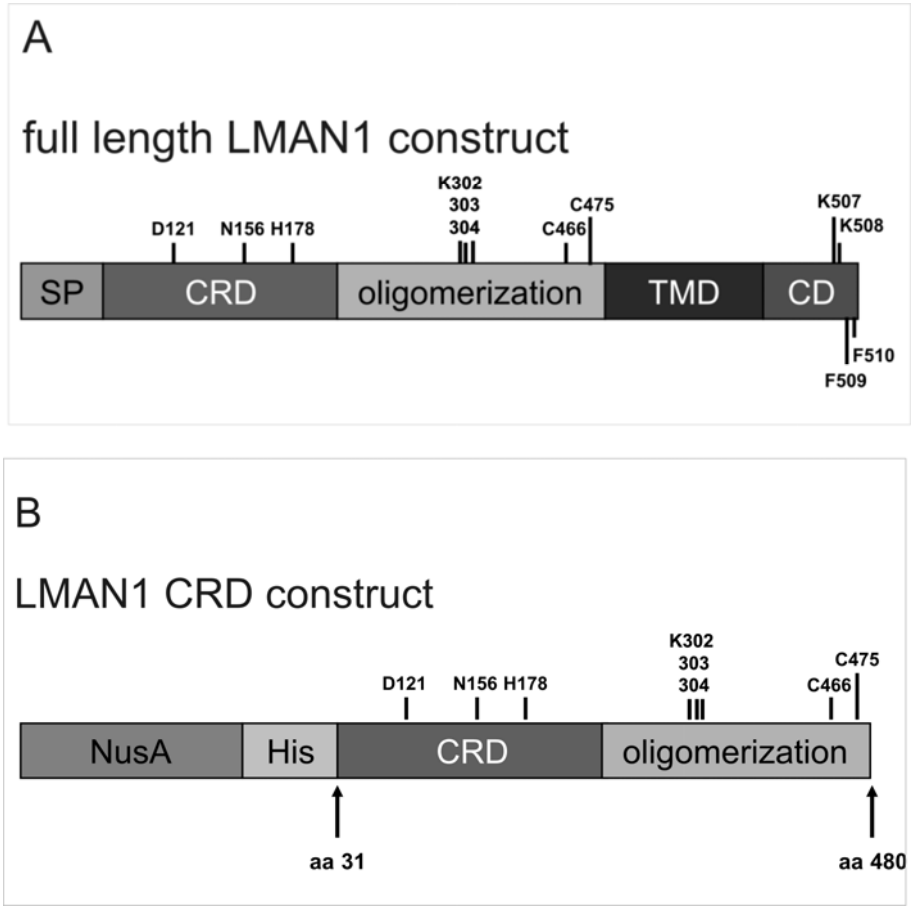


Figure 6.1 LMAN1 constructs [Hauri et al., 2000]. **A.** Full length *LMAN1* used for mammalian protein expression and **B.** presenting the *LMAN1*-CRD with oligomerization domain and N-terminal fusion domains for expression in bacteria. In both constructs the location of the *LMAN1* cMNR is shown (A9 → K302-304). Double lysines in C-terminal position (K507+K508) are required as ER-retention motif the double phenylalanine motif (F509+F510) is an ER-exit determinant. Disulfide linked oligomerization is mediated by two luminal cysteine residues C466 and C475. D121 and N156 are two conserved residues required for the interaction with Ca^{2+} ions and the monosaccharides over hydrogen bonds. H 178, acts as a pH-sensor, is necessary for Ca^{2+}

binding and substrate release in the ERGIC because of a more acidic environment [Zhang et al., 2009]. SP: signal peptide; CRD: carbohydrate recognition domain; NusA: fusion tag; His: histidine tag; TMD: transmembrane domain; CD: cytosolic domain.

Pipetting instruction for Phusion-Tag polymerase

Component	Concentration	Volume
Reaction buffer	5 x stock	4.0 µl
dNTP-Mix	10 mM	0.4 µl
Primer_sense	5 µM	2.0 µl
Primer_antisense	5 µM	2.0 µl
Phusion DNA polymerase	2 U/µl	0.2 µl
H_2O HPLC-grade	-	9.4 µl
Template DNA*	10 - 500 ng/µl	1.0 µl
* 10-20 ng for plasmid-DNA and 200-500 ng for genomic DNA		

Cycling conditions for Phusion-Tag polymerase

Step	Number of cycles	Temperature	Time
Initial denaturation	1	98°C	30 sec
Cyclic amplification (1)	25 – 35	98°C	5 – 10 sec
(2)		58°C – 62°C	10 – 30 sec
(3)		72°C	15 – 30 sec/1kbp
Final elongation	1	72°C	5 – 10 min
(1) denaturation, (2) annealing, (3) elongation			

Vector pcDNA3.1 was also digested with KpnI and NotI and purified by gel extraction. For the blunt end cloning of *LMAN1*-CRD pET44a(+) vector was digested with PshAI and dephosphorylated using the calf intestinal phosphatase (CIP) to prevent self-ligation. One unit CIP dephosphorylates 1 pmol 5`- terminal phosphorylated DNA fragments in 60 min at 37°C.

Calculation for the dephosphorylation of 10 µg (10 x 10^{-6}g) digested plasmid

MW of one nucleotide (1bp) = 330 g/mol

MW of pET44a(+) single stranded = 7311 bp = 2.412.630 g/mol

MW of pET44a(+) double stranded = 4,82526 x 10^6 g/mol

10×10^{-6} g / $4{,}75215 \times 10^{6}$ g/mol = $2{,}1 \times 10^{-12}$ mol

As a double stranded DNA fragment possessed two 5`- ends the pET44a(+)- vector contains altogether 2.1 pmol 5`-ends. 2.1 units CIP were required to dephosphorylate 10 µg of pET44a(+) vector.

Procedure

Step	Component	Incubation Time
Dephosphorylation	**10 µg digested vector DNA**	**1 hour at 37°C (water bath)**
	3 units CIP	
	5 µl 10 x CIP buffer	
	Total volume 50 µl	
Heat inactivation	**200 mM EGTA**	**10 min 75°C**

Ligation

After restriction digest and dephosphorylation steps vector and PCR product were ligated at a molar ratio of vector-DNA to insert-DNA of 1 : 3. The ligation catalyzes the formation of phosphodiester bonds between a 5`-phosphate and a 3`-hydroxyl terminus in a digested double DNA strand. Blunt end ligation was performed at 16°C and sticky end ligation at 4°C over night with T4-ligase and the appropriate 10 x reaction buffer. As a re-ligation control vector without insert underwent the same procedure.

Calculation of the molar amounts of PCR products and vectors

DNA amount (mol) = weight (g) / (660 g/mol x bp)

Transformation into bacteria

After ligation the ligation reactions are transformed into bacteria for the production of plasmid-DNA. Chemically competent DH5α-cells were stored at – 80°C and thawed on ice directly prior to use. 2 µl of a ligation reaction was added and gently stirred. The mixture was incubated on ice for 30 min followed by a heat-shock in a water bath

at 42°C for exactly 30 seconds and storage on ice for 2 min. Subsequently, 250 µl pre-warmed SOC-medium was added to the transformation reaction and mixed. After 1 hour incubation at 37°C in a shaking incubator 50 µl to 100 µl of each transformation reaction were plated on agar plates containing 100 µg/ml ampicillin, used for the positive selection of bacteria that had taken up the vector during transformation. Grown clones were checked by Colony-PCR with insert-specific primers and only clones containing the insert as determined by gel electrophoresis were used for the isolation of plasmids. For Colony-PCR bacterial clones were picked into a reaction tube containing water and were heated for 10 min at 95°C. 2 µl of bacteria solution was used in a Standard-PCR described in 6.1.2. Positive clones were grown in 5 ml over-night cultures LB medium (100 µg/ml ampicillin) at 37°C.

Plasmid preperation

Plasmids were purified from 4 ml of the 6 ml over-night culture using the NucleoSpin Plasmid kit according to the manufacturer's instructions. Cloned insert cDNAs of *LMAN1* or *LMAN1-CRD* in the plasmid-preparation were verified by direct sequencing (6.1.3) of overlapping fragments before starting transfection or expression experiments. For transfection of plasmid-DNA into mammalian cells, bacteria were grown in 100 ml over-night cultures and plamids were purified under endotoxin-free conditions using the Endo-free Plasmid Maxi kit.

Cryopreservation of positive bacteria clones

For long term storage, bacteria were cryo-preserved by mixing 1 ml of cultured bacteria with 250 µl glycerol and storage at –80°C.

6.2 Biochemical methods

6.2.1 Determination of protein concentration

Bradford Assay

The Bradford protein assay is based on an absorbance shift in the dye Coomassie brilliant blue G. Coomassie binds to basic (e.g. arginine) and aromatic amino acid (aa) residues of proteins thereby shifting itsabsorbance from 465 nm to 595 nm. The increase of absorbance at 595 nm is proportional to the protein concentration present

in the sample. Bovine serum albumin was used as a standard (0, 1, 2, 5 and 10 µg/ml) protein sample was 10 x diluted in H_2O dest. The Bradford Assay was performed using the BioRad Protein assay reagent according to manufacturer`s instructions.

Lowry Assay

The Lowry protein assay is based on Cu^+-ions binding to peptide bonds under alkaline conditions with the oxidation of aromatic protein residues. Cu^+ -ions build an unstable blue complex with the yellow Folin-Ciocalteau phenol reagent proportional to the protein concentration present in the sample. The concentration of the reduced Folin reagent is measured by absorbance at 630 nm [LOWRY et al., 1951]. Protein standards and protein samples were used as described above. After precipitation of proteins with an equal volume of 10% TCA (Trichloroacetic acid), pellets were solved in 50 µl 200 mM NaOH at 65°C for 10 min. Proteins were first incubated with 500 µl reagent A (5.6) for 20 min at room temperature, followed by an incubation step with 500 µl reagent B (5.6) for 30 min at room temperature.

For both methods the protein absorbance was measured using the Ultrospec™ 3300 *pro* UV/Visible Spectrophotometer.

6.2.2 Polyacrylamide Gelelectrophoresis

Cell pellets were lysed in RIPA-buffer, sonicated and subsequently ultracentrifugated at 430.000 g for 15 min at 4°C (Beckmann TLA 100.2 rotor). Protein concentration was determined using the Bradford or Lowry method (6.2.1) and 13 to 65 µg protein was incubated in NuPAGE LDS Sample Buffer at 95°C for 5 min. Proteins were separated using NuPAGE® Bis-Tris Mini gels from Invitrogen (acrylamide concentration 4 – 12%; 1.0 mm x 12 well). The NuPAGE® System is based on a Bis-Tris-HCl buffered (pH 6.4) polyacrylamide gel, with a separating gel that operates at pH 7.0. Denatured protein was loaded together with a molecular weight marker. For Western blot analysis the SeeBlue® Plus Pre-stained protein standard and for silver gels the Mark12™ unstained standard were applied. As running buffer NuPage MES running buffer was used. For run conditions the manufacturer`s instructions were followed.

6.2.3 Western blot analysis

Protein preparation and separation was perormed as mentioned in 6.2.4. Protein concentration was determined by Lowry method (6.2.1). 65 µg of protein was seperated as described in 6.2.2 followed by Western blot analysis using primary antibody directed against LMAN1 or as a control anti-actin antibody. Horseradish peroxidase (HRP) – conjugated antibodies were used as secondary antibodies (Anti-Rabbit IgG or Anti-Mouse IgG; see 5.4). Visualization was performed with Western Lighting Chemiluminescence Reagent Plus on Kodak BioMax light films.

6.2.4 Silver staining of polyacrylamide gels

Silver staining is the most sensitive method for permanent visible staining of proteins in polyacrylamide gels with a detection limit of 1 – 5 ng protein. Silver ions (Ag^+) build complexes with Glu-, Asp- and Cys-residues of proteins. By treatment with formaldehyde, Ag^+-ions are reduced and form an insoluble brown precipitate of metallic silver. 65 µg of protein was separated as described in 6.2.2 and subsequently stained as described below. All solutions were prepared in H_2O dest and complete staining protocol was performed at room temperature. Gels were preserved in Gel-Dry™-Drying solution for long-term storage.

Solutions

Fixing solution		
Fixation 1	Fixation 2	Fixation 3
10% (v/v) 96% acetic acid	5%(v/v) 96% acetic acid	10%(v/v) ethanol absolute
40%(v/v) methanol	10%(v/v) ethanol absolute	

Sensitizing solution		**Silver solution**
0.02% (w/v) $Na_2S_2O_3$		0.2% (w/v) $AgNO_3$
0.02% (v/v) 37% formaldehyde		0.05% (v/v) 37% formaldehyde

Developing solution		**Stop solution**
6% (w/v) Na_2CO_3		10% (v/v) 96% ethanol
0.4% (w/v) $Na_2S_2O_3$		
0.02% (v/v) 37% formaldehyde		

Staining protocol

Staining solution	Repeats	Incubation time
Fixation 1	1 x	1 h
Fixation 2	1 x	30 min
Fixation 3	1 x	15 min
H_2O dest	3 x	1 min
Sensitizing solution	1 x	1 min
H_2O dest	1 x	1 min
Silver solution	1 x	15 min
H_2O dest	3 x	20 sec
Developing solution	1 x	5-10 min†
H_2O dest	3 x	1 min
Stop solution	1 x	15 min
H_2O dest	3 x	5 min
Gel-Dry Drying solution	1 x	15 min*

† Monitor development and change the solution when protein bands are visible and the background is just starting to darken.

* Gels were placed between two cellophane sheets and spanned into a frame for drying over night at RT.

6.2.5 2D-gel electrophoresis

During the 2D-gel-electrophoresis proteins are separated by isoelectric focusing followed by a molecular weight separation. For the first dimension proteins were dissolved in an appropriate buffer (8 M urea, 2% CHAPS, 0.002% Bromophenol Blue, 0.5% ZOOM® Carrier Ampholytes pH 3 -10 non-linear) for isoeletric focusing containing ampholytes. Ampholytes are small, soluble molecules with positive and negative charged groups, which sort based on their isoelectric points in an electric field. Isoelectric focusing was established in a ZOOM® IPG Runner cassette according to manufacturer´s instructions. For the second dimension focused IPG stripe was loaded into a NuPAGE® Bis-Tris Mini gel together with a molecular weight marker (6.2.2) and proteins were separated by molecular weight. The manufacturer`s instructions were followed.

6.2.6 Protein expression and purification

6.2.6.1 Bacterial expression

LMAN1-CRD was blunt end cloned as described in 6.1.7. The gene product included aa 31 – aa 480 of full length LMAN1 protein (see Fig. 3.15). The pET-44 vectors are designed for cloning and high-level expression of peptide sequences fused with the 495 aa NusA-Tag™ protein and encode an additional N-terminal His-Tag. NusA (N utilization substance A) is a fusion protein that itself has a high solubility thus enhancing the solubility of the fusion partner, *LMAN1*-CRD. E. coli lysogenic strain BL21 (DE3) cells were used for plasmid expression and grown in terrific broth medium (TB) for liquid culture. Media contained 100 µg/ml ampicillin. 1 litre of shake-flask culture was incubated at 37°C for constitutive expression of LMAN1-CRD construct. Protein expression was induced by 100 µM IPTG (final concentration) in the exponential growing phase (OD = 0.6 – 0.7) and culture was grown over night. Bacteria were harvested by centrifugation (3000 rpm, 30 min, 4°C; Sorvall) and after cell lysis (50 mM Tris-HCl pH 8.0, 50 mM NaCl, 0.5 mM EDTA, 5% (v/v) glycerol, 1 mM TCEP and Lysonase Bioprocessing Reagent 20 µl/1g cell paste), a series of detergent and buffer washes removed membrane components and contaminating proteins from the IB pellet (see Novagen User Protocol). Purified IBs were denatured with the reducing agent Tris-(2-carboxyethyl)-phosphine HCl (TCEP) and the detergent N-Laurosylsarcosine followed by dialysis (10 mM Tris-HCl, 50 µM EDTA, 100 µM TCEP, 0.06% N-Laurosylsarcosine solution, pH 8.0) and subsequent testing of different refolding conditions using the iFOLD™ Protein Refolding System 1. The iFOLD™ Protein Refolding System 1 provides reagents for inclusion body (IB) purification and denaturation, combined with a plate-based protein refolding buffer matrix. One 96-well plate contains 92 unique refolding solutions (see Novagen user protocol). 92 refolding conditions were checked by a radioactive binding assay with A1AT-^{125}I (6.3.2). Eight refolding conditions showed high binding capacity and were further analysed (Tab 6.1).

***Table 6.1** iFOLD buffer compositions for 8 buffers showing high binding capacity for LMAN1*

iFold buffer 1	**iFold buffer 2**
50 mM Tris pH 7.0 (RT)	50 mM Tris pH 7.0 (RT)
12.5 mM β-Cyclodextrine	100 mM NaCl
1 mM TCEP	1 mM $CaCl_2$
0,5 M Guanidine-Hydrochloride (GuHCl)	1 mM $MgCl_2$
	12.5 mM β-Cyclodextrine
iFold buffer 3	**iFold buffer 4**
50 mM Tris pH 7.5 (RT)	50 mM Tris pH 7.5 (RT)
100 mM NaCl	100 mM NaCl
1 mM TCEP	0.5 M Guanidine-Hydrochloride (GuHCl)
0.1% PEG-6000 (w/v)	
iFold buffer 5	**iFold buffer 6**
50 mM Tris pH 7.5 (RT)	50 mM Tris pH 8.0 (RT)
250 mM NaCl	12.5 mM β-Cyclodextrine
12.5 mM β-Cyclodextrin	0.5 M Guanidine-Hydrochloride (GuHCl)
0.1% PEG-6000 (w/v)	
iFold buffer 7	**iFold buffer 8**
50 mM Tris pH 7.5 (RT)	50 mM Tris pH 7.5 (RT)
100 mM NaCl	100 mM NaCl
12.5 mM β-cyclodextrine	12.5 mM β-cyclodextrine
0.1% PEG-6000 (w/v)	1 mM TCEP
	0.5 M L-Arginin-Hydrochloride (L-Arg)

Previous to refolding, LMAN1-CRD was purified over Ni^{2+}-Sepharose using the N-terminal His-tag of the construct (section 6.2.2.3). N-terminal NusA- and His-tags were afterwards removed by Enterokinase. Digested protein was refolded over night by room temperature using the best refolding conditions. The pre-purified, digested and refolded *LMAN1*-CRD protein was functionally purified using either mannose- or A1AT-sepharose.

6.2.6.2 Mammalian protein expression

293T cells were transiently transfected by the $CaPO_4$-method, gene product was expressed for 48 hours and cells were harvested under native conditions using Triton-X-100 as detergent. Cells were harvested using trypsin and centrifuged for pelleting at 1200 rpm, 10 min at RT. Cells were lysed under native conditions using Triton-X-100 as detergent (cell lysis buffer: 20 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM $CaCl_2$, 1% Triton-X-100), sonicated and ultracentrifuged at 430.000g 4°C for 15 min (UZ, Beckmann). Supernatant was dialysed against the same buffer containing 0.04% Triton-X-100. Dialysed protein lysate (dialysis buffer: 20 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM $CaCl_2$, 0.4% Triton-X-100) was functionally purified over mannose-sepharose.

6.2.6.3 Protein purification with ÄktaPurifier FPLC System

For purification of proteins Ni^{2+}-sepharose self-packed mannose-sepharose or self-packed A1AT-sepharose were used. All protein solutions were purified using the ÄktaPurifier FPLC system.

His-FPLC conditions

Column was equilibrated using 2 column volumes (CV) washing buffer and protein solution was either loaded over the Superloop or the syringe, depending on protein solution volume. A flow rate of 0.5 ml/min was used (maximal pressure: 0.6 MPa). Bound protein was washed with 5 CV washing buffer and eluted with 5 CV elution buffer. Fractions were separated on a 4 – 12% Bis-Tris Gel and analysed by Western blot (6.2.1) using primary antibody directed against LMAN1 or by silver staining (6.2.2).

His-FPLC washing buffer	**His-FPLC elution buffer**
50 mM Tris-HCl pH 7.4	50 mM Tris-HCl pH 7.4
200 mM NaCl	200 mM NaCl
1 mM $CaCl_2$	1 mM $CaCl_2$
20 mM Imidazole	500 mM Imidazole

Mannose- and A1AT-FPLC conditions

Column was equilibrated using 2 CV washing buffer and protein solution was either loaded over the Superloop or the syringe, depending on protein solution volume. A flow rate of 0.1 ml/min was used (maximal pressure: 0.6 MPa). Bound protein was washed with 5 CV washing buffer and eluted with 5 CV elution buffer. Mannose- and A1AT-buffers were sampled for both sepharoses. Fractions were separated on a 4 – 12% Bis-Tris Gel followed by Western blot analysis (6.2.1) using primary antibody directed against LMAN1 or by silver staining (6.2.2).

Mannose FPLC washing buffer	**Mannose FPLC elution buffer**
20 mM Tris pH 7.4 (4°C)	20 mM Tris pH 7.4 (4°C)
150 mM NaCl	150 mM NaCl
1 mM $CaCl_2$	1 mM $CaCl_2$
0.04% Triton-X-100	0.04% Triton-X-100
-	20 mM EGTA
-	100 mM Mannose

A1AT FPLC washing buffer	**A1AT FPLC elution buffer**
10 mM HEPES pH 7.4 (4°C)	10 mM Tris pH 7.4 (4°C)
10 mM MES	150 mM NaCl
150 mM NaCl	0.04% Triton-X-100
1 mM $CaCl_2$	20 mM EGTA
0,04% Triton-X-100	-

6.2.7 Alpha-1-antitrypsin ELISA

Cells were grown in T25-flasks to approximately 80% confluence (corresponding to approximately 7 µg DNA) and conditioned cell culture medium was harvested after 24 h. For the Enzyme-Linked-Immuno-Sorbent-Assay (ELISA) 100 µl undiluted conditioned medium was used. ELISA was established according to manufacturer´s instructions.

6.2.8 ER-Golgi Fractionation

ER-Golgi fraction was isolated using the Endoplasmic reticulum isolation kit. The kit is based on a differential centrifugation step for the isolation of the ER. Complete Post Mitochondrial Fraction (PMF) was established according to manufacturer´s instructions. Pulse Chase Experiments were performed using a defined protein amount of whole PMF fractions.

6.3 Radioactive labelling

6.3.1 Liquid scintillation counting (LSC)

Liquid scintillation counting is a standard technique for measuring radiation (α-, ß and □-radiation). The radiolabelled sample is incorporated into uniform distribution with a liquid chemical medium capable of converting the kinetic energy of nuclear emissions into light energy. Ultima Gold was used as a liquid chemical medium. Ultima Gold is a liquid scintillation cocktail for a wide range of aqueous and non-aqueous samples. Aromatic molecules (scintillators) in LSC cocktails are excited by radiation, recognized by scintillation or light emission respectively. Actually pi-electrons in aromatic molecules are easily excited and are able to convert and transfer energy. By mixing the Ultima Gold solution with the radiolabelled sample the energy of radiation, which is the nuclear decay in form of particles emitted freely, is converted into excitation energy. After transfer of energy to a primary scinitillator molecule, the excited molecule emits light in the range 300 – 400 nm and transits into the ground state energy. Emitted light is absorbed by secondary scintillator molecules and delivered as fluorescent light with a wavelength of 400 nm – 470 nm. Fluorescent light is converted into electrical impulses by photo multipliers (PMT tubes) and subsequent translated into a measurable value (counts per minute CPM). Dpm calculations were based on the tSIE-method.

6.3.2 LMAN1 binding to ^{125}I−A1AT

Radioactive iodine (^{125}I) was incorporated into alpha-1-antitrypsin (A1AT) by electrophilic substitution using Pierce ® Iodination Beads. Tyrosine residues of

human A1AT were radioactively labelled by the chosen method. First ^{125}I was incubated with iodination beads to be in the oxidized form and subsequent positive charged iodine reacts with tyrosines to form 125I- tyrosine. 100 µg pure A1AT protein (in Tris, NaCl pH 7.8; 10 mg/ml each) was incubated with 74MBq ^{125}I and potassium phosphate buffer [Markwell, 1982]. Functional *LMAN1*-CRD binding was verified by the constructs specificity. The N-terminal His-tag of the *LMAN1*-CRD construct is able to bind Ni^{2+}-sepharose whereas functional binding over the LMAN1 CRD-domain should be performed by A1AT-^{125}I (Fig.6.2).

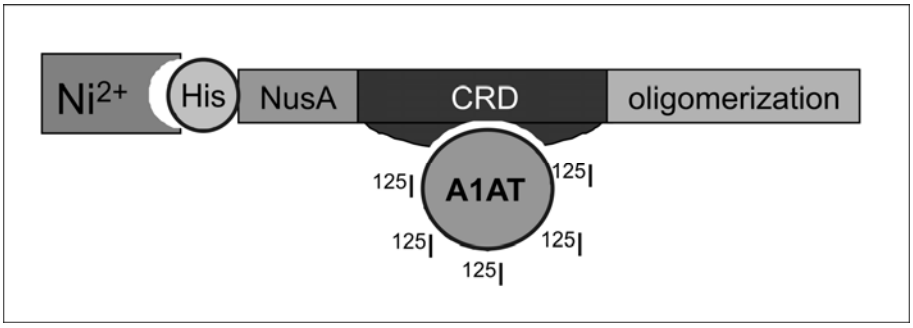


Figure 6.2 Radioactive binding assay for verification of functional LMAN1 binding. In case of LMAN1 CRD binding to its substrate A1AT, which was iodinated by ^{125}I, radioactivity could be measured in the elution fraction. Bound protein was eluted by 1 M imidazole.

25 µl Ni^{2+}-sepharose were incubated with radioactive ^{125}I -A1AT (100.000 CPM) and refolded protein lysate (92 different buffer conditions; iFold System) for one hour at 4°C. After washing steps bound protein was eluted using 1 M imidazole. Radioactivity in the eluted fractionmeasured with UltimaGold in a liquid scintillation counter (LSC TRI-CARB 2900 TR) indicates ^{125}I-A1AT binding. Non-specific binding of ^{125}I-A1AT to the agarose was detected in control experiments applying 1M imidazol as competitive inhibitor of specific binding.

6.3.3 Metabolic labelling with ^{3}H–Xylose

HDC9 cells were transiently transfected with either the empty vector control or *XYLT2*-pcDNA3.1 as described in 6.4.3.1 and 24 hours after transfection cells were fed with RPMI-medium containing 5 µCi ^{3}H-Xylose. After 48 hours incorporation time cells were harvested with 0.9% NaCl and trypsin. Cells were centrifuged at 4000 rpm

for 5 min RT and cell pellet was dissolved in 1 N NaOH for 1 hour at 56°C, sonicated and subsequently neutralized in 2.5 N acetic acid. One part was mixed with 10 ml Ultima Gold and measured in a LSC. For the other part protein concentration was measured using the Lowry method. Results were evaluated as decay per minute for a given protein amount in mg.

6.3.4 Pulse Chase Experiment

Six million parental LoVo cells or *LMAN1*-transfected LoVo cells were sowed in 10-cm dishes with methionine-free medium containing 10% dialysed FCS and 1% glutamine. After 24 hours cells were pulsed with the same medium containing 100µCi/ml ^{35}S-methionine. After 30 min chase time cells were harvested and fractionated using the Endoplasmic Reticulum Isolation Kit (6.2.8). 30 µg isolated ER/Golgi proteins (PMF; between 1.000.000 to 500.000 DPM) were separated by 2-D gel electrophoresis (6.2.5) using a nonlinear pH-gradient from 3 – 10 as first dimension and 4 – 12% polyacryamide gel were used in the second dimension. Gels were fixed in fixation solution 1 (10% (v/v) 96% ethanole, 2.5% (v/v) methanole in H_2O dest) for 1 hour or over night at room temperature. After washing steps in distilled water (H_2O dest) gels were fasten-dried on whatman filter papers using undertow, afterwards dried gels were exposed to a phosphoimager screen for 1 week. During phosphorimaging radioactive material can be localized and quantified on an imaging plate (IP). IPs are coated with barium fluoro-halide phosphor crystals (BaF(Br,I):Eu^{2+}). Europium crystals (Eu^{2+}) within the IP stored emitted radiation of ^{35}S-proteins on 2-D-gels in form of trapped electrons with an absorption band at about 600 nm. Exposed IP was scanned with a He-Ne (633 nm) laser within the FLA-3000 Phosphorimager, released electrons were recombined with Eu^{2+} crystals releasing photons at 400 nm. The blue light (400 nm) emitted was collected by Phosphorimager software producing a digitized image (FLA3000 settings: Sample Mode IP; gradiation 256 (8bit); resolution 100 pixel). Digitized images were evaluated using the Proteomweaver Software for 2-D gel analysis (2004 Version 3.0.0 beta; Definiens AG, Munich, Germany). Proteomweaver software helps in precise verifying specific spots on 2-D gel images.

6.4 Cell Culture experiments

6.4.1 Human cancer cell lines

Most of the MSI-H CRC cell lines were either obtained from ATCC (Manassas, VA, USA) or the German Cancer Research Center (DKFZ, Heidelberg, Germany). Cell line Vaco6 was a kind gift of J.K.V. Wilson, Cleveland. The MSI-H CRC cell line K073A was established in our laboratory. MSI-H status of all colorectal tumor tissues used in the present study (n = 170) has been determined previously using the National Cancer Institute/ICGHNPCC reference marker panel including *BAT25* and *BAT26* [Boland et al., 1998] and one additional mononucleotide marker *CAT25* [Findeisen et al., 2005].

6.4.2 Cell culture

All cell lines were grown under standard conditions in RPMI 1640 supplemented with 10% FCS, 100 U/mL penicillin, and 0.1 mg/mL streptomycin in a humified incubator with 5% CO_2 at 37°C. Depending on the cell growth, cells were subcultured twice a week at a ratio of 1:5 to 1:20, in order to keep the confluence of the cells between approximately 50% and 80%. Mycoplasma contamination of cultured cells was excluded by PCR. For cryopreservation, tumor cells were harvested and frozen at -80°C. Typically, 1 x 10^7 to 3 x 10^7 cells were stored in cryo-vials with 1.5 ml of 10% DMSO in FCS.

6.4.3 Transfection methods

6.4.3.1 Electroporation

Tumor cells (n = 10^7) were transiently transfected by electroporation using the Amaxa Nucleofector 1, program T-20 with solution V and 5 µg DNA (program and solution dependent on cell line, see User manual). Because of a high mortality rate during electroporation medium was changed after 24 hours. Cells were grown in small plastic flasks (T25) and pelleted cells or conditioned medium were harvested after 24 to 48 h. This method was used for transient transfection of Colo60H and HDC9 cells.

6.4.3.2 Lipofection

One day before transfection, 0.5 to 2 x 10^5 cells (LoVo cells 2 x 10^5) were plated on 24-well plate in 500 µl of growth medium without antibiotics so that cells had a confluency of 90 – 95% at the time of transfection. For each transfection sample complexes were prepared as follows: DNA was diluted in 50 µl DMEM medium without serum in Eppendorf tubes. Lipofectamine™ 2000 was gently mixed before usage, then the appropriate amount (ratio DNA : Lipofectamin 1 : 2.5) was diluted in 50 µl DMEM and incubated for 5 minutes at RT in polystyrole tubes. Afterwards, diluted DNA was combined with diluted Lipofectamine™ 2000 (total volume = 100 µl). Complex formation was incubated for 10 min at RT. 100 µl of complexes were pipetted drop-wise to each well containing cells and medium. Cells were incubated at 37°C for 18 – 48 hours prior to testing for transgene expression. Medium was changed after 4 – 6 hours.

For generating stable cells, cells were selected with G418 (neomycin resistance) 24 – 48 hours after transfection. Appropriate concentration of G418 was tested prior to transfection. For LoVo cells the selection was performed using 400 µg/ml G418. Treatment with G418 was continued for a minimum of 14 days. After the selection period, single clones were picked by scraping with a sterile pipet tip and transferred on 48-well plates. Alternatively, the cells were harvested by trypsination and limiting dilution of the bulk culture was performed for the selection of single cell clones on 96-well plates.

6.4.3.3 $CaPO_4$ transfection

293T cells were cultured in ten 10 cm dishes to a confluence of approximately 80%. For each transfection 1 ml 1 x HBS buffer, 10 µg plasmid DNA and 50 µl 2.5 M $CaCl_2$ were used. *LMAN1*-pcDNA3.1 was mixed with HBS buffer in a polystyrole tube, $CaCl_2$ was added to the bottom of the tube and mixed by rotating the tube until disappearing of lines. Complex formation (precipitate) was incubated for 5 min. Calcium phosphate forms an insoluble precipitate with DNA, which attaches to the cell surface and is taken up by the cells via endocytosis [Jordan et al., 1996]. Formed Calcium phosphate/DNA precipitate was incubated on 293T cells for 10 min at RT. Subsequent culture medium was added and cells were incubated for 48 h in 37°C with 5% CO_2, after 24 h medium was changed.

6.4.4 Cell Proliferation Assay

100 µl 5 x 10^4 cells/ml were sowed into a 96-well plate. Parental LoVo cells and mock-transfected as well as LMAN1-transfected LoVo cells were included. After one, two, and four days cell viability and growth was determined by adding 20 µl MTS/formazan. The MTS tetrazolium compound (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) is bio-reduced by cells into a colored formazan product that is soluble in tissue culture medium. The quantity of formazan product as measured by the absorbance at 490 nm is directly proportional to the number of living cells in culture. MTS/formazan substrate was incubated in the cell medium for exactly one hour at 37°C and 5% CO_2. Absorbance was measured at 485 nm in an ELISA-Reader. Cells were measured in triplets and mean values were charted against time in days.

6.4.5 Lectin-FACS analysis

For cell surface glycan profiling a fluorescence-activated cell sorter (FACS) was used. FACS analysis distinguishes cells based on size and granularity as well as fluorescence. A laser beam is directed on a stream of fluid containing stained cells for analysis. By cells passing the beam, the beam light is scattered and fluorescent chemicals are excited. Cells can be determined by cell volume (FCS Forward

scatter), granularity (SSC Side Scatter) and by fluorescence (f.e. filters for GFP and Phycoerithrin PE). Quantitative determination of carbohydrate-dependent lectin binding using streptavidin / R-phycoerythrin as a fluorescent marker was performed by flow cytofluorometry in a FACSCalibur instrument using the CellQuest pro software (Becton-Dickinson). Solutions with 4 x 10^5 cells per sample were carefully washed with PBS and incubated with biotinylated lectins for 30 min at 4°C. After washing steps in 1 x PBS/0.1% BSA cells were incubated with streptavidin / R-phycoerythrin (1 : 40). LoVo cells without staining and LoVo cells incubated with streptavidin / R-phycoerythrin were used for adjusting the parameter settings of FACSCalibur. Mean fluorescence intensity was normalized to control values. Statistical analysis was performed on four independent experiments using Student's t-test (independent by groups; grouping variable 0=mock-transfected LoVo cells; 1=LMAN1-transfected LoVo cells) calculated with the program STATISTICA (v7.0 Stat Soft.Inc; Tulsa. OK, USA).

6.5 Human tissues

Human tissues were obtained from the local tissue bank established within the German Collaborative Group on HNPCC. Informed consent was obtained from all patients and the study protocol was approved by the local Ethics Committee. Human tissues were formalin-fixed and paraffin-embedded (FFPE) and sections of 2 µm for immunohistochemistry and 5 µm for microdissection were used.

6.5.1 Immunohistochemistry

For immunohistochemical staining, 2-µm FFPE sample sections were deparaffinized and rehydrated. For antigen retrieval, slides were boiled in 10 mmol/L citrate buffer (pH 6) for 3 x 5 min and endogenous peroxidase activity was blocked with hydrogen peroxide (0.6%) for 20 min. Slides were rinsed with deionized water and then washed in PBS/0.1% Tween 20 (PBS-T) for 5 min. The slides were then stained by 2-h incubation at room temperature with anti-LMAN1 primary antibody (1:200) in 1% horse serum/PBS-T. A biotinylated anti-mouse/anti-rabbit antibody in 1% horse serum/PBS-T was used as secondary antibody for 30 min at room temperature and staining procedure was subsequently amplified by the A/B-complex (Vectastain Elite

ABC kit) for 30 min at room temperature. Immunoreactions were visualized with 3,3-diaminobenzidine, followed by counterstaining with hematoxylin. Tumor-infiltrating lymphocytes, muscle cells, and normal colonic epithelia served as positive controls.

6.5.2 Microdissection of Hemalaun- and Eosin- stained tissues

For morphological assessment of FFPE samples and separate microdissection of normal and tumor tissue, slides were stained with hemalaun and Eosin (H&E staining), resulting in blue staining of acidic structures such as cell nuclei and red staining of eosinophilic structures such as the cytoplasm. Regional microdissection was performed to isolate normal or tumor tissue. Whole tumor areas were microdissected using the H&E staining, but for a protein-specific microdissection, tumor tissues were stained with an appropriate antibody and these brown stained areas were microdissected. First, normal control tissue was isolated by use of a 0.9 x 40 mm needle pre-wetted with ATL buffer (DNeasy Blood & Tissue kit, Qiagen) and under a stereomicroscope. In the second step, tumor tissue was removed from the slide using a fresh needle. The removed tissue was collected in 180 µl of ATL buffer in 1.5 ml reaction tubes. 12 µl of Protease K were added to each tube and the samples were incubated at 56°C on a thermomixer over night. On the next day genomic DNA was isolated according to the manufacturer`s instructions.

6.5 Database analysis

Three databases and four filters were used for candidate gene selection (Fig. 3.1A). Information about ER/Golgi resident proteins was retrieved from the LOCATE database (http://locate.imb.uq.edu.au/), whose XML source was downloaded (version human_v3_20070620) and prepared for local usage as a MySQL database. Protein subcellular localization in LOCATE v3 was defined by 463 GO terms, of which 30 GO terms were related to ER/Golgi subcellular localization (Filter 1). From this candidate subset, the MNR_ensembl database (http://www.seltarbase.org/?topic= MNR_ensembl; version 45.36 g) allowed further specification of candidate genes by restriction of cMNR lengths to a minimum of seven repeat units (Filter 2). Subsequent exclusion of all cMNRs previously investigated revealed a list of 431 candidate genes (Filter 3; SelTarbase, 4th release 200707). Final selection by annotation of genes

encoding proteins of the cellular glycosylation/deglycosylation and glyoprotein transport system led to 28 genes that were used for subsequent analyses (Filter 4). Automated primer design was performed by a perl script using primer3_core (Primer3 version 0.1; http://primer3.sourceforge.net/) in combination with a self-constructed human mispriming repeat library (see Tab. 5.5.1). The bioinformatics-based approach is outlined schematically in Fig. 3.1.A.

7. References

1. Akizuki M, Fukutomi T, Takasugi M, Takahashi S, Sato T et al. (2007) Prognostic significance of immunoreactive neutrophil elastase in human breast cancer: long-term follow-up results in 313 patients. Neoplasia 9:260-264

2. Allen A, Hutton DA, and Pearson JP (1998) The MUC2 gene product: a human intestinal mucin. Int J Biochem Cell Biol 30:797-801

3. Allgayer H, Babic R, Grutzner KU, Beyer BC, Tarabichi A et al. (1998) Tumor-associated proteases and inhibitors in gastric cancer: analysis of prognostic impact and individual risk protease patterns. Clin Exp Metastasis 16:62-73

4. Anelli T, Ceppi S, Bergamelli L, Cortini M, Masciarelli S et al. (2007) Sequential steps and checkpoints in the early exocytic compartment during secretory IgM biogenesis. EMBO J 26:4177-4188

5. Appenzeller C, Andersson H, Kappeler F, and Hauri HP (1999) The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. Nat Cell Biol 1:330-334

6. Appenzeller-Herzog C, Roche AC, Nufer O, and Hauri HP (2004) pH-induced conversion of the transport lectin ERGIC-53 triggers glycoprotein release. J Biol Chem 279:12943-12950

7. Apweiler R, Hermjakob H, and Sharon N (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta 1473:4-8

8. Aychek T, Miller K, Sagi-Assif O, Levy-Nissenbaum O, Israeli-Amit M et al. (2008) E-selectin regulates gene expression in metastatic colorectal carcinoma cells and enhances HMGB1 release. Int J Cancer 123:1741-1750

9. Bakker H, Oka T, Ashikov A, Yadav A, Berger M et al. (2009) Functional UDP-xylose transport across the endoplasmic reticulum/Golgi membrane in a Chinese hamster ovary cell mutant defective in UDP-xylose Synthase. J Biol Chem 284:2576-2583

10. Barresi R and Campbell KP (2006) Dystroglycan: from biosynthesis to pathogenesis of human disease. J Cell Sci 119:199-207

11. Bengtsson E, Neame PJ, Heinegard D, and Sommarin Y (1995) The primary structure of a basic leucine-rich repeat protein, PRELP, found in connective tissues. J Biol Chem 270:25639-25644

12. Bernfield M, Kokenyesi R, Kato M, Hinkes MT, Spring J et al. (1992) Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. Annu Rev Cell Biol 8:365-393

13. Bogenrieder T and Herlyn M (2003) Axis of evil: molecular mechanisms of cancer metastasis. Oncogene 22:6524-6536

14. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR et al. (1998) A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 58:5248-5257

15. Bonin-Debs AL, Boche I, Gille H, and Brinkmann U (2004) Development of secreted proteins as biotherapeutic agents. Expert Opin Biol Ther 4:551-558

16. Brockhausen I. (2007) Glycobiology, Chapter: Glycosyltranferases specific for the synthesis of mucin-type O-glycans. Scion Publishing, Bloxham (Eds.: Sansom C, Markman O):217-234

17. Brockhausen I (1999) Pathways of O-glycan biosynthesis in cancer cells. Biochim Biophys Acta 1473:67-95

18. Brockhausen I (2006) Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions. EMBO Rep 7:599-604

19. Buckowitz A, Knaebel HP, Benner A, Blaker H, Gebert J et al. (2005) Microsatellite instability in colorectal cancer is associated with local lymphocyte infiltration and low frequency of distant metastases. Br J Cancer 92:1746-1753

20. Carethers JM, Chauhan DP, Fink D, Nebel S, Bresalier RS et al. (1999) Mismatch repair proficiency and in vitro response to 5-fluorouracil. Gastroenterology 117:123-131

21. Carlson JA, Rogers BB, Sifers RN, Hawkins HK, Finegold MJ et al. (1988) Multiple tissues express alpha 1-antitrypsin in transgenic mice and man. J Clin Invest 82:26-36

22. Corfield AP. (2007) Glycobiology, Chapter: Glycobiology of mucins in the human gastrointestinal tract. Scion Publishing, Bloxham (Eds.: Sansom C, Markman O):248-260

23. Cowan DA, Gay D, Bieler BM, Zhao H, Yoshino A et al. (2002) Characterization of mouse tGolgin-1 (golgin-245/trans-golgi p230/256 kD golgin) and its upregulation during oligodendrocyte development. DNA Cell Biol 21:505-517

24. Cuellar K, Chuong H, Hubbell SM, and Hinsdale ME (2007) Biosynthesis of chondroitin and heparan sulfate in chinese hamster ovary cells depends on xylosyltransferase II. J Biol Chem 282:5195-5200

25. Day AJ (1999) The structure and regulation of hyaluronan-binding proteins. Biochem Soc Trans 27:115-121

26. Demetriou M, Granovsky M, Quaggin S, and Dennis JW (2001) Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation. Nature 409:733-739

27. Demetriou M, Nabi IR, Coppolino M, Dedhar S, and Dennis JW (1995) Reduced contact-inhibition and substratum adhesion in epithelial cells expressing GlcNAc-transferase V. J Cell Biol 130:383-392

28. Dennis JW, Granovsky M, and Warren CE (1999) Glycoprotein glycosylation and cancer progression. Biochim Biophys Acta 1473:21-34

29. Dolcetti R, Viel A, Doglioni C, Russo A, Guidoboni M et al. (1999) High prevalence of activated intraepithelial cytotoxic T lymphocytes and increased neoplastic cell apoptosis in colorectal carcinomas with microsatellite instability. Am J Pathol 154:1805-1813

30. Dube DH and Bertozzi CR (2005) Glycans in cancer and inflammation--potential for therapeutics and diagnostics. Nat Rev Drug Discov 4:477-488

31. Duval A and Hamelin R (2002) Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. Cancer Res 62:2447-2454

32. Dwek MV and Brooks SA (2004) Harnessing changes in cellular glycosylation in new cancer treatment strategies. Curr Cancer Drug Targets 4:425-442

33. El-Bchiri J, Buhard O, Penard-Lacronique V, Thomas G, Hamelin R et al. (2005) Differential nonsense mediated decay of mutated mRNAs in mismatch repair deficient colorectal cancers. Hum Mol Genet 14:2435-2442

34. El-Bchiri J, Guilloux A, Dartigues P, Loire E, Mercier D et al. (2008) Nonsense-mediated mRNA decay impacts MSI-driven carcinogenesis and anti-tumor immunity in colorectal cancers. PLoS One 3:e2583-

35. Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5:435-445

36. Esko JD, Stewart TE, and Taylor WH (1985) Animal cell mutants defective in glycosaminoglycan biosynthesis. Proc Natl Acad Sci U S A 82:3197-3201

37. Fearon ER and Vogelstein B (1990) A genetic model for colorectal tumorigenesis. Cell 61:759-767

38. Findeisen P, Kloor M, Merx S, Sutter C, Woerner SM et al. (2005) T25 repeat in the 3' untranslated region of the CASP2 gene: a sensitive and specific marker for microsatellite instability in colorectal cancer. Cancer Res 65:8072-8078

39. Fraldi A, Zito E, Annunziata F, Lombardi A, Cozzolino M et al. (2008) Multistep, sequential control of the trafficking and function of the multiple sulfatase deficiency gene product, SUMF1 by PDI, ERGIC-53 and ERp44. Hum Mol Genet 17:2610-2621

40. Fukushima K, Satoh T, Baba S, and Yamashita K (2010) alpha1,2-Fucosylated and beta-N-acetylgalactosaminylated prostate-specific antigen as an efficient marker of prostatic cancer. Glycobiology 20:452-460

41. Furlan D, Casati B, Cerutti R, Facco C, Terracciano L et al. (2002) Genetic progression in sporadic endometrial and gastrointestinal cancers with high microsatellite instability. J Pathol 197:603-609

42. Fuster MM and Esko JD (2005) The sweet and sour of cancer: glycans as novel therapeutic targets. Nat Rev Cancer 5:526-542

43. Gabius H-J (2009) The sugar code: The fundamentals of Glycosciences. Wiley VCH, Weinheim, 1st edition.

44. Geboes K, Rutgeerts P, Vantrappen G, and Desmet VJ (1983) Immunoreactivity of alpha-1-antitrypsin in the human colon. Hepatogastroenterology 30:24-26

45. Goetz JG (2009) Bidirectional control of the inner dynamics of focal adhesions promotes cell migration. Cell Adh Migr 3:185-190

46. Gotoh M, Sato T, Kiyohara K, Kameyama A, Kikuchi N et al. (2004) Molecular cloning and characterization of beta1,4-N-acetylgalactosaminyltransferases IV synthesizing N,N'-diacetyllactosediamine. FEBS Lett 562:134-140

47. Götting C, Kuhn J, Brinkmann T, and Kleesiek K (2002) Xylosyltransferase activity in seminal plasma of infertile men. Clin Chim Acta 317:199-202

48. Götting C, Kuhn J, and Kleesiek K (2007) Human xylosyltransferases in health and disease. Cell Mol Life Sci 64:1498-1517

49. Götting C, Kuhn J, Sollberg S, Huerkamp C, Brinkmann T et al. (2000) Elevated serum xylosyltransferase activity correlates with a high level of hyaluronate in patients with systemic sclerosis. Acta Derm Venereol 80:60-61

50. Götting C, Sollberg S, Kuhn J, Weilke C, Huerkamp C et al. (1999) Serum xylosyltransferase: a new biochemical marker of the sclerotic process in systemic sclerosis. J Invest Dermatol 112:919-924

51. Greenson JK, Bonner JD, Ben-Yzhak O, Cohen HI, Miselevich I et al. (2003) Phenotype of microsatellite unstable colorectal carcinomas: Well-differentiated and focally mucinous tumors and the absence of dirty necrosis correlate with microsatellite instability. Am J Surg Pathol 27:563-570

52. Hampel H, Stephens JA, Pukkala E, Sankila R, Aaltonen LA et al. (2005) Cancer risk in hereditary nonpolyposis colorectal cancer syndrome: later age of onset. Gastroenterology 129:415-421

53. Hart GW, Housley MP, and Slawson C (2007) Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins. Nature 446:1017-1022

54. Hascall VC and Heinegard D (1974) Aggregation of cartilage proteoglycans. I. The role of hyaluronic acid. J Biol Chem 249:4232-4241

55. Hauri HP, Kappeler F, Andersson H, and Appenzeller C (2000) ERGIC-53 and traffic in the secretory pathway. J Cell Sci 113 (Pt 4):587-596

56. Hebert DN, Garman SC, and Molinari M (2005) The glycan code of the endoplasmic reticulum: asparagine-linked carbohydrates as protein maturation and quality-control tags. Trends Cell Biol 15:364-370

57. Heutinck KM, Ten B, I, Hack CE, Hamann J, and Rowshani AT (2010) Serine proteases of the human immune system in health and disease. Mol Immunol 47:1943-1955

58. Hollingsworth MA and Swanson BJ (2004) Mucins in cancer: protection and control of the cell surface. Nat Rev Cancer 4:45-60

59. Huang H, Campbell SC, Nelius T, Bedford DF, Veliceasa D et al. (2004) Alpha1-antitrypsin inhibits angiogenesis and tumor growth. Int J Cancer 112:1042-1048

60. Hurtley SM and Helenius A (1989) Protein oligomerization in the endoplasmic reticulum. Annu Rev Cell Biol 5:277-307

61. Ionov Y, Nowak N, Perucho M, Markowitz S, and Cowell JK (2004) Manipulation of nonsense mediated decay identifies gene mutations in colon cancer Cells with microsatellite instability. Oncogene 23:639-645

62. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, and Perucho M (1993) Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature 363:558-561

63. Iozzo RV (1998) Matrix proteoglycans: from molecular design to cellular function. Annu Rev Biochem 67:609-652

64. Iozzo RV (1999) The biology of the small leucine-rich proteoglycans. Functional network of interactive proteins. J Biol Chem 274:18843-18846

65. Iozzo RV (2005) Basement membrane proteoglycans: from cellar to ceiling. Nat Rev Mol Cell Biol 6:646-656

66. Iozzo RV and Cohen I (1993) Altered proteoglycan gene expression and the tumor stroma. Experientia 49:447-455

67. Iozzo RV and Murdoch AD (1996) Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. FASEB J 10:598-614

68. Itzkowitz SH, Bloom EJ, Kokal WA, Modin G, Hakomori S et al. (1990) Sialosyl-Tn. A novel mucin antigen associated with prognosis in colorectal cancer patients. Cancer 66:1960-1966

69. Jass JR (1998) Diagnosis of hereditary non-polyposis colorectal cancer. Histopathology 32:491-497

70. Jiang X and Couchman JR (2003) Perlecan and tumor angiogenesis. J Histochem Cytochem 51:1393-1410

71. Jordan M, Schallhorn A, and Wurm FM (1996) Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. Nucleic Acids Res 24:596-601

72. Kalsheker NA (2009) alpha1-Antitrypsin deficiency: best clinical practice. J Clin Pathol 62:865-869

73. Kamiya Y, Kamiya D, Yamamoto K, Nyfeler B, Hauri HP et al. (2008) Molecular basis of sugar recognition by the human L-type lectins ERGIC-53, VIPL, and VIP36. J Biol Chem 283:1857-1861

74. Kane MF, Loda M, Gaida GM, Lipman J, Mishra R et al. (1997) Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res 57:808-811

75. Kim H, Nam SW, Rhee H, Shan LL, Ju KH et al. (2004) Different gene expression profiles between microsatellite instability-high and microsatellite stable colorectal carcinomas. Oncogene 23:6218-6225

76. Kim IJ, Kang HC, Park JH, Shin Y, Ku JL et al. (2003) Development and applications of a beta-catenin oligonucleotide microarray: beta-catenin mutations are dominantly found in the proximal colon cancers with microsatellite instability. Clin Cancer Res 9:2920-2925

77. Kinsella MG, Bressler SL, and Wight TN (2004) The regulated synthesis of versican, decorin, and biglycan: extracellular matrix proteoglycans that influence cellular phenotype. Crit Rev Eukaryot Gene Expr 14:203-234

78. Kloor M, von Knebel Doeberitz M, and Gebert JF (2005) Molecular testing for microsatellite instability and its value in tumor characterization. Expert Rev Mol Diagn 5:599-611

79. Korkmaz B, Moreau T, and Gauthier F (2008) Neutrophil elastase, proteinase 3 and cathepsin G: physicochemical properties, activity and physiopathological functions. Biochimie 90:227-242

80. Kufe DW (2009) Mucins in cancer: function, prognosis and therapy. Nat Rev Cancer 9:874-885

81. Lau KS, Partridge EA, Grigorian A, Silvescu CI, Reinhold VN et al. (2007) Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. Cell 129:123-134

82. Leavitt RD, Felsted RL, and Bachur NR (1977) Biological and biochemical properties of Phaseolus vulgaris isolectins. J Biol Chem 252:2961-2966

83. Lee MJ, Lee HS, Kim WH, Choi Y, and Yang M (2003) Expression of mucins and cytokeratins in primary carcinomas of the digestive system. Mod Pathol 16:403-410

84. Lengauer C, Kinzler KW, and Vogelstein B (1997) Genetic instability in colorectal cancers. Nature 386:623-627

85. Lesley J, Hyman R, English N, Catterall JB, and Turner GA (1997) CD44 in inflammation and metastasis. Glycoconj J 14:611-622

86. Li YC, Korol AB, Fahima T, and Nevo E (2004) Microsatellites within genes: structure, function, and evolution. Mol Biol Evol 21:991-1007

87. Liefhebber JM, Punt S, Spaan WJ, and van Leeuwen HC (2010) The human collagen beta(1-O)galactosyltransferase, GLT25D1, is a soluble endoplasmic reticulum localized protein. BMC Cell Biol 11:33-

88. Lievens PM, De SB, Garofalo S, Lunstrum GP, Horton WA et al. (2008) Transient dimerization and interaction with ERGIC-53 occur in the fibroblast growth factor receptor 3 early secretory pathway. Int J Biochem Cell Biol 40:2649-2659

89. Linnebacher M, Gebert J, Rudy W, Woerner S, Yuan YP et al. (2001) Frameshift peptide-derived T-cell epitopes: a source of novel tumor-specific antigens. Int J Cancer 93:6-11

90. Liu Y, Choudhury P, Cabral CM, and Sifers RN (1999) Oligosaccharide modification in the early secretory pathway directs the selection of a misfolded glycoprotein for degradation by the proteasome. J Biol Chem 274:5861-5867

91. Lowe JB and Marth JD (2003) A genetic approach to Mammalian glycan function. Annu Rev Biochem 72:643-691

92. Lowry OH, Rosebrough NJ, Farr AL, and Randall RJ (1951) Protein measurement with the Folin phenol reagent. J Biol Chem 193:265-275

93. Lugli A, Zlobec I, Baker K, Minoo P, Tornillo L et al. (2007) Prognostic significance of mucins in colorectal cancer with different DNA mismatch-repair status. J Clin Pathol 60:534-539

94. Luther KB and Haltiwanger RS (2009) Role of unusual O-glycans in intercellular signaling. Int J Biochem Cell Biol 41:1011-1024

95. Lynch HT and de la Chapelle (1999) Genetic susceptibility to non-polyposis colorectal cancer. J Med Genet 36:801-818

96. Lynch JP and Hoops TC (2002) The genetic pathogenesis of colorectal cancer. Hematol Oncol Clin North Am 16:775-810

97. Ma B, Simala-Grant JL, and Taylor DE (2006) Fucosylation in prokaryotes and eukaryotes. Glycobiology 16:158R-184R

98. Mainardi CL, Dixit SN, and Kang AH (1980) Degradation of type IV (basement membrane) collagen by a proteinase isolated from human polymorphonuclear leukocyte granules. J Biol Chem 255:5435-5441

99. Markowitz S, Wang J, Myeroff L, Parsons R, Sun L et al. (1995) Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. Science 268:1336-1338

100. Markwell MA (1982) A new solid-state reagent to iodinate proteins. I. Conditions for the efficient labeling of antiserum. Anal Biochem 125:427-432

101. Matsushima N, Ohyanagi T, Tanaka T, and Kretsinger RH (2000) Super-motifs and evolution of tandem leucine-rich repeats within the small proteoglycans--biglycan, decorin, lumican, fibromodulin, PRELP, keratocan, osteoadherin, epiphycan, and osteoglycin. Proteins 38:210-225

102. McDonald JA and Kelley DG (1980) Degradation of fibronectin by human leukocyte elastase. Release of biologically active fragments. J Biol Chem 255:8848-8858

103. McGivern A, Wynter CV, Whitehall VL, Kambara T, Spring KJ et al. (2004) Promoter hypermethylation frequency and BRAF mutations distinguish hereditary non-polyposis colon cancer from sporadic MSI-H colon cancer. Fam Cancer 3:101-107

104. McGowan SE and Thompson RJ (1989) Extracellular matrix proteoglycan degradation by human alveolar macrophages and neutrophils. J Appl Physiol 66:400-409

105. McKenzie E, Tyson K, Stamps A, Smith P, Turner P et al. (2000) Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member. Biochem Biophys Res Commun 276:1170-1177

106. Mendelsohn R, Cheung P, Berger L, Partridge E, Lau K et al. (2007) Complex N-glycan and metabolic control in tumor cells. Cancer Res 67:9771-9780

107. Monsigny M, Roche AC, Sene C, Maget-Dana R, and Delmotte F (1980) Sugar-lectin interactions: how does wheat-germ agglutinin bind sialoglycoconjugates? Eur J Biochem 104:147-153

108. Moussalli M, Pipe SW, Hauri HP, Nichols WC, Ginsburg D et al. (1999) Mannose-dependent endoplasmic reticulum (ER)-Golgi intermediate compartment-53-mediated ER to Golgi trafficking of coagulation factors V and VIII. J Biol Chem 274:32539-32542

109. Neve EP, Lahtinen U, and Pettersson RF (2005) Oligomerization and interacellular localization of the glycoprotein receptor ERGIC-53 is independent of disulfide bonds. J Mol Biol 354:556-568

110. Novaretti MC, Domingues AE, Manhani R, Pinto EM, Dorlhiac-Llacer PE et al. (2008) ABO genotyping in leukemia patients reveals new ABO variant alleles. Genet Mol Res 7:87-94

111. Nufer O, Kappeler F, Guldbrandsen S, and Hauri HP (2003) ER export of ERGIC-53 is controlled by cooperation of targeting determinants in all three of its domains. J Cell Sci 116:4429-4440

112. Nyfeler B, Reiterer V, Wendeler MW, Stefan E, Zhang B et al. (2008) Identification of ERGIC-53 as an intracellular transport receptor of alpha1-antitrypsin. J Cell Biol 180:705-712

113. Nyfeler B, Zhang B, Ginsburg D, Kaufman RJ, and Hauri HP (2006) Cargo selectivity of the ERGIC-53/MCFD2 transport receptor complex. Traffic 7:1473-1481

114. Ohtsubo K and Marth JD (2006) Glycosylation in cellular mechanisms of health and disease. Cell 126:855-867

115. Parish CR, Freeman C, and Hulett MD (2001) Heparanase: a key enzyme involved in cell invasion. Biochim Biophys Acta 1471:M99-108

116. Parkin DM, Bray F, Ferlay J, and Pisani P (2005) Global cancer statistics, 2002. CA Cancer J Clin 55:74-108

117. Pastrello C, Santarosa M, Fornasarig M, Sigon R, Perin T et al. (2005) MUC gene abnormalities in sporadic and hereditary mucinous colon cancers with microsatellite instability. Dis Markers 21:121-126

118. Patsos G, Andre S, Roeckel N, Gromes R, Gebert J et al. (2009) Compensation of loss of protein function in microsatellite-unstable colon cancer cells (HCT116): a gene-dependent effect on the cell surface glycan profile. Glycobiology 19:726-734

119. Potapenko IO, Haakensen VD, Luders T, Helland A, Bukholm I et al. (2010) Glycan gene expression signatures in normal and malignant breast tissue; possible role in diagnosis and progression. Mol Oncol 4:98-118

120. Przybylo M, Pochec E, Link-Lenczowski P, and Litynska A (2008) Beta1-6 branching of cell surface glycoproteins may contribute to uveal melanoma progression by up-regulating cell motility. Mol Vis 14:625-636

121. Rajagopalan H, Bardelli A, Lengauer C, Kinzler KW, Vogelstein B et al. (2002) Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status. Nature 418:934-

122. Rao RM, Betz TV, Lamont DJ, Kim MB, Shaw SK et al. (2004) Elastase release by transmigrating neutrophils deactivates endothelial-bound SDF-1alpha and attenuates subsequent T lymphocyte transendothelial migration. J Exp Med 200:713-724

123. Roch C, Kuhn J, Kleesiek K, and Götting C (2010) Differences in gene expression of human xylosyltransferases and determination of acceptor specificities for various proteoglycans. Biochem Biophys Res Commun 391:685-691

124. Roeckel N, Woerner SM, Kloor M, Yuan YP, Patsos G et al. (2009) High frequency of LMAN1 abnormalities in colorectal tumors with microsatellite instability. Cancer Res 69:292-299

125. Samowitz WS, Holden JA, Curtin K, Edwards SL, Walker AR et al. (2001) Inverse relationship between microsatellite instability and K-ras and p53 gene alterations in colon cancer. Am J Pathol 158:1517-1524

126. Scamuffa N, Siegfried G, Bontemps Y, Ma L, Basak A et al. (2008) Selective inhibition of proprotein convertases represses the metastatic potential of human colorectal tumor cells. J Clin Invest 118:352-363

127. Schwitalle Y, Kloor M, Eiermann S, Linnebacher M, Kienle P et al. (2008) Immune response against frameshift-induced neopeptides in HNPCC patients and healthy HNPCC mutation carriers. Gastroenterology 134:988-997

128. Seelentag WK, Li WP, Schmitz SF, Metzger U, Aeberhard P et al. (1998) Prognostic value of beta1,6-branched oligosaccharides in human colorectal carcinoma. Cancer Res 58:5559-5564

129. Shah SN, Hile SE, and Eckert KA (2010) Defective mismatch repair, microsatellite mutation bias, and variability in clinical cancer phenotypes. Cancer Res 70:431-435

130. Smyrk TC, Watson P, Kaul K, and Lynch HT (2001) Tumor-infiltrating lymphocytes are a marker for microsatellite instability in colorectal carcinoma. Cancer 91:2417-2422

131. Söreide K, Janssen EA, Soiland H, Korner H, and Baak JP (2006) Microsatellite instability in colorectal cancer. Br J Surg 93:395-406

132. Spicer AP, Joo A, and Bowling RA, Jr. (2003) A hyaluronan binding link protein gene family whose members are physically linked adjacent to chondroitin sulfate proteoglycan core protein genes: the missing links. J Biol Chem 278:21083-21091

133. Spreafico M and Peyvandi F (2008) Combined FV and FVIII deficiency. Haemophilia 14:1201-1208

134. Springer GF (1984) T and Tn, general carcinoma autoantigens. Science 224:1198-1206

135. Stephens DJ and Pepperkok R (2001) Illuminating the secretory pathway: when do we need vesicles? J Cell Sci 114:1053-1059

136. Takenaka Y, Fukumori T, and Raz A (2004) Galectin-3 and metastasis. Glycoconj J 19:543-549

137. Tang BL, Wang Y, Ong YS, and Hong W (2005) COPII and exit from the endoplasmic reticulum. Biochim Biophys Acta 1744:293-303

138. Tateno H, Nakamura-Tsuruta S, and Hirabayashi J (2009) Comparative analysis of core-fucose-binding lectins from Lens culinaris and Pisum sativum using frontal affinity chromatography. Glycobiology 19:527-536

139. Ten Hagen KG, Fritz TA, and Tabak LA (2003) All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. Glycobiology 13:1R-16R

140. Thibodeau SN, Bren G, and Schaid D (1993) Microsatellite instability in cancer of the proximal colon. Science 260:816-819

141. Trombetta ES and Helenius A (1998) Lectins as chaperones in glycoprotein folding. Curr Opin Struct Biol 8:587-592

142. Umar A (2004) Lynch syndrome (HNPCC) and microsatellite instability. Dis Markers 20:179-180

143. Umar A and Srivastava S (2004) The promise of biomarkers in colorectal cancer detection. Dis Markers 20:87-96

144. Van den Steen P., Rudd PM, Dwek RA, and Opdenakker G (1998) Concepts and principles of O-linked glycosylation. Crit Rev Biochem Mol Biol 33:151-208

145. Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P et al. (2009) Glycobiology Chapter: Glycans in Glycoprotein Quality Control. Cold Spring Harbor Laboratory Press, New York, 2nd edition. Figure 2.

146. Vlad AM and Finn OJ (2004) Glycoprotein tumor antigens for immunotherapy of breast cancer. Breast Dis 20:73-79

147. Vlodavsky I and Friedmann Y (2001) Molecular properties and involvement of heparanase in cancer metastasis and angiogenesis. J Clin Invest 108:341-347

148. Vollenweider F, Kappeler F, Itin C, and Hauri HP (1998) Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. J Cell Biol 142:377-389

149. Weitz J, Koch M, Debus J, Hohler T, Galle PR et al. (2005) Colorectal cancer. Lancet 365:153-165

150. Woerner SM, Benner A, Sutter C, Schiller M, Yuan YP et al. (2003) Pathogenesis of DNA repair-deficient cancers: a statistical meta-analysis of putative Real Common Target genes. Oncogene 22:2226-2235

151. Woerner SM, Gebert J, Yuan YP, Sutter C, Ridder R et al. (2001) Systematic identification of genes with coding microsatellites mutated in DNA mismatch repair-deficient cancer cells. Int J Cancer 93:12-19

152. Woerner SM, Yuan YP, Benner A, Korff S, von Knebel Doeberitz M et al. (2010) SelTarbase, a database of human mononucleotide-microsatellite mutations and their potential impact to tumorigenesis and immunology. Nucleic Acids Res 38:D682-D689

153. Wolpin BM, Kraft P, Gross M, Helzlsouer K, Bueno-de-Mesquita HB et al. (2010) Pancreatic cancer risk and ABO blood group alleles: results from the pancreatic cancer cohort consortium. Cancer Res 70:1015-1023

154. Yang P, Cunningham JM, Halling KC, Lesnick TG, Burgart LJ et al. (2000) Higher risk of mismatch repair-deficient colorectal cancer in alpha(1)-antitrypsin deficiency carriers and cigarette smokers. Mol Genet Metab 71:639-645

155. Yang WH, Kim JE, Nam HW, Ju JW, Kim HS et al. (2006) Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. Nat Cell Biol 8:1074-1083

156. You KT, Li LS, Kim NG, Kang HJ, Koh KH et al. (2007) Selective translational repression of truncated proteins from frameshift mutation-derived mRNAs in tumors. PLoS Biol 5:e109-

157. Young J, Simms LA, Biden KG, Wynter C, Whitehall V et al. (2001) Features of colorectal cancers with high-level microsatellite instability occurring in familial and sporadic settings: parallel pathways of tumorigenesis. Am J Pathol 159:2107-2116

158. Zcharia E, Metzger S, Chajek-Shaul T, Aingorn H, Elkin M et al. (2004) Transgenic expression of mammalian heparanase uncovers physiological functions of heparan sulfate in tissue morphogenesis, vascularization, and feeding behavior. FASEB J 18:252-263

159. Zhang YC, Zhou Y, Yang CZ, and Xiong DS (2009) A review of ERGIC-53: its structure, functions, regulation and relations with diseases. Histol Histopathol 24:1193-1204

8. Appendix

8.1 Abbrevations

APC	Adenomatous Polyposis Coli
Asp	Asparagine
B4GALT1	Beta-1,4-Galactosyltransferase
BAX	Bcl-2 Associated X-Protein
BRAF	V-raf murine sarcoma viral oncogene homolog B1
cMNR	Coding Mononucleotide Repeat Mutations
cMS	Coding Microsatellite
CR	Calreticulin
CRC	Colorectal Cancer
CS	Chondroitin Sulfate
CX	Calnexin
DS	Dermatan Sulfate
ECM	Extracellular Matrix
ER	Endoplasmic Reticulum
ERAD	ER associated degradation
FGF2	Fibroblast Growth Factor Receptor 2
FGFR3	Fibroblast Growth Factor Receptor 3
Fuc	Fucose
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	N-Acetylgalactosamine
GALNT5	N-Acetylgalactosaminyltransferase 5
Glc	Glucose
GlcA	Glucoronic Acid
GlcNAc	N-Acetylglucosamine
HNPCC	Hereditary Non-Polyposis Colon Cancer
HS	Heparan Sulfate
IdoA	Iduronic Acid
KS	Keratan Sulfate
LDS	Lithium Dodecyl Sulfate
Man	Mannose
MAPK	mitogen-activated protein kinase

MCFD2	Multiple Coagulation Factor Deficiency Protein 2
MLH1	MutL homolog 1
MMR	Mismatch Repair
MSH2	MutS homolog 2
MSH6	MutS homolog 6
MSI	Microsatellite Instability
NMD	Nonsense-mediated mRNA decay
NE	Neutrophil Elastase
NeuAc	N-Acetyl-Neuraminic Acid (Sialic Acid)
PDGF	Platelet Derived Growth Factor
PDGFR	Platelet Derived Growth Factor Receptor
PG	Proteoglycan
PLR3	Phosphatase Of Regenerating Liver-3
PMF	Post-Mitochondrial Fraction
PMS2	
PMS	Post-Meiotic Segregation
Ser	Serine
SIAT4B	Sialyltransferase 4B
SUMF1	Sulfatase Modifying Factor 1
TGFBR	Transforming Growth Factor Beta Receptor
TGFß	Transforming Growth Factor
Thr	Threonin
VEGF	Vascular Endothelial Growth Factor
Xyl	Xylose
XYLT	Xylosyltransferase

8.2 Own publications, presentations and posters

Original Papers

Roeckel N, Woerner SM, Kloor M, Yuan YP, Patsos G, Gromes R, Kopitz J, Gebert J. (2009) High frequency of LMAN1 abnormalities in colorectal tumors with microsatellite instability. Cancer Res 69:292-299.

Patsos G, André S, Roeckel N, Gromes R, Gebert J, Kopitz J, Gabius HJ. Compensation of loss of protein function in microsatellite-unstable colon cancer cells (HCT116): a gene-dependent effect on the cell surface glycan profile. (2009) Glycobiology. 19:726-734

Roeckel N, Weber A, Gromes R, Himmelsbach S, Kopitz J, Gebert J. (2010). XYLT2, a glycosyltransferase as a target in MSI-H tumorigenesis. In preperation.

Poster presentation

Roeckel N, Woerner SM, Kloor M, Patsos G, Gromes R, Karl M, Kopitz J, Gebert J. Mutational inactivation of glycoprotein transporter (LMAN1) in MSI colorectal adenomas and carcinomas. International society for gastrointestinal hereditary tumours (InSiGHT), poster presentation, 24. – 27.06.2009, Düssedorf.

Oral presentation

Roeckel N, Woerner SM, Kloor M, Yuan YP, Patsos G, Gromes R, Karl M, Kopitz J, Gebert J. Loss of LMAN1 glycoprotein carrier function in microsatellite unstable colorectal tumors. American Association of Cancer Research (AACR) Annual Meeting, minisymposium session, oral presentation, 18.- 22.04.2009, Denver, Colorado, USA.

9. Acknowledgements

First of all, I would like to thank my supervisors Prof. Dr. rer. nat. Jürgen Kopitz and Dr. rer. nat. Johannes Gebert for their scientific and practical support. Whenever there was need they had time to answer my questions and they provided valuable guidance in all aspects of work and further more. They helped me by improving my critical scientific point of view, gave me the opportunity to participate in international congresses and furthered my career in all other directions.

I would like to thank Prof. Dr. med. Magnus von Knebel Doeberitz, who gave me the possibility to work in his group and always supported me.

I would like to thank Prof. Dr. rer. nat. Hans-Joachim Gabius (Faculty of Veterinary Medicine of the Ludwig-Maximilians-University, Munich) who provided us with lectins for lectin-FACS analysis and gave scientific advice. Dr.-Ing. Niels Grabe and Thora Pommerenke (Hamamatsu Tissue Imaging and Analysis Center, Heidelberg) I want to thank for digitalization of IHC stainings.

I am grateful to all members of the Department of Applied Tumor Biology, for their kind reception in the lab. Everbody was always open to give advice in scientific and practical areas during the lab working day. I am grateful to Dr. med. Matthias Kloor and Dr. med. Stefan Woerner who introduced me in frameshift mutation analysis and gave me support in evaluation of immunohistochemistry. Dr. rer. nat. Roland Gromes and Dr. med. Matthias Kloor I would like to thank for reading of the manuscript and for helpful comments. Together with Dr. med. Miriam Reuschenbach, Dr. sc. hum. Sara Michel and Kathrin Bauer they made poor lab days more sunny, always providing a good sense of humor. A special thank goes to Marcel Karl and Sigrun Himmelsbach, my two main team players during lab life, who always understood to motivate and encourage me. It was a great pleasure to be part of this team.

The deepest gratitude is dedicated to my parents, who always supported me in what I`ve made up my mind concerning my career. I am also indebted to my friends for their understanding, especially I want to name Kerstin, Marcel and Judith.