Immunology of a cancer mucin:

Proteasomal processing and presentation of human MUC1 tandem repeat glycopeptides in the MHC class I pathway

INAUGURAL-DISSERTATION
zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
Tanja Ninković
aus Belgrad, Serbien
2006

Referees/Berichterstatter Prof. Dr. Franz-Georg Hanisch
Prof. Dr. Dietmar Schomburg

Date of oral examination/ 26.10.2006
Tag der mündlichen Prüfung

The present research work was carried out under the supervision and the direction of Prof. Dr. Franz-Georg Hanisch in the Institute of Biochemistry II, Medical Faculty, University of Cologne, Cologne, Germany, from February 2003 to October 2006.
Diese Arbeit wurde von Februar 2003 bis Oktober 2006 am Institut für Biochemie II der Medizinischen Fakultät der Universität zu Köln unter der Leitung und der Betreuung von Prof. Dr. Franz-Georg Hanisch durchgeführt.

To my parents

Silent gratitude isn't much use to anyone.

~ Gladys Browyn Stern

Therefore,

I would like to thank to my PhD supervisor Prof Hanisch for excellent mentorship and advices, endless patience and for teaching me to think critically. I appreciate very much that he gave me freedom to develop in my own pace and to find my own ways and interests. His support for independent thinking and crisp analysis led me to the path of being a scientist. Above all, I am happy that I had a supervisor that true all the years of his experience managed to keep passion for science and flexible and open intellect.

Also, I would cordially like to thank my Prof Dimitrijević, supervisor during my master studies who made me fell in love in cancer research and to my biology teacher Jelisavka Luković who "infected" me with the "biology virus". My success I also owe to Nasta Dedović and Nikola Tanić who taught me how important understandment of experimental techniques can be.

The great gratitude I owe to Mirko, my soul mate, who shared with me sunshine and fog of scientific work and always took my hand to help me come to the solution and provided calm space to think.

Many, many thanks to my dear parents and my darling brother Ivan who supported all my decisions and helped me deal with all problems and challenges. They taught me to trust myself and all my self-confidence I owe to them.

Here I would also like to thank my friends in "exile", Dragana, Mirjana, Natasa, Marija, Aneta, Marko and Hilal who made my stay in Germany easier. And many thanks to my Belgrade friends who always wait for me at home with a smile.

Many thanks to my colleagues in the lab and institute for exchanging knowledge and offering help, especially to Hanieh, Katja, Marcus, Regine, Isabelle, Sebastian, Tilo, Yeliz, Uschi and Nefres.

Also, I would like to thank Prof Paulsson for offering me the possibility to do my PhD at the excellently organized Institute for Biochemistry and many thanks to Frau Pelzer and people from the Service lab, with the special gratitude to Stefan Müller and Udo Roth.

This work would also not have been possible without help of Aleksej Popov and Kondo Eisei from the Molecular Tumor Biology and Immunology and kind cooperation of their supervisor Prof Joachim Schultze.

ABREVIATIONS

MUC1	mucine 1
RNA	ribonucleic acid
VNTR	variable number of tandem repeats
GalNAc	N-acetylgalactosamine
GlcNac	N-acetylglucosamine
MHC	Major histocompatibility complex
HLA-A,B,C	human leukocyte antigen
$H2K^b$	mouse MHC class I allele
CTL	cytotoxic T lymphocytes
IgG, IgM	Immunoglobuline G, Immunoglobuline M
APC	antigen presenting cells
DC	dendritic cells
IPs	immunoproteasomes
INF γ	Interferon gamma
IL-4	Interleukin-4
TNF-α	Tumor Necrosis Factor-alpha
GM-CSF	Granulocyte-Macrophage Colony Stimulating Factor
ER	Endoplasmatic Reticulum
ERAD	Endoplasmatic Reticulum Associated Degradation
Tg mice	transgenic mice
HPLC	high pressure liquid chromatography

Units of Measure and Prefixes

°C	degree Celsius
Da	Dalton
g	gram
h	hour
L	litre
m	meter
min	minute
s	sec
V	volt
MW	molecular weight

Symbol Prefix (Factor)
k kilo (103)
c centi (10-2)
m milli (10-3)
μ micro (10-6)
n nano (10-9)
p pico (10-12)

1 INTRODUCTION

1.1 Breast cancer

Breast cancer is the primary cause of death of women in the western countries, accounting 1 out of 10 women. Just over 1 million of people per year get diagnosed with breast cancer and around 400.000 die. Despite a 60% survival rate 1100 persons die daily due to breast cancer (Parkin DM et al.,2001).

The small size of most cancer nodes (below two centimetres) results in late detection, and localisation in the proximate neighbourhood of lymph nodes increases the metastatic probability. The best prognosis for breast cancer patients is still dependent on an early detection. Although many efforts have been made to develop new diagnostic and therapeutic tools, surgery with chemotherapy and radiation are still the methods of choice for breast cancer treatment.

Breast cancer is actually a common name for different malignant incidences which can be classified in ductal, lobular, medular, tubular and mucinous carcinoma or as inflammatory breast cancer, Paget's disease of the nipple and phyllodes tumor (Berg JW et al., 1995).

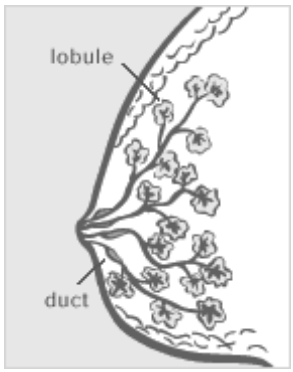

Figure 1: Brest anatomy showing ducti and lobule. From the web site http://www.breastbiopsy.com/bioresults_typesofbreastcancer2.jsp

Ductal carcinoma originates in the milk ductus and after conveying metastatic potential, it can spread through the body. Around 80% of all breast cancers are invasive ductal carcinoma.

Lobular carcinoma originates in the milk producing glands. In invasive form it spreads to surrounding fatty tissue and can metastasize. 10-15% of all breast cancers are of invasive lobular type.

Other types of breast cancer are rare and represent between 1 and 5% of all breast cancer cases.

A large number of tumor markers in the sera of patients indicate breast cancer. These include MUC1 (e.g. CA15-3), CEA, HER-2oncogene, p53 tumor-suppressor, glycolytic enzymes and cytokeratins. Compared to healthy individuals and patients with benign breast diseases, patients with breast cancer have elevated levels of breast antigens circulating in their plasma and/or serum.

1.2 MUC1 and mucins

Mucins are high-molecular weight glycoproteins, representing the main constituents of mucus and the glycocalix of epithelial and some non-epithelial cells. Common functions of mucins are lubrication and protection. The family of mucin proteins consists of up to now 19 defined proteins (Entrez-Protein database, Table 1), MUC1-MUC20. Their MW of 100.000 till 1.000.000 is by 50-90% contributed by carbohydrates.

Mucin-Name	**S**ecretory **M**embraneous	Chromosomal localisation	Expressing tissue	Literature
Mucin 1	m,s	1q21-q24	Ubiquitous	Gendler *et al.*, 1990
Mucin 2	s	11p15	GI tract, respiratory tract	Gum *et al.*, 1994
Mucin 3	m,s	7q22	GI tract	van Klinken *et al.*, 1997
Mucin 4	m,s	3q29	GI tract	Nollet *et al.*, 1998
Mucin 5AC	s	11p15	Respiratory tract	Guyonnet Duperat *et al.*, 1995
Mucin 5B	s	11p15	Respiratory tract, Stomach	Meerzaman *et al.*, 1994
Mucin 6	s	11p15.5-15.4	Bile duct	Toribara *et al.*, 1993
Mucin 7	s	4q13-q21	GI tract	Bobek *et al.*, 1993
Mucin 8	/	12q24.3	Saliva gland	Shankar *et al.*, 1994
Mucin 9	/	1p13	Trachea	Lapensee *et al.*, 1997
Mucin 11	/	7q22	Oviduct	Williams *et al.*, 1999
Mucin 12	m,s	7q22	GI tract, Uro-genital tract	Williams *et al.*, 1999
Mucin 13	m	3q13.3	GI tract, Uro-genital tract	Williams et al., 2001
Mucin 15	/	11p14.3	GI tract	Pallesen et al., 2002
Mucin 16	m	19p13.3	Genital tract, GI tract, lymph organs	Yin et al., 2001
Mucin 17	m	7q22	Ovaries	Gum et al., 2002
Mucin 19	/	12q12	Trachea, Saliva gland	Chen et al., 2003
Mucin 20	/	3q29	Kidneys, Placenta, Lungs, Prostate, Liver	Higuchi et al., 2004

Table 1. The list of currently identified mucins, with chromosomal localisation, tissue expression and the mark about secretory or transmembrane form of the protein

The currently known MUC species can be subdivided into two groups based on their structural aspects and biosynthetic routes: membrane bound and secretory mucins.

Membrane bound mucins (MUC1, MUC3, MUC4, MUC12) are essential components of the glycocalix. Their hydrophobic transmembrane domains are anchoring the protein within the lipid bilayer with C-terminus oriented toward the cytosol (Table 1). Most of the membrane bound mucins show tissue-specific expression patterns (Hanisch FG et al., 2000).

Secretory mucins are the main constituents of the visco-elastic mucus gel and are mainly coded by a gene cluster on chromosome 11p15.5. All secretory mucins (MUC2, MUC5AC, MUC5B, MUC6) except MUC7 possess one or several von Willebrandt factor-like D domains (Hanisch FG et al., 2000), cysteine-rich peptides, which function in the oligomerization of mucin monomers and in the packaging into secretory vesicles (Perez-Vilar J et al., 1999). Other mucin genes are randomly scattered over the genome.

The special characteristic of all mucins is the presence of a "tandem repeats domain", several tens till hundreds of repeats of a peptide (Fowler J et al., 2001). The number of tandem repeats varies due to the diversity of mRNA length from 0,4 to 20kb transcribed from the single gene. Tandem repeats are heavily O-glycosylated by long glycan chains (up to 20 monosaccharides per chain) on numerous serines and threonines. At terminal position negatively charged sugars, like sialic acid or sulphated sugars, are often found. The sugar chain composition and length are tissue specific. Because of the high proline content, the complete molecule is extended and protruding from the cell surface (Figure 2).

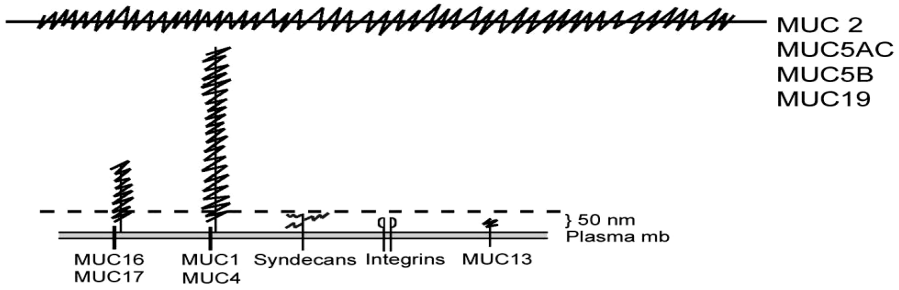

Figure 2. Model of mucins and other cell surface glycoproteins/proteoglycans. The figure schematically portrays structures of four size classes of mucins as well as consensus structures for syndecans and integrins. All extracellular portions of these molecules are roughly drawn to scale. Blue squiggles represent O-linked oligosaccharides while red squiggles represent glycosaminoglycans. Integrins, syndecans and most other surface receptors do not extend beyond 50 nm from the cell surface; however, with the exception of MUC13, mucins extend much further due to the extended structures contributed by the proline-rich heavily O-glycosylated tandem repeat domains. MUC1 and MUC4 are the largest transmembrane mucins, extending >200 nm from the cell surface. The ectodomain structures of MUC16 and MUC17 are considerably shorter than those of MUC1 and MUC4, but still much larger than those of other surface glycoproteins. Soluble mucins, such as MUC2, 5AC, 5B and 19 are even larger, reading 500–100 nm in length (Brayman M et al., 2004).

The functions of mucins are mainly influenced by sugar presence. The major common function of mucins is protection of the cellular surface of mechanical injuries and micro-organisms, prevention of dehydration as well as lubrication, especially in the gastrointestinal, respiratory or reproductive tracts (Perez-Vilar et al., 1999). These functions are contributed mainly by secretory forms of mucins, which reside close to the apical surface of epithelia cells in the form of oligomers connected by disulfide bridges. Membrane bound mucins are as monomers attached to the cell membrane, where beside protective functions they influence cell-cell interactions. The long, protruding structure of mucins (Bramwell ME et al.,1986) introduces anti-adhesive effects and prevents cell-cell and cell-extracellular matrix contacts (Ligtenberg MJ et al., 1992), thereby taking part in epithelial morphogenesis (Hilkens J et al., 1992), migration, proliferation of cells, carcinogenesis and metastasis. Specific interactions with E-selectine and integrins involve mucins in inflammation, development and lymphocytes homing (Walz et al., 1990; Yuen CT et al.,1992; Schimizu Y et al.,1993).

1.2.1 MUC1

The first mucin characterized during the 80ies was MUC1, a large glycoprotein existing in both secretory and transmembrane forms. It was described under numerous names: Polymorphic epithelial mucin, PEM, PEMT, Episialin, Tumor-associated mucin, Carcinoma-associated mucin, Tumor-associated epithelial membrane antigen, EMA, H23AG, Peanut-reactive urinary mucin, PUM, Breast carcinoma-associated antigen DF3, CD227 antigen. Finally it was systematically designated as MUC1 (Figure 3).

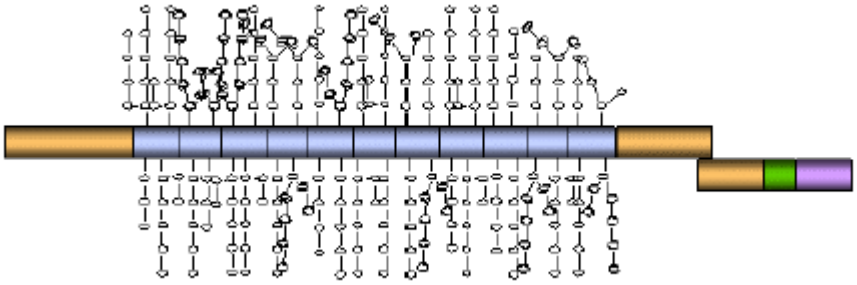

Figure 3. The structure of MUC1. The 72 amino acids cytosolic domain is shown in purple. The 28 amino acids domain is shown in green. The VNTR region is shown in blue while the rest of the protein is shown in tan. The extensive O-linked glycosylation in the VNTR region is indicated with the stick figures.

1.2.1.1 MUC1 localisation

MUC1 exhibits a ubiquitous organ distribution (Fig. 4), (Zotter *et al.*, 1988) including mammary gland acini and ducts, salivary gland ducts and serous acini, but not mucinous acini; squamous epithelium of the esophagus; parietal cells, canaliculi and peptic cells of the stomach; acini and ducts of the pancreas; bile ducts in the liver; enterocytes of the duodenum, but not the large intestine; respiratory and ciliated epithelium of the lungs, serous bronchial glands, but not mucinous ones; distal tubules of the kidney and collecting ducts, but not proximal tubules; bladder urothelium; prostate gland epithelium; resting endometrium of the uterus; rete testis; (activated) mesothelium. Important negative tissues are the skin epithelium and all kinds of mesenchymal tissues. Beside epithelial tissues, MUC1 was found on lymphoid cells and lymphomas, especially on plasma cells and myelomas (Zotter *et al.*, 1988; Treon *et al.*, 1999) as well as on erythroblasts. MUC1 presence is moreover reported for activated T cell and mature dendritic cells, but not on resting T cells and immature dendritic cells.(Clossen S et al.,2004; Agrawal *et al.*, 1998)

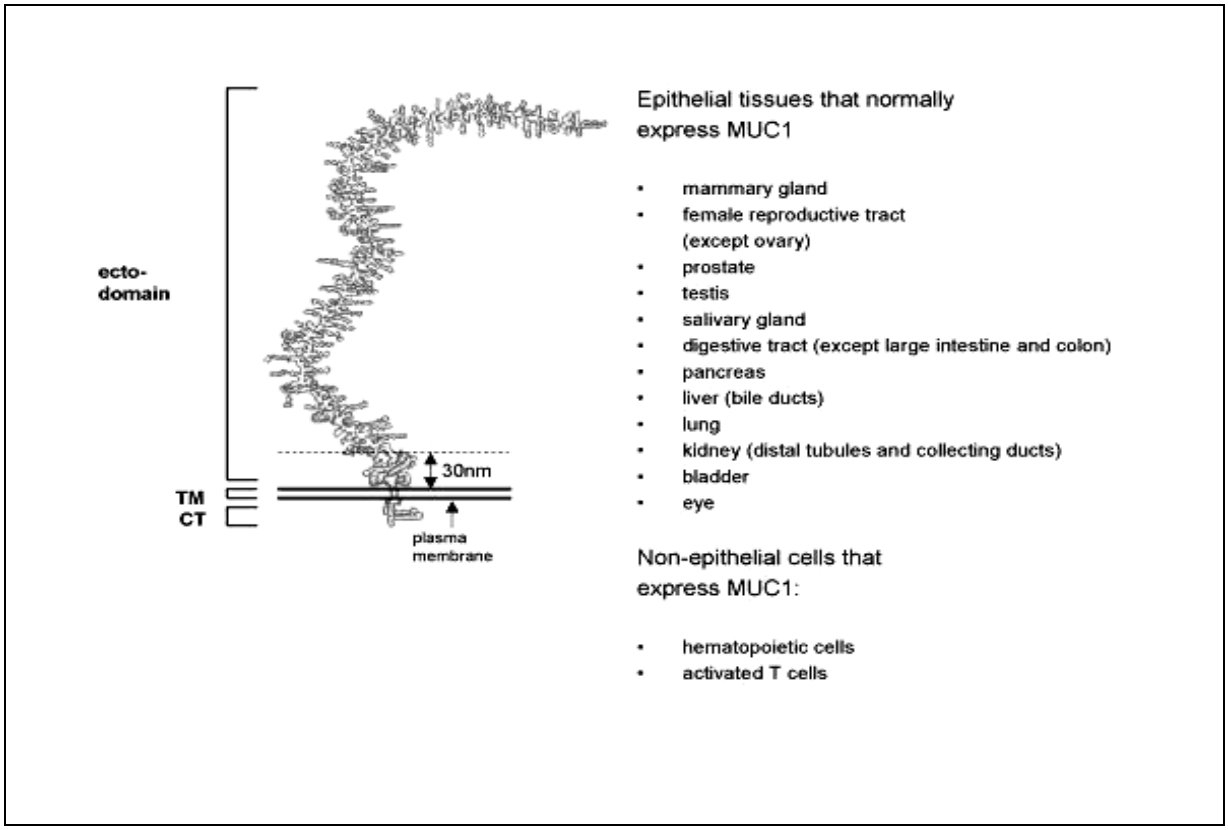

Figure 4. Structure of MUC1 in the transmembrane form and the list of MUC1 positive tissues

1.2.1.2 MUC1 Gene

MUC1 is coded by a gene localised on the long arm of chromosome 1 in the 1q21-q24 region. The gene is composed of seven exons that together comprise 4,2 to 7,0 kb of genomic DNA. The alternative splicing of mRNA gives rise to transcripts of 3.7 to 6.4 kb coding for the different isoforms of the MUC1 protein, transmembrane, secretory, and MUC1 X/Y/Z isoforms.

1.2.1.3 MUC1 protein structure

The alternative splicing results in the generation of several isoforms of MUC1 protein, and therefore we speak of MUC1 polymorphism. Several forms of MUC1 are formed:

1. MUC1/REP: heterodimeric complex of a large tandem repeat ectodomain component and a small transmembrane component (Ligtenberg MJ et al., JBC 1992)
2. MUC1/SEC that contains tandem repeat but no transmembrane domain (Smorodinsky N et al., 1996)
3. Short MUC1/X/Y/Z membrane anchored isoforms generated by alternative splicing and devoid of tandem repeats. (Zrihan-Licht S et al., 1994; Baruch A et al., 1997; Oosterkamp HM et al., 1997)

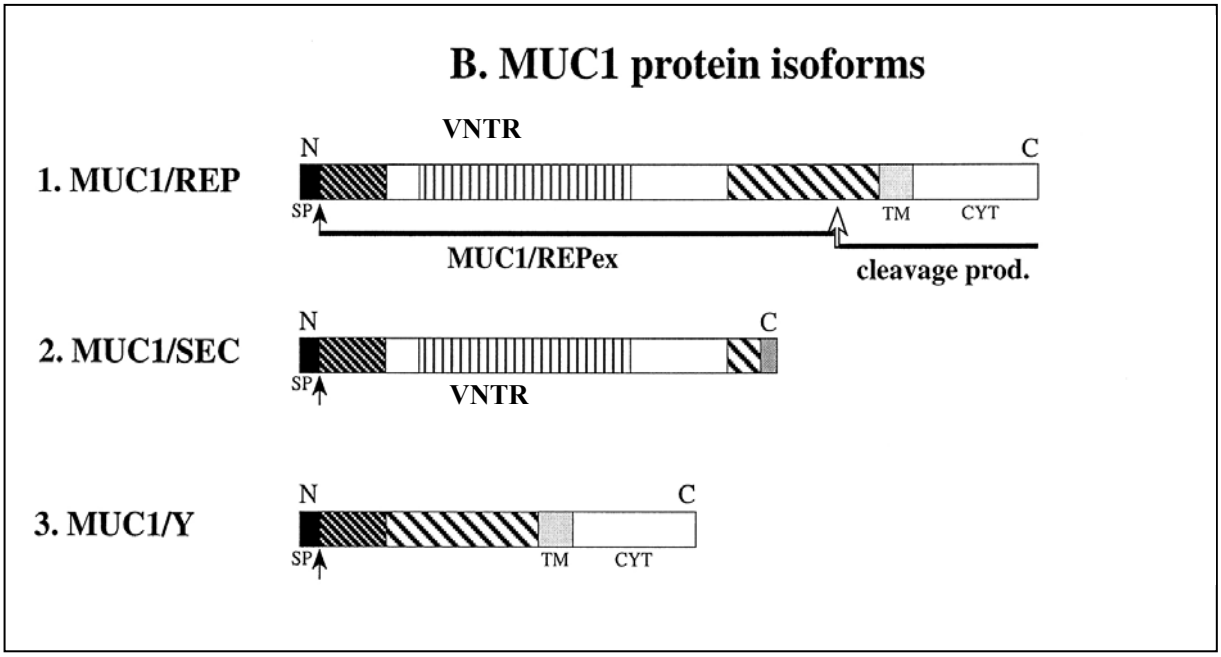

Figure 5. MUC1 protein isoforms. Domains re-occurring in various MUC1 isoforms are marked with the pattern. N and C represent amino and carboxy termini of the protein, TM – transmembrane domain; CYT-cytoplasmic domain, SP – signal peptide. The proteolytic cleavage site in the MUC1/REP protein is indicated by the open arrow. Cleavage divides the protein into the large tandem repeat array (VNTR)-containing extracellular domain and the smaller proteolytic cleavage product, which contains the cytoplasmic domain.

The best characterized *MUC1* gene product is a long membrane associated protein, named also a MUC1/REP. It is a type 1 transmembrane mucin-like protein (Fig. 5) consisting of two sub-domains: large extracellular domain and short cytoplasmic tail. After a translation of MUC1 mRNA, the MUC1/REP protein is autolytically cleaved in the endoplasmatic reticulum. This cleavage generates two peptide chains that get noncovalently re-associated in the proximity of the cleavage site (Ligtenberg MJL et al., 1992). The larger of the cleavage products is the ectodomain component, mainly built by the variable number of tandem repeats (VNTR), flanked by short degenerate sequences and non-repetitive sequences at both sides of the protein (Wreschner DH et al., 1990; Ligtenberg MJL et al., 1990; Gendler SJ et al., 1990; Abe M et al., 1993). The VNTR region contains 20-125 repeats of the 20-meric amino acid sequence PAPGSTAPPAHGVTSAPDTR repeat which represents 50-80% of the peptide mass (reviewed in Gendler SJ et al., 1995; Lagow E et al., 1999). MUC1 protein is highly polymorphic in the population with respect to the number of tandem repeats as well as the presence of sequence variations within the tandem repeats (major being DT>ES). Shorter cleavage product carries the cytoplasmic and transmembrane domain.

The MUC1/REP extracellular domain can be shed from the cell by a yet not completely understood mechanism and found in the sera of breast cancer patients at increased levels (Baruch A et al., 1999). During its biosynthesis, the MUC1/REP protein is extensively modified by O- and N-glycosylation. O-glycosylation takes mainly place within tandem repeats of the VNTR region. Further on, MUC1/REP proteins can be subdivided into four variants A-D, which arise from the single nucleotide polymorphisms.

The MUC1/SEC protein contains full length sequence corresponding to the extracellular domain of MUC1/REP, but lacks the membrane anchoring hydrophobic domain as well as the intracellular domains. (Fig. 5)

In contrast to MUC1/SEC, short MUC1/X/Y/Z splice isoforms contains a transmembrane and cytoplasmic domains identical to the MUC1/REP, but their extracellular part contains neither the repeat array domain nor its flanking region and, hence, are devoid of the hallmark mucin-like features (Fig. 5). (Baruch A et al., 1999)

Due to the abundant glycosylation of the VNTR region with frequent negatively charged terminal sugars and because of the high number of proline residues, MUC1 has an extended, rigid form protruding 200-500 nm from the cell surface into the lumen. (Fig.2 Brayman et al., 2004).

1.2.1.4 MUC1 functions

Beside functions common to all glycocalix-building proteins, like lubrication, hydration, and protection from microbes, MUC1 fulfils several other more complex roles. So is MUC1 reported to take part in signal transduction, binding of actin cytoskeleton and it influences cell-cell and cell-matrix contacts and metastases.

It has been shown that MUC1 mucin undergoes phosphorylation on both tyrosine and serine residues of the cytosolic domain and, as a consequence, can potentially bind second messenger or adaptor proteins such as Grb2 (Wreschner DH et al., 1994, Zrihan-Licht S et al., 1994). Kufe D has reported the participation of MUC1 in the Grb2/Sos/ras signalling pathway, whereas interaction of MUC1 with erbB1 during lactation was reported by Gendler SJ (Pandey P et al., 1995; Schroeder JA et al., 2001). Moreover, high levels of MUC1 in mammary gland promote MAPK (ERK1/2) signal transduction (Schreder JA et al., 2001). Also, the cytoplasmic domain of the MUC1 protein interacts directly with ß-catenin and therefore with the actin cytoskeleton (Yamamoto M et al., 1997). Besides this, the secretory MUC1 can interact with MUC1/Y and trigger Ser/Thr phosphorylation of its cytoplasmic part which involves MUC1 in signal transduction (Baruch A et al., 1999). These results suggest that MUC1 protein participates in signal transduction and may well play an active role in the oncogenic process.

The tandem repeat domain is reported to interact with ICAM-1 of endothelial and epithelial cells (Hayashi T. et al., 2001), and therefore to participate in cell migration and metastasis. MUC1 has also influence on the cell-cell and cell-ECM contacts by sterically introducing anti-adhesive effects (Hilkens J et al., 1992, Wesseling J et al., 1995) or promoting adhesion by specific interaction of carbohydrate moieties with adhesion molecules like selectines (Zhangh K et al., 1996).

1.2.1.4.1 MUC1 glycosylation

MUC1 is post-translationally modified by both N- and O- type of glycosylation.

MUC1 has five potential N-glycosylation sites, three of which are located in the C terminal transmembrane component of MUC1, and two proximate to the VNTR domain. However, the major type of MUC1 glycosylation is O-glycosylation on numerous Ser and Thr residues within the VNTR region. Each 20-meric tandem repeat contains five possible glycosylation sites (Gendler SJ et al., 1995; Hanisch FG et al., 2000), two serines and three threonines (Fig.6).

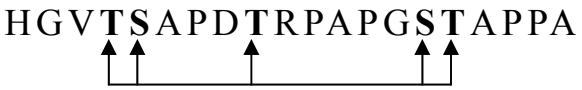

Figure 6. MUC1 tandem repeat amino acid sequence with label potential O-glycosylation sites

The average glycosylation density and structure vary *in vivo* and are cell and tissue specific. O-glycosylation is performed by the ordered action of glycosyltransferases adding carbohydrates to the protein backbone. The reaction initiates with addition of N-acetylgalactosamine on side chains of hydroxyl-amino acids Ser or Thr by one isoform of the 20 different polypeptide N-acetylgalactosaminyltransferases. The product, peptide linked GalNAc, is named Tn antigen and presents a starting point for the generation of more complex glycan chains. The elongation of the carbohydrate chain is performed by one of the numerous glycosyltransferases isoforms (Varki A). Based on the structure of two or three sugars attached to the Tn antigen eight basic *core* structures can be defined. Core structures are further elongated by sialylation, acetylation, sulphatation and (poly)lactosamine extension. The membrane attached protein appears as incompletely glycosylated form on the membrane and is repetitively re-internalised for additional glycosylations probably in the trans Golgi network. The mature form of the protein can remain attached to the membrane or become shed without the short, membrane attached subunit. The secretory form of MUC1 is not re-internalised and reaches the extracellular space via secretory vesicles with only one passage of the Golgi.

The subcellular location where O-glycosylation initiates is the Golgi, since the up to 20 polypeptide GalNAc-transferases are localized throughout the cis to trans compartments. The O-glycosylation position within the polypeptide chain is not easily predictable and depends on the availability of GalNAc-transferases, the amino acids sequences around the putative glycosylation site and already existing glycosylation on proximal Ser/Thr residues. (Elhammer AP et al., 1999). Up to now, at least 14 polypeptide GalNAc-transferases potentially involved in glycosylation of MUC1 are cloned and available for *in vitro* glycosylation (Clausen H et al., 1996; Bennett EP et al., 1998; Ten Hagen KG et al., 1999). Each of these enzymes has a specific site preference, but some overlapping between them exists.

The complexity of glycan chains is tissue specific, and characteristic patterns were described for a number of organ sites and tissue specific secretions. So the predominant presence of core2-based linear or branched glycan chains was detected in lactation-associated MUC1 compared to short glycans, mainly of core 1 type, found in urine associated MUC1.

1.2.1.5 MUC1 as a tumor antigen

MUC1, as well as ovarian tumor antigen CA-125, HER2neu and Carcinoembrynic Antigen (CEA), are glycosylated tumor antigens with proven potential in clinical diagnostics or in immunotherapy.

Tumor antigens are those antigens that are produced predominantly or exclusively by tumor cells and are recognised by the immune system. When presented only by tumor cells, but never by normal cells, they are called **tumor-specific antigens** and typically result from a tumor specific mutation. More common are antigens presented by both tumor and normal cells, named **tumor-associated antigens**. Cytotoxic T lymphocytes that recognize these antigens may be able to destroy the tumor cells before they proliferate or metastasize.

Two research groups showed by using specific monoclonal antibodies that sera from tumor patients were positive for MUC1/SEC and MUC1/REP (Smorodinsky N et al., Hilkens J et al., 1986). Thus, both the MUC1/SEC protein and the MUC1/REP extracellular cleavage product can be secreted or shed from breast cancer cells and used for monitoring during the patients recovery. Beside this, it was shown that the MUC1/Y protein is expressed by various human secretory epithelial tumors, but is not detectable in the adjacent normal tissue (Zrichan-Licht S et al., 1994; Hartman M et al., 1999; Baruch A et al., 1997)

In the mid 80ies it was shown that the tumor-associated MUC1 is **over-expressed in a non-polarised manner** and **shed** into the circulation in **increased amounts** (Hilkens J et al., 1986; Wesseling J et al., 1996). Due to pericellular expression of the several hundred nanometre long ectodomain, the MUC1 on carcinoma cells represents a primary immune target in anti-cancer strategies.

During the 90ies the MUC1-specific immune response in human sera was detected. Natural B cell responses were detected in breast cancer patients as well as in pregnant women or in women after passage of one lactation phase. These weak humoral responses were shown to provide a certain degree of survival benefit for breast cancer patients (Kotera Y et al., 1994). More specific, cellular responses of MUC1 specific T lymphocytes were detected in the cancer patients. These T-cells were able to lyse the cancer cell, but in an MHC unrestricted manner (Jerome KR et al., 1991). Due to their low efficiency, these detected immune responses to MUC1 are not sufficiently strong to eradicate the growing tumor.

Both humoral and cellular immune responses found in cancer patients were directed toward the tandem repeat domain of MUC1, mainly to its PDTR motif. Over 50 monoclonal antibodies were generated in mice and more than half of these were directed to this immunodominant motif (von Mensdorff-Pouilly S et al., 2000). Initially regarded as non-glycosylated, the PDTR peptide motif turned out to be O-glycosylated on the tumor-associated MUC1 (Muller S et al., 1997). Many of the peptide-specific antibodies bind with higher affinity to the glycosylated PDTR motif (Karsten U et al., 1998, 2004), indicating a glycosylation-induced stabilisation of the epitope.

The antibodies and T lymphocytes detected in cancer patients were specific for the cancer associated MUC1. Cancer associated MUC1 differs from lactation-associated MUC1 on the level of O-glycosylation (Muller S et al., 1999). The aberrant glycosylation is the structural difference that characterises the cancer-associated MUC1 as a tumor-specific antigen (Fig 7). It is visible in the form of three structural changes:

Glycans of the tumor-cell associated MUC1 are shorter than the polylactosamine-type chains found on lactation-associated MUC1. They are predominantly core 1-based, with numerous Tn (core1-GalNAc) or T antigen (core1-disaccharide Gal-GalNAc) and in particular their sialylated derivatives. The percentage of sialylated glycans is higher in cancer- than in lactation-associated glycoform.

This tumor-associated truncation could be caused by reduced levels of core2-specific enzymes, ß-*N*-acetylgalactosaminyltransferase or by the increased sialylation (over-expression of 3GalST) (Dalziel M et al., 2001; Brockhausen I et al., 1995), which presents the stop signal for further addition of sugars (Whitehouse C et al., 1997).

The third remarkable difference in MUC1 glycosylation occurring on tumor cells is the increased glycosylation density. The density of glycosylation is significantly higher on tumor-associated MUC1 than on lactation-.associated MUC1. The number of glycosylated Ser/Thr in the tumor MUC1 VNTR domain is remarkably increased with 4,8 glycan chains per repeat found in breast carcinoma cell lines compared to 2,6 glycan chains per repeat in lactation-associated MUC1 (Muller S et al., 1997,1999). The shortening of sugar chains directly correlated with the density of glycosylation in the different breast cancer cell lines.

Together with increased and unpolarized expression, this deviation unmasks the core peptide epitopes and exposes otherwise (sterically) hidden glycan epitopes. (Karsten U et al., 2005*)*.

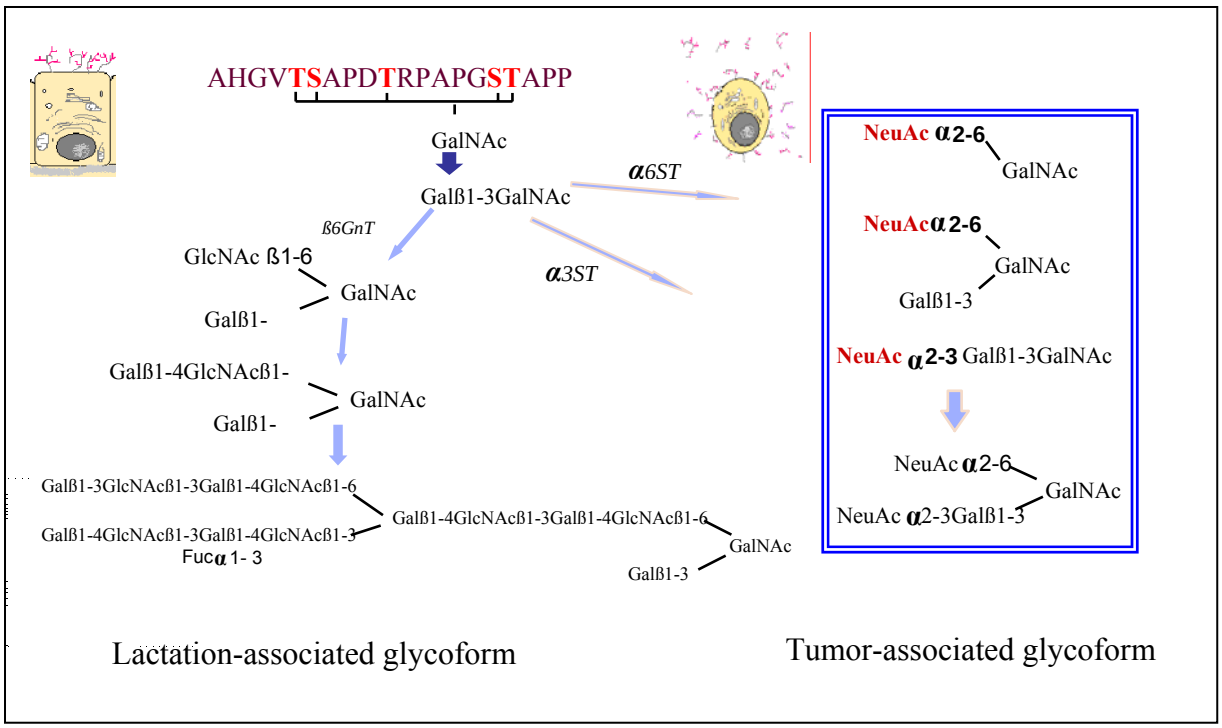

Figure 7. Glycosylation of lactation- and tumor-associated MUC1. Glycosylation positions within a tandem repeat are bolded. Glycan chains attached to Tn antigen characteristic for lactation- or tumor-associated MUC1 are depicted on opposite sites of the figure.

1.2.1.5.1 MUC1 functions in tumor cells

The aberrant expression and glycosylation levels of MUC1 have consequences for the functions and interactions of the protein.

The pericellular expression of tumor-associated MUC1 protein plays a distinctive role in sterically reducing cell-cell and cell-extracellular matrix interactions. It has been postulated that MUC1 may be involved in the metastatic spread of cancer cells from the initial tumor site (Ligtenberg MJ et al., 1992; Hartman M et al., 1992; Wesseling J et al., 1995). It was demonstrated that MUC1 expression positively correlates with tumor formation and progression in knock out mice (Spicer W et al., 1995). Also the short form of MUC1 protein, MUC1/Y is reported to stimulate oncogenesis and tumor progression (Baruch A et al., 1997).

The negative charge of abundant sialyl groups present on the terminal ends of MUC1 sugar chains facilitates the interaction of tumor cells with E-and P-selectine, promoting the adhesion to endothelial cells and extravasation into the surrounding tissue during metastasis. Moreover, the negative charge is supporting spreading and dissemination of tumor cells by introducing repulsing force between them.

The over-expressed MUC1 on tumor cells shields the cell from immune recognition by the cellular part of the immune system, and thus favours tumor cell survival and metastasis.

Therefore role played by MUC1 in cancer progression represents two sides of one coin: On the one end, loss of polarity and over-expression of MUC1 in cancer cells interferes with cell-adhesion and shields the tumor cell from immune recognition by the cellular arm of the immune system, thus favouring metastases. On the other hand, MUC1, in essence a self-antigen, is displaced and altered in malignancy and induces immune responses to it. Tumor-associated MUC1 has short carbohydrate side-chains and exposed epitopes on its peptide core; it gains access to the circulation and comes in contact with the immune system provoking humoral and cellular immune responses to it.

1.2.1.6 Immunity to MUC1

Natural antibodies to MUC1 are present in the circulation of cancer patients, both as free antibodies and as immune complexes, bound to MUC1. Both IgM and IgG antibodies were detected in breast cancer patients and their early appearance is connected with good prognosis (Fig 8).

Natural MUC1 antibodies from breast cancer patients reacted more strongly to the MUC1 peptides with cancer-specific glycosylation (GalNAc) than to the naked VNTR sequence.

Therefore, as the naturally existing tumor-related MUC1 is glycosylated, the immune response against breast cancer should be directed toward glycosylated MUC1 epitopes.

Initially, the glycosylation was considered to prevent antibody binding to the peptide backbone. Besides this, it was shown that circulating tumor MUC1 antigen found in tumor patients can not be processed by dendritic cells and does not elicit either MHC class II- (Hiltbold EM et al., 2000) nor CTL response (Hiltbold EM et al., 1999).

These findings were shown to be related to heavy glycosylation (Hiltbold et al., 2000) and not to a tolerance to MUC1 (Hiltbold et al., 1998).

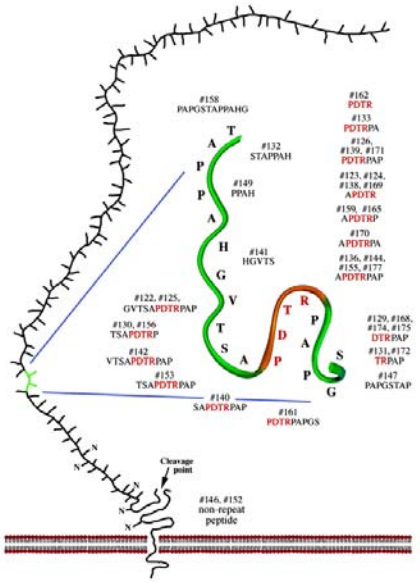

Figure 8. MUC1 VNTR specific antibodies.

The high avidity interaction of MUC1 glycan chains with the mannose receptor residing in early endosomes prevents the protein to traffic to late endosomes after endocytosis where processing occurs (Hiltbold EM et al., 2000).

Additionally, the digestion of densely glycosylated MUC1 by the enzyme Cathepsin L, involved in MHC class II epitope generation is prevented or restricted by site-specific O-glycosylation (Hanisch FG et al., 2003). Consequently, under-glycosylation was regarded as the prerequisite for an effective immune response against tumor MUC1. In particular, it was postulated that the immunodominant domain within tandem repeats, the PDTR sequence, was not glycosylated. As a consequence, cancer vaccines were developed based on unglycosylated MUC1 (Karsten U et al., 2005).

MUC1 positive tumors were targeted in a series of vaccination attempts in mouse models and in clinical trials. The clinical trials comprised

- gene-based vaccine approaches, where MUC1 is encoded by a highly attenuated vaccinia virus vector (Scholl SM et al., 2000),
- antibody-based vaccines (Heuser C et al., 2003, Wilkinson RW 2000),
- peptide-based vaccination (Ramanathan RK et al., 2005; Wierecky Y et al., 2006; Yamamoto K et al., 2005; Kohlgraf KG et al., 2004 ; Plamer M et al., 2001),
- MUC1/interleukins synthetic chimera (Snyder LA et al., 2006),
- fusion of tumor and dendritic cell (Tanaka Y et al., 2005) or
- vaccination based on autologous, peptide-pulsed dendritic cells (Loveland BE et al., 2006; Wierecky Y et al., 2006).

The peptide-based strategies stimulated strong humoral IgM and IgG responses, but with no or with weak cellular immune response (Soares MM et al., 2001; Stepensky D et al., 2006). The developed IgM and IgG reactivities were peptide-specific and therefore failed to recognize glycosylated tumor related MUC1. The approach using autologous dendritic cells loaded with nonglycosylated-MUC1 stimulated peptide-specific CTL in 50% of the patients with metastatic renal cell carcinoma (Wierecky Y et al., 2006).

Such a moderate success of peptide vaccines can presumably be ascribed to the structural discrepancy between nonglycosylated epitopes used as immunogens and the glycosylated

antigen of targeted cancer cells. As glycosylation, albeit aberrant, is still present in the tumor-associated mucin the optimal MUC1 vaccine is not a peptide, but a glycopeptide.

1.2.1.7 Immunity to tumor and MUC1 tumor vaccines

The theory of *immune surveillance* claims that the immune system continually recognizes and eliminates tumor cells. Immune surveillance is most likely to be successful against virus-induced tumors, which express foreign antigens. But even the tumors with antigens which can be recognized by the host immune system can evade immune elimination. There are numerous mechanisms used by tumor cells to evade immune elimination: (1) expression of anti-apoptotic genes blocking the activity of CTLs and NK cells as well the effects of chemotherapy, (2) secretion of inhibitory cytokines TGFβ, IL-10, VEGF which prevent maturation of dendritic cells and recruit regulatory T lymphocytes inhibiting activation of helper and cytotoxic T cells that arrived at the tumor site. (3) reduction / lack of MHC expression, (4) lack of co-stimulation, (5) antigen shedding, (6) sequestration of tissue from surveillance , (7) lymphocyte killing…

To overcome these sabotaging mechanisms of the tumor, efforts are made to develop tumor vaccines for the induction of a stronger and more specific anti-cancer cell response. Tumor vaccines should utilize tumor specific antigens and stimulate dendritic cell maturation and T lymphocyte activation outside of the tumor micro-surrounding and thereby overcome some of the possible evasion mechanisms. Therefore, for the design of anti-breast cancer vaccine the best MUC1-derived antigens and glycosylation forms have to be known and interaction of glycopeptides with the immune machinery has to be elucidated.

Processing of glycosyalted MUC1 antigen is one of the crucial steps in generation of MHC epitopes and stimulation of MUC1 specific T-cells. Previously, Vlad et al (2002) showed that O-linked glycans are not removed from MUC1 during processing in the MHC class II pathway in dendritic cell. The glycans control the extent of MUC1 proteolysis in endosomes and site specificity of Cathepsin L, the major endosomal protease involved in MHC class II antigen processing (Hanisch FG et al., 2003). T-helper cell responses specific for MUC1 glyco-peptides were reported by Vlad et al (2002). A MUC1 glycopeptide was also shown to bind to MHC class I and to induce activation of CTLs which was not glycopeptide specific (Apostolopoulos V et al., 2003).

1.2.2 Antigen presenting cells

The most potent antigen-presenting cells (APCs) involved in the tumor defence are dendritic cells. Their role is to internalise circulating antigens and present them on MHC class II and I molecules. During the process of internalisation and processing of antigen, immature dendritic cells undergo maturation, expressing high amounts of MHC and co-stimulatory molecules. Co-stimulatory molecules are necessary for the activation of naïve T lymphocytes, in the T cell areas of lymph nodes.

Dendritic cells are able to present endocytosed antigens in MHC class II and class I fashion. This process of presentation of extracellular antigens in the MHC class I pathway is called cross-presentation of antigen or cross-priming of APCs. It refers to two possible situations of processing of endocytosed extracellular antigen: (1) antigen can escape from endosomes and enter the cytosol where it reaches proteasomes and other cytosolic enzymes (Hearn AR et al., 2004; Seifert U et al., 2003) or (2) antigen gets processed within endosomes by cathepsins residing there and loaded onto recycled MHC class I molecules (Chen L et al., 2004, Shen l 2004 immunity). Both of these pathways can generate 8 to 11-meric peptides potentially fitting to the groove of MHC class I molecules.

1.2.3 Proteasome

Proteasomes are multi-enzyme complexes that digest proteins. They were discovered by Alfred Goldberg and Martin Rechtsteiner in 1980s. The average human cell contains 20.000 to 30.000 proteasomes present in both cytoplasmic and nuclear compartments, which in 90% degrade missfolded proteins, but also proteins with high turnover rates are degraded by IPs. By degrading antigen, proteasomes create immune epitopes, which is their second function.

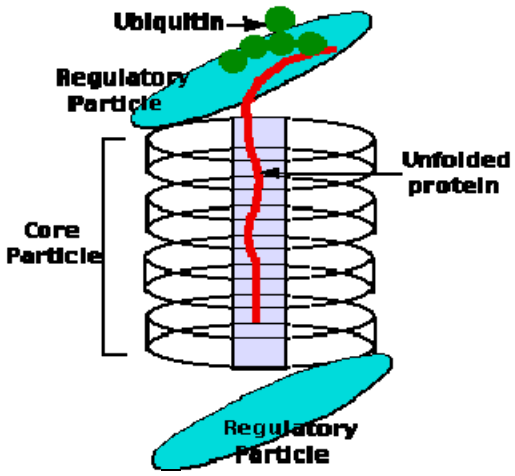

Figure 9. Structure of 26S proteasome

The general structure of a proteasome involves three subunits: two regulatory 19S particles and catalytic 20S core particle (Fig. 9). At both ends of the 20S core are 19S caps, also known as regulatory units since they recognise proteins to be degraded by the attached polyubiquitin chain and unfold the proteins before it can enter the cavity of the 20S core for active hydrolysis of peptide bonds.

The dimensions of the proteasome cavity (148 Å) and entrance (131 Å) are such that only unfolded proteins can enter (Unno M et al., 2002). The unfolding of polypeptide chains of native proteins by cap is ATP driven.

The basic 20 S core is evolutionary conserved from archebacteria through yeast down to humans, whereas the associated ubiquitin system is found only in eukaryotes. The 20S core consists of a barrel shaped cylinder made of four annular rings of protein subunits: the end ones known as alpha rings sandwich the two beta rings (Fig 10). Beta rings contain proteolytic enzymes on the inner side of the cylinder; the alpha rings have no known enzyme activity.

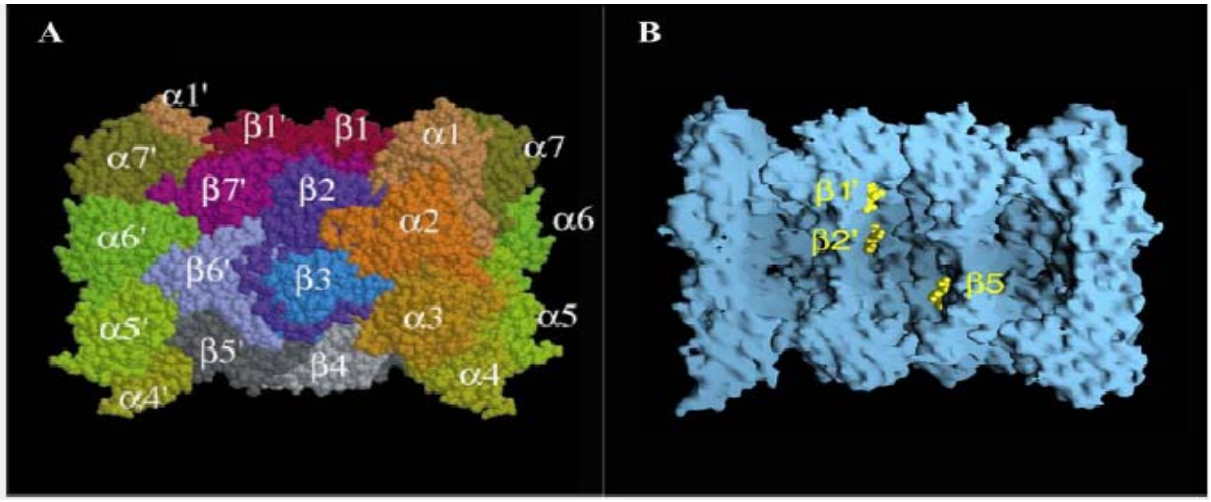

Figure 10. Structure of the *Saccharomyces cerevisia*e 20S proteasome
A. Slightly tilt side view in space-filling mode showing the overall subunit arrangement, which results in a C2 symmetry of the particle. All seven α- and β-type subunits are marked and shown in different colors.
B. Surface view of the proteasome molecule clipped along the cylinder axis shows the central active chamber (formed by two face to face-oriented β-rings) which is separated from two antechambers (formed by the β- and α-ring) by narrow constrictions. The catalytic Thr1 residues of active subunits in the central active chamber have calpain inhibitor molecules bound (shown as space-filling models in yellow).

There are two kinds of proteasomes existing: constitutive and immuno-proteasomes. Constitute proteasomes are expressed in each cell type. Immunoproteasomes are found mainly in antigen-presenting cells after stimulation of these by INF-γ. Human immunoproteasomes represent a subset of inducible 20S proteasomes, occurring by substitution of three proteolytically active subunits LMP2, LMP7 and MECL-1 upon interferon-γ induction.

Due to their cytosolic and nuclear localisation, immuno-proteasomes are digesting mainly non-glycosylated proteins and the only glycosylation form they come into contact with is the O-linked GlcNAc, a substoichiometric modification at the monosaccharide level that is involved in the regulation of protein function, similar to, but mostly antagonistic to

phosphorylation. It was demonstrated that proteasome can process O-GlcNAc modified proteins. Although no data showed that proteasomes can deal with longer glycan chains, this information indicated that proteasomes have some sort of tolerance to at least short glycan moieties. As the cancer related MUC1 is glycosylated with short glycan chains, there is a possibility that MUC1 glycopeptides or complete glycoprotein can be substrates of immunoproteasomes.

1.2.3.1 The proteasome specificity

The two central β rings of proteasomes have six threonine protease catalytic active sites. Two of these active sites have a *chymotrypsin like activity* and digest the peptide after hydrophobic amino acids, two sites have *trypsin like activity* digesting after basic amino acid residues, and the residual two sites cleave behind acidic amino acid residues *similar to caspase activity*. The active sites take part in antigen degradation and production of immune-epitopes. The efficiency of immune-epitope generation is rather low, and requires digestion of 1000 molecules for production of a single MHC binding epitope (Boes M et al., 2004). Short 8 to 11-meric peptides leave proteasome and with the help of TAP protein (transporter associated with antigen processing) enter ER where they are loaded onto empty MHC class I proteins.

The proteasome is a threonine protease (Seemuller E et al.,1995). Its active centre is build by an N-terminally located threonine, one lysine, two aspartic acids and two serines. The two serines and one aspartic acid carry electronegative oxygen atoms that are attracting hydrogen of the amino group of the terminal threonine. The lysine with its positive charge makes a salt bridge with second aspartic acid, and this positive charge lowers the pKa value of the hydroxyl group of terminal Thr (Chen P et al., 1996; Schmidtke G et al., 1996; Seemüller E et al., 1996; Heinemeyer Wet al., 1997). The hydrogen of the Thr-OH group is transferred to its amino group, probably with the help of water molecule, and a reactive oxygen can nucleophilically attack the peptide bond of the substrate. The cleavage is terminated by degradation of generated acyl-ester intermediate mediated by water molecules. Thereby is proteasomal threonine released for the next catalytic cycle.

1.2.4 MHC molecules

The Major Histocompatibility Complex (MHC) is a set of molecules displayed on cell surfaces of all nucleated cells. Their main function is a presentation of "self" and "non-self" antigens and activation of T lymphocytes. There are app. 50,000 - 100,000 MHC molecules on a typical cell and most of them (class I) are occupied by self peptides.

MHC molecules can be divided in two classes: MHC class I and class II. MHC class I molecules are present on all nucleated cells, whereas MHC class II are restricted to antigen-presenting cells. Each class of human MHC is represented by three gene loci that are called "HLA" for Human Leukocyte Antigen. The class I loci are HLA-A,-B and -C and the class II loci HLA-DR, -DQ and -DP. In the human genome more than 200 MHC alleles that spread over HLA loci are described. This polymorphism allows a variety of peptide antigens to bind to MHC molecules. The population of app. 100.000 MHC class I molecules on a somatic cell displays >1000 different peptides. One single peptide-MHC complexe can be represented in widely different amounts from 1 - 5000 molecules/cell (mean~100).

The allelic variation of MHC molecules is functionally reflected in the selection of peptides, which can bind; each allelic product has a unique set of peptides, which it can bind with high affinity (though rarely particular peptides may bind to more than one MHC allele).

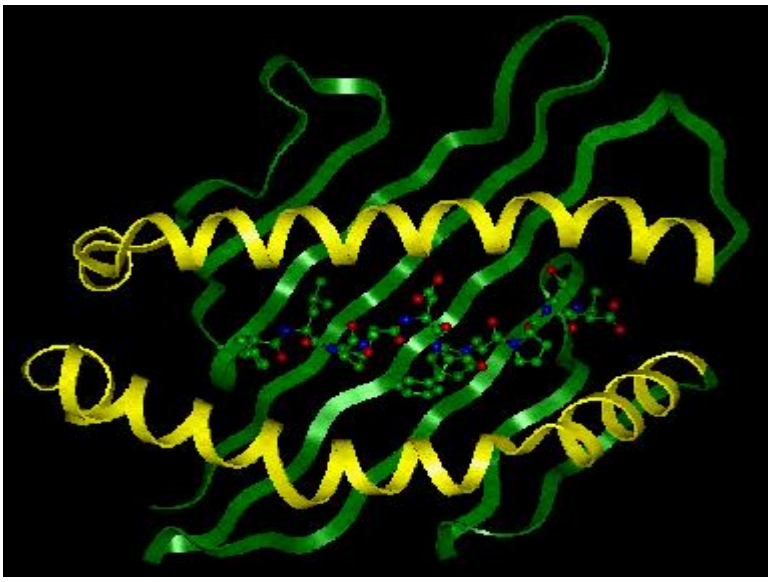

Figure 11. MHC class I molecule from top with peptide bound to the protein groove. Yellow spirals present two alpha helixes, green bands – beta-plated sheet. Between two alpha helixes is green/blue/red peptide chain (antigen).

MHC Class I is a heterodimeric membrane molecule consisting of a heavy alpha chain and a light beta-microglobulin chain (Fig. 11). The alpha chain can be divided into three extracellular domains, alpha1, alpha2 and alpha3, in addition to the transmembranous and cytoplasmic domains. Between the alpha-1 and alpha-2 domains a peptide binding groove is located that is formed by beta-pleated sheet on the bottom and two alpha helices on the sides. This region has several pockets and is capable of binding a small peptide via non-covalent interactions. The beta subunit, a beta-Microglobulin, associates primarily with the alpha-3 domain and is necessary for MHC stability.

Class I molecules usually bind nonapeptides, and less frequently octapeptides or decapeptides. The peptides are anchored within "pockets" in the peptide-binding groove of MHC class I only at a few positions of the peptides, which are called "anchor residues". Other side chains of the bound peptide also make contact with the groove, but apparently not constrained by any specific pocket. Nonapeptides bind in an extended conformation with a kink near P4 (fourth amino acid of the peptide). Octapeptides bind with less acute kink, whilst decapeptides and longer peptides bind with more pronounced kinks or “zigzags” in the extended conformation.

The MHC/peptide complexes play a fundamental role in regulation of immune responses. T-cells can recognize antigens normaly as a complex with MHC molecules. Rare MHC-independent activation of T-cells was also reported. T-cells vary in the need for number of specific MHC/peptide complex from a few complexes per cell to a few thousand, depending on the affinity or activation state of the T-cell and on the antigen presenting cell. Recognition of MHC/peptide complex by T-cells requires antigen to be processed by unfolding and proteasomal digestion before it complexes with the MHC molecule. Once formed the complex of antigenic peptide and MHC is generally very stable (half life ~ 24h).

1.3 Aim of the study

Breast cancer patients often develop weak but helpful immune response directed toward the tumor antigen MUC1. This protein is in cancer cells expressed in increased and unpolarized manner. Moreover, it is aberrantly O-glycosylated with shorter, denser and more often sialylated sugar chains than those of normal epithelial cells. These structural changes make MUC1 a tumor antigen with diagnostic and therapeutic potential. Several studies showed that the most immunodominat part of the protein is exactly this aberrantly O-glycosylated domain with tandem repeats of a 20-meric peptide sequence. The stimulation of immune responses against the cancer-related form of MUC1 is rationalized to have therapeutic effects in breast cancer patients. For the development of a breast cancer vaccine, those glycoforms should be identified, which are able to induce the strongest cytotoxic and humoral response to the natural target. Hence, we attempted in this study to:

- Identify the influence of glycan structure and position on antigen processing by immunoproteasomes of antigen-presenting cells
- Identify the structure of MUC1 derived glycopeptides that are the best substrate for immunoproteasomes
- Define those MUC1 glycopeptides which are processed to MHC class I binding glyco-epitopes
- Identify glycosylated peptides with high binding capacity to the most frequent MHC class I allele in the human population (HLA-A0201)
- Understand the weak immunogenicity of natural glycoforms of tumor-associated MUC1.

2 MATERIALS AND METHODS

MATERIAL

- 2,5-Dihydroxybenzoic acid, DHB, Bruker Daltonics
- 2-mercaptoethanol 50mM, Gibco, UK
- Accutase PAA Laboratories GmbH, Germany
- Acetic acid, Fluka Chemie, Germany
- Acetone HPLC grade, JT Baker, Germany
- Acetonitrile Chromasolv for HPLC, Sigma-Aldrich, Germany
- Alpha-dithiotreithol, DTT, minimum 98%, Sigma-Aldrich, Germany
- Beta-2 micro globulin from human urine, 90%, Sigma-Aldrich, Germany
- BSA Bovine serum albumin fraction V, Carl Roth GmbH, Germany
- Cacodylic acid sodium salt, 98%, Sigma Chemical
- Chloroform, anhydrous, 99+%, Aldrich, Germany
- Cholesterol, min 99%, Sigma-Aldrich, Germany
- Citric acid monohydrate, KMF optichem, Laborchemie Handels
- Dimethyl sulphoxide, DMSO, hybrid-max, Sigma, Germany
- Ethanol, Absolut Chromasolv, Sigma-Aldrich, Germany
- Ethylendiaminetetraacetic acid, EDTA, disodium salt dihydrat, Fluka Chemie, Germany
- FITC Fluorescein isothiocyanate isomer I, Sigma-Aldrich, Germany
- Foetal Bovine Serum, FCS, heat inactivated, Gibco, Invitrogen Corporation
- Formaldehyde solution, min 35%, extra pure, Merck, Germany
- HEPES pufferan, >99,5%, Roth, Germany
- Hydrochloric acid, HCl, 1 mol/l, volumetric solution, Sigma-Aldrich, Germany
- Magnesium chloride, $MgCl_2$, 99%, Sigma-Aldrich, Germany
  Manganese (II) chloride dehydrate, $MnCl_2$, Merck, Germany
- Methanol dried, Sigma-Aldrich, Germany
- Sodium hydroxide solution, NaOH Fluka
- Nonessential amino acids, 100X, Biochrom AG
- Nonidet-P40, Fluka BioChemika
- PBS Dulbeco, Biochrom AG
- PEG-PE 2000 1,2-Dipalmitoyl-sn-Glycero-3-Phosphoethanolamine-N-[Methoxy(Polyethylene glycol)-2000], sodium salt, Avanti Polar Lipids, USA
- Penicillin-Streptomycin solution, Sigma Aldrich, Germany
- Phosphoric acid, orto-phosphoric acid 85%, Merck, Germany, KGaA
- Poly-lactic acid Fluka Chemie GmbH
- Polyvinyl alcohol Sigma Chemicals
- POPC, 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphocholine, Avanti Polar Lipids, USA
- Potasium chloride, KCl, >99,5%, Roth, Germany
- Potasium hydroxide, KOH, Carl Roth GmbH
- Primulin, Sigma, Germany
- Protein A Sepharose 4 Fast Flow Amersham Pharmacia Biotech, Sweden
- Sodium azide, Fluka Chemica

- Sodium chloride, NaCl, KMF Laborchemie Handels GmbH, Germany
- Sodium hydrogen phosphate, dehydrate, Merck
- Trifluoroacetic acid, TFA Fluka Chemie
- Tris-(hydroxymethyl)-aminomethan, Merck
- Trypsin/EDTA solution, 0,5%/0,2% (w/v) in PBS (10x), Biochrom AG

SPECIFIC CHEMICALS

- 20S immunoproteasomes, Immatics;Germany
- Albumin from hen egg white, Fluka Biochemika
- Ficoll-Paque PLUS, Amersham Bioscience, Sweden
- Fluorescent mounting medium, Dako Cytomation;Danemark
- GM-CSF, Recombinant human, BD Bioscience, Pharmingen, Germany
- IL-4, recombinant human, Cell Genix Tecnologie Transfer GmbH Germany
- Peptides, Mimotopes/Perbio
- Protease inhibitor cocktail tablets, Complete Mini, Roche Diagnostics GmbH
- Proteasome inhibitor Ada(Ahx)3-(Leu)3-vinyl sulphone, Biomol
- Sepharose 4b, Pharmacia Bioscience
- TNF alpha, Tumor necrosis Factor-alpha, Recombinant, Sigma-Aldrich, Germany

SOFTWARE

- Confocal Microscope software, Leica confocal software
- ESI/MS/MS software MassLynx 4.0 Waters
- FACS software, Expo 32 ADC XL4 Colour/Analysis, Bruker Daltonics
- Find Pep http://expasy.org/tools/findpept.html
- HPLC software Gold chromatography data software, Beckman Coulter
- MALDI software Flex Control /Analysis, Bruker Daltonics
- NetChop http://www.CBS.dtu.dk/services/NetChop/
- Syfpeithi http://www.syfpeithi.de

PARTICULARS

- Cell culture plates: 6 well, 48 well, 150cm², 300cm², TPP Technoplastic Products AG, Switzerland
- Chromabond HR-P glass column, Carl Roth, Germany
- Cryo-tube vials, NUNC
- Glass coverslides, microscope cover glasses, Marienfeld, Germany
- Glass slides, microscope slides, precleaned polysine ^{tm} , Menzel GmbH, Germany
- Maldi plate MTP 384 massive Au coated, Bruker Daltonics
- Membrane for extruder, polycarbonate membrane pore diameter 100nm, Avestin Inc
- MACS separation column, Miltenyi Biotech, Germany
- Preseparation Filters, Miltenyi Biotech, Germany
- Midi-MACS separation unit, Miltenyi Biotech, Germany

- Nanospray needle, spray capillaries, Medium, for Micromass Qtof, Poxeon Biosystems
- Nanospray source, Qtof II, Micromass
- Previal C18 5u rpHPLC column, 150mm x 10mm, Alltech, Munich, Germany
- Solid phase extraction backerbond C18 SPE 200mg/3ml, JT Baker
- Syringe sterile filter 0,22µm fast flow PES membrane, TPP Technoplastic Products AG, Switzerland
- Thin layer chromatography plates HPTLC aluminium sheets silica gel 60 20x20cm, MERCK, Germany
- Ultrafiltration tubes, Centriprep centrifugal filer devices, Millipore Corporation, USA
- Ultrasphere HPLC ODS column, 2 mm x 15 cm, Beckman Coulter

MASHINES

- Centrifuge Z323K Hermle
- Confocal microscope, Leica CTR MIC, Leica Microsystems GmbH
- ESI/MS/MS Qtof II, Micromass
- Extruder, Lipofast with stabilisator, Avestin Inc
- FACS- flow cytometer EPICS XL4C, Beckman Coulter
- HPLC system, Beckman system gold, Beckman Coulter
- Incubators- cell culture, Heraeus 6000 Heraeus Instruments
- Incubators- dry , Function line, Heraeus Instruments
- MALDI-TOF system, Reflex IV, Bruker Daltonics
- Peristaltic pump P-1, Pharmacia, Biotech
- pH-Meter pH 538 Multical Wisseschaflich-Technische Werkstätten GmbH, Germany
- Sonicator, Brandsonic
- Ultracentrifuge Optima-L70 Beckman Coulter
- UV lamp UVIS multipurpose equipment for UV254/366 nm Desaga, Heidelberg,Germany
- Vacuum centrifuge, DNA plus, Heto

CELL CULTURE MEDIA

- CellGrow DC medium, Cell Genix Technologie Transfer GmbH, Germany
- DMEM Dulbeco's MEM with stabile glutamine, Biochrom AG
- IMDM, basal Iscove medium with stabile glutamine, Biochrom AG
- Phenol-red free RPMI 1640 medium, Biochrom AG
- RPMI 1640 medium, Biochrom AG

ANTIBODY

- Alexa Fluor 488, goat anti-mouse IgG, Molecular Probes
- Anti-CD14 microbeads, human, Miltenyi Biotech, Germany
- W6/32 monoclonal anti-human HLA class I antigen clone W6/32, Sigma; Germany

2.1 Cell-biological methods

2.1.1 Culture of mouse dendritic cell line

DC2.4 cell line was obtained from the American Type Culture Collection and grown in DMEM medium supplemented with 10% FCS, penicillin/streptomycin (200 IU/ml-200 µg/ml), 1% nonessential amino acids, 50 µM 2-mercaptoethanol at 37°C and 5% CO_2. Cells were detached by 10 min incubation in PBS containing 2 mM EDTA, centrifuged at 180xg / 5min and split 1:10. In all experimental procedures followed by mass spectrometric analyses, cultivation of cells was carried on in media without phenol red.

2.1.2 Culture of T2 cell line

The T2 cell line is a negative mutant for "Transporter associated with antigen processing" (TAP protein), expressing empty HLA-A0201 allele of MHC class I molecule on the cell surface. The cell line was a kind gift from Prof. Jonathan Howard and Dr. Michael Knittler, University of Cologne, and was initially grown at 37°C and 7,5% CO_2 in IMDM media supplemented with 10% FCS and penicillin/streptomycin (200 IU/ml-200 µg/ml). For the MHC class I stabilisation experiments, T2 cells were slowly adapted to reduced serum conditions: 2% FCS.

2.1.3 Culture of human mammary carcinoma cell lines

The mammary carcinoma cell line T-47D and MCF7 were obtained from the American Type Culture Collection. T-47D cell line was grown in RPMI 1640 with 10% FCS and penicillin/streptomycin (200 IU/ml-200 µg/ml), at 37 °C and 5% CO_2. MCF7 cell line was grown in DMEM with 10% FCS and penicillin/streptomycin (200 IU/ml – 200 µg/ml). For splitting, cells were washed with PBS and detached by incubation with 1x trypsin or Accutase solution for 10 minutes in the cell culture incubator. Medium was replaced at intervals of 2-3 days. In all experimental procedures followed by mass spectrometric analyses, cultivation of cells was carried on in media without phenol red.

2.1.4 Long time storage of cells

For long time storage cells were frozen in liquid nitrogen. A minimum of one million cells were resuspended in 90% fully supplemented medium or in pure FCS (for T2 cells) and 10% DMSO. Cells were placed in Cryo Tube vials, gradually cooled to -80°C in Cryo freezing box for minimum 2 h but not longer than 24h and stored in liquid nitrogen tanks.

2.2 Biochemical methods

2.2.1 MUC1 derived peptides and glycopeptides

MUC1 derived glycopeptides were chemically synthesised and kindly provided by Prof. Hans Paulson (GP1-16; GGP1-4; SGGP1-5). The P1 100-mer peptide and P2 ESR61 were provided by Dr. Olivera J. Finn and *in vitro* glycosylated with GalNAc using purified polypeptide GalNAc-transferases T1 and T2 (kindly provided by Dr. Henrik Clausen, School of Dentistry, University of Copenhagen, Denmark). Other nonglycosylated peptides were ordered from Mimotopes/Perbio and glycosylated *in vitro*. TAP25 was synthesized in a local facility at the Institute of Biochemistry, Cologne, Germany.

		NAME	SEQUENCE
PEPTIDE		P1(DTR100)	(HGVTSAPDTRPAPGSTAPPA)x5
		P2(ESR61)	A(HGVTSAPESRPAPGSTAPPA)x3
		P3(AHG21DTR)	AHGVTSAPDTRPAPGSTAPPA
		P4(AHG21ESR)	AHGVTSAPESRPAPGSTAPPA
GLYCOPEPTIDES	MONOSACCHARIDE	GP1(Tn100)	(HGV**T**SAPDTRPAPG**ST**APPA)x5
		GP2 (H11)	AHGV**T**SAPDTRPAPGSTAPPA
		GP3(H12)	AHGVTSAPD**T**RPAPGSTAPPA
		GP4(H13)	AHGVTSAPDTRPAPGS**T**APPA
		GP5(H14)	AHGV**T**SAPESRPAPGSTAPPA
		GP6(H15)	AHGVTSAPE**S**RPAPGSTAPPA
		GP7(H16)	AHGVTSAPESRPAPGS**T**APPA
		GP8(A4GALNAC)	AHGVT**S**APESRPAPGSTAPPA
		GP9(SAP20-1)	**S**APDTRPAPGS**T**APPAHGVT
		GP10(SAP20-2)	**S**APD**T**RPAPGSTAPPAHGVT
		GP11(SAP20-3)	**S**APD**T**RPAPGS**T**APPAHGVT
		GP12(SAP20-4)	SAPD**T**RPAPGSTAPPAHGVT
		GP13(SAP20-5)	SAPD**T**RPAPGS**T**APPAHGVT
		GP14(SAP20-6)	SAPDTRPAPGS**T**APPAHGVT
		GP15(A11)	HGV**T**SAPDTRPAPGS**T**APPA
		GP16(A12)	HGV**TS**APD**T**RPAPG**ST**APPA
	DISACCH	GGP1(A1)	AHGV**T**SAPDTRPAPGSTAPPA
		GGP2(A2)	AHGVTSAPD**T**RPAPGSTAPPA
		GGP3(A3)	AHGVTSAPDTRPAPGS**T**APPA
		GGP4(A4)	AHGVT**S**APDTRPAPGSTAPPA
	SIALYL	SGGP1 (A4ST)	AHGV***T***SAPDTRPAPGSTAPPA
		SGGP2 (A13ST)	AHGV***TS***APDTRPAPGSTAPPA
		SGGP3 (A14ST)	AHGVTSAPDTRPAPG***ST***APPA
		SGGP5 (A8ST)	AHGV***TS***APD***T***RPAPG***ST***APPA

Table 2. List of MUC1 VNTR peptides and glycopeptides used in processing studies. Glycosylated amino acids are labelled with yellow and with bold (GalNAc carrying), underlined bold (GalGalNAc carrying) or italic bold underlined (Sialyl-GalGalNAc carrying)

2.2.2 Enzymatic glycosylation of peptides *in vitro*

For *in vitro* glycosylation of all five Ser/Thr in the 20-meric tandem repeat peptide a minimum of three different polypeptide GalNAc transferases should be used: ppGalNAc-T1, ppGalNAc-T2 and ppGalNAc-T4. ppGalNAc-T1 initiates glycosylation of Thr in the VTSA sequence of the tandem repeat, but can also glycosylate Ser and Thr in the GST motif (Wandall HHet al, 1997; Hanisch FG et al, 1999). ppGalNAc-T2 rapidly initiates glycosylation of Thr in the GST motif, and can less efficiently glycosylate Ser in the same region and Thr in the VTSA motif. The glycosylation of Thr in DTR and Ser in VTSA can be performed by using ppGalNAc-T4, which is exclusively dependent on glycans already present on the substrate peptide (Bennett EP et all, 1998).

The peptide substrates (50 nmol) were dried from solutions in water by vacuum centrifugation and solubilized in 368 µl of a 25 mM cacodylat buffer containing 10 mM $MnCl_2$ (final pH 7.4). Addition of the cosubstrate UDP-GalNAc (0,28 mg = 112 µl of stock solution containing 2,5 mg/ml of cacodylat buffer) was followed by the respective enzyme preparations ppGalNAc-T1 and ppGalNAc-T2, 20 µl of ppGalNAc-T1 (2.04 mU/ml) and/or ppGalNAc-T2 (4.04 mU/ml) to yield a total volume of 500 µl. The reaction mixtures were incubated for 48 h at 37°C by adding fresh aliquots of cosubstrate and enzyme(s) after a 24 h interval. The reaction was stopped by acidification with TFA to pH 2. The glycosylated peptide was purified from the reaction mixture by reversed-phase HPLC.

2.2.2.1 Purification of i*n vitro* glycosylated peptides by rpHPLC

Aliquots (50-100 µl) of the reaction mixtures were injected onto a narrow-bore ODS Ultrasphere column (150 × 2 mm) for analytical chromatography or onto a Prevail C18 column for preparative chromatography. Chromatography was performed on an HPLC system by gradient elution in a mixture of Solvent A (H_2O with 2%ACN and 0,1%TFA, v/v) and Solvent B (80%ACN and 20 % H_2O with 0,1%TFA, v/v). The gradient for analytical chromatography increased from 0% solvent B to 6% solvent B within 3 min, followed by an increase from 6% to 36% solvent B within 40 min. The gradient applied in preparative chromatography increased from 5% to 50% solvent B within 30 minutes. The glycopeptides were run at a flow rate of 0,3 ml/min on the narrow-bore ODS column and at a flow rate of 2 ml/min on the Prevail column. The photometrical detection was performed at a wavelength of 214 nm.

2.2.3 Proteasomal processing of MUC1 peptides and glycopeptides

2.2.3.1 Control of Proteasome Purity

Proteasome purity was controlled by the incubation of proteasome and substrate in the presence of a proteasome specific inhibitor and residual activity was measured by using the non-glycosylated reference peptide P1 as substrate. The assay was performed by incubating 1 µg of human 20S immunoproteasomes with 150 µM proteasome inhibitor Ada(Ahx)3-(Leu)3-Vinyl sulphone in 10 µl digestion buffer B for 10 min at room temperature, followed by the addition of 10 µg of P1 peptide and incubation at 37ºC for 24 h.

2.2.3.2 Optimisation of In Vitro Proteasomal Processing Assay:

MUC1 VNTR peptide and glycopeptides P1, GP1 and GP11 (10 µg) were incubated with 1µg of human 20S immunoproteasomes in 10 µl of digestion buffer A, B or C.

Buffer A: 50 mM Tris-HCl, pH 7.6; 25 mM KCl; 10 mM NaCl ; 1 mM $MgCl_2$;1 mM DTT;

Buffer B: 20 mM HEPES-KOH, pH7.6; 2 mM MgAc2

Buffer C: 20 mM HEPES/KOH, pH7.6; 2 mM MgAc2; 1 mM DTT buffer.

The reaction was performed at 37ºC for 48h.

The processing efficiency was measured by rpHPLC as described above for analytical chromatography of glycopeptides (section 2.2.2.1). When the nonglycosylated reference peptide P1 was quantitatively fragmented, the digestion was stopped by freezing the sample at -80ºC.

2.2.3.3 Analyses and Separation of the Cleavage Products by rpHPLC:

Each sample was analysed and cleavage products were fractionated by rpHPLC (as described for analytical chromatography (section 2.2.2.1). The gradient elution was performed by increasing the proportion of solvent B as described above to 36% (for longer peptides) or to 26% (for shorter peptides) within a period of 30 min.

Fractions were collected manually, dried and resolved in 5 µl H_2O containing 0,1% TFA. Chromatograms were analysed by the Gold Chromatography software Data System, version 1.6.

Subsequent analyses of collected fractions by positive ion matrix-assisted laser desorption ionisation (MALDI/TOF/MS) mass spectrometry and/or by LC/ESI-MS/MS spectrometry were performed.

2.2.4 Mass spectrometric analyses of processed peptides

Matrix-assisted laser desorption/ionization (MALDI/MS) is a soft ionisation technique used in mass spectrometry, allowing the ionisation of biomolecules such as proteins, peptides and sugars. MALDI/MS analyses were perfomed on a Bruker–Reflex IV instrument by positive ion detection in the reflectron mode. Ionisation of co-crystallized analytes was induced with a pulsed nitrogen laser beam (337 nm) and the ions were accelerated in a field of 20 kV and reflected at 23 kV. For MALDI/MS analyses 1 µl of the peptide or glycopeptide sample, resolved in H_2O with 0,1% TFA (v/v), was applied on a stainless steel or gold-plate target by mixing with the same volume of matrix. A saturated solution of 2,5-dihydroxybenzoic acid (DHB) containing 0,1%TFA in ACN (2:1 v/v) was used as a matrix.

Electrospray mass spectrometry (ESI/MS/MS) was performed on a Q-Tof 2 quadrupole time-of-flight mass spectrometer equipped with a Z spray source. Samples were injected into nanospray source by nanospray capillary. Results were evaluated by Mascot search engine.

2.2.5 Amino acid sequence analysis of generated fragments

Amino acid sequence of generated fragments was elucidated from the series of ions of ESI/LS/MS/MS fragmentation spectra or from the molecular masses of ions detected by MALDI/MS. The molecular masses detected by MALDI/MS were analysed by the FindPep software application from Expasy (www.expasy.org). This software identifies the amino acid sequences of peptides from their experimentally measured masses resulting from unspecific cleavage of proteins. The program takes into account chemical modifications, post-translational modifications (PTM) and autolytic cleavage of protease. The experimentally measured peptide masses are compared with the theoretical peptide masses calculated from the user-entered sequence or from an entry specified in the Swiss-Prot/TrEMBL database.

The suggested amino acid sequence was assigned to the detected molecular mass, if the deviation of the experimental and theoretical mass was not more than 1 Da. If several sequences corresponded to the same mass, all of them were taken into consideration.

2.3 Mathematical analyses

2.3.1 Quantification of generated peptide fragments

To quantify the amount of generated proteasome degradation products, we performed further analyses of MALDI/MS spectra and rpHPLC chromatograms. The mass spectrometric analyses of the HPLC fractions showed that several fragments co-eluted from the HPLC column at the same retention time.

Principally, we calculated the starting amount of substrate as 100% (Table 3). This amount was, due to proteasomal digestion, separated into several rpHPLC peaks. Each rpHPLC peak contained several co-eluting ion fractions, and the amount of substrate fragment found in one rpHPLC peak was estimated based on their relative intensities. In this way we obtained the relative amounts of each proteasome degradation product.

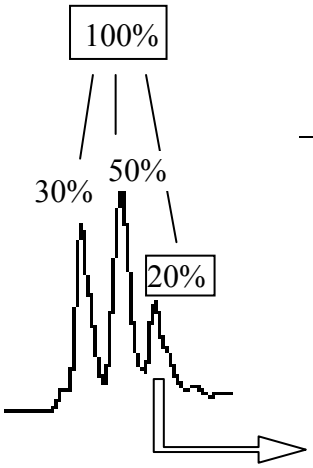

Table 3. Example of the calculation of relative peptide amount

Ion masses	Relative MALDI/MS ion peak intensity	Relative fragment amount = Percent of the starting substrate amount
1028Da	10%	2%
1436Da	50%	10%
1559Da	40%	4%
Sum	100%	20% (= the amount found in this HPLC peak)

Relative HPLC peak area was calculated as :

= peak area x 100 / sum of peak areas

Relative MALDI/MS ion peak intensity was calculated as:

=MALDI/MS ion intensity x 100/ sum of all ion intensities in analyses HPLC peak

Relative fragment amount was calculated as:

= relative HPLC area x relative MALDI/ MS ion intensity / 100

For example, the fraction 11 of P1 proteasome degradation HPLC chromatogram contained 5,74 % of the starting substrate distributed on 9 ion species, of which 5 (MW 1471.4; 1476.55; 1483.42; 1527.73; 2677.76) could be assigned to the MUC1 sequence (Table 4).

Measured MW	Corresponding MUC1 VNTR amino acid sequence	MALDI/MS ion intensity	Relative ion intensity	Relative amount of peptide
1471,40	PDTRPAPGSTAPPAH	11569,85	65,32575	3,749698
1473,05	*not MUC1 related*	2125,27	11,99971	0,688783
1476,55	TAPPAHGVTSAPDTR	316,9	1,789282	0,102705
1483,42	GVTSAPDTRPAPGSTA	553,3	3,124045	0,17932
1527,73	PDTRPAPGSTAPPAHG	174,95	0,987804	0,0567
2286,02	*not MUC1 related*	1656,85	9,354915	0,536972
2432,17	*not MUC1 related*	358,75	2,025576	0,116268
2677,76	PDT-PAPG	630,35	3,559086	0,204292
3294,64	*not MUC1 related*	324,79	1,833831	0,105262
SUM		17711,01	100	5,74

Table 4. Example of HPLC/MALDI/MS based quantification of generated fragments

The relative amount of fragment 1471,4 Da of relative ion intensity 65,3 can be calculated by comparison with total amount of fragments in the HPLC peak (5,74%) as follows:

5,74:100=x:65,3; resulting in 3,75 units.

Relative intensities of reaction products were calculated for all 29 processed MUC1 (glyco)peptide substrates.

2.3.2 Calculation of the frequency of digestion between two amino acids

The calculated relative amounts of fragments with identical N- or C-terminal motifs were summed up, and the higher one of these two sums (further on referred as “a”) was chosen for calculation of cleavage frequency at each peptide bond. The cleavage frequency was calculated by expressing the value “a” for each peptide bond as a percent of total “a” values in the substrate.

2.3.3 Calculation of average fragment length

The average length of generated fragments was calculated as the arithmetic mean of the fragment lengths.

2.3.4 Calculation of peptide length distribution

All detected fragments within one processed sample were grouped according to their length. The relative amounts of fragments within the same length group were summed up and expressed as the percentage from the total amount of all detected fragments within the sample, not counting the amount of non-processed substrate.

2.3.5 Calculation of relative amounts of short 8 to11-meric peptides

The relative amounts of fragments 8 to 11 amino acids long were summed up and their contribution to the complete amount of fragments expressed as a percentage.

2.3.6 Calculation of relative amounts of glycopeptides within 8 to 11-mers

The amount of glycopeptides within the group of 8 to 11mers was expressed as the percent of the total amount of 8 to 11meric peptides.

2.3.7 Analyses of the origin region of 8 to 11mers

8 to 11mers were sorted in three groups based on their starting motif:

1. Peptides starting N-terminally of the SAP motif within A-H-G-V

2. Peptides starting in proximity of the SAP region within T-S-A-P-D-T

3. Peptides starting within R-P-A-P-G-S- terminating often with unprocessed C terminus of the substrate.

The amounts of fragments for each of these three groups separate were summed up and expressed as the percentage of the total amount of 8 to 11mers.

2.3.8 Algorithm prediction of peptide ligation strength to MHC molecule

The amino acid sequences of all fragments in the size range of 8 to 11 amino acids (peptides and glycopeptides) were analysed by SYFPEITHI software application (http://www.syfpeithi.de/) without considering glycosylation. SYFPEITHI is a database comprising more than 4500 peptide sequences known to bind class I and class II MHC molecules. The entries are compiled from published reports only. The database contains information on peptide sequences, anchor positions, MHC specificity, source of proteins, source of organisms and publication references.

The database predicts the ligation strength of epitope to a defined HLA type. The algorithms used are based on the book "MHC Ligands and Peptide Motifs" by Rammensee H.G et al. The probability of being processed and presented is given in order to predict T-cell epitopes.

The probability of binding to the most common allele in Caucasian population HLA-A0201 allele for each MALDI/MS detected 8 to 11 amino acids long fragment sequence was screened. Additionally, the prediction of potential HLA-A0201-ligands from the complete

MUC1 VNTR sequence was performed. Peptides with the ligation strength higher than 9 were considered as good binders.

2.4 Immunological methods

2.4.1 Selection of short peptides

The peptide fragments predicted by SYFPEITHY to have high ligation strength were synthesized by Mimotopes/Perbio and in vitro glycosylated. Both the peptides found in proteasomal digests and the peptides that were just theoretically predicted to bind with strong ligation strength to HLA-A0201 were included. A group of peptides that could not be in vitro glycosylated due to restrictions of ppGalNAc-T4 enzyme was not included into further analysis although the peptides were predicted to be good binders. Glycosylation was performed as described above at VTSA and GST motifs based on the site-specificity of ppGalNAc-T1 and/or ppGalNAc-T2. The ligation strength of glycosylated peptides to HLA-A0201 was measured by MHC stabilisation assay using the human T2 cell line and compared to the ligation strength of their non-glycosylated forms and a viral reference peptide.

2.4.2 MHC stabilization assay

MHC stabilisation assay is based on the increased affinity of W6/32 antibody for the stabilised conformation of MHC molecules after binding of a peptide-ligand. The T2 cells express empty MHC molecules with short half life. Binding of the peptide ligand to the groove of MHC molecules stabilizes their tertiary conformation and prevents fast recycling. Stabilised tertiary conformation of MHC class I molecules is recognized by W6/32 anti-HLA-A, B, C monoclonal antibody. The binding affinity of W6/32 antibody directly reflects the number of stabilized MHC molecules upon binding of tested peptide. Thereby the affinity of the peptide can be measured and compared to that of negative and positive controls.

For MHC stabilisation assay T2 cells were grown in IMDM medium supplemented with 2% FCS. T2 cells were washed in PBS from the residual serum, and resuspended at 1x106/ml of IMDM supplemented with 0,1 % FCS and 25mM HEPES. Half of the million of cells were incubated in 48-well cell culture plate with 1 µg of ß2-microglobulin and 100 µg of peptide for 16 h at 27°C and 8 % CO2. Cells were washed with 5 ml PBS, centrifuged at 180xg for 5 min and the pellet resuspended in 10µl PBS with 1 µg W6/32 antibody and incubated at 4°C for 30 minutes. Subsequently cells were washed in PBS with 1%BSA (3x5 ml) and the pellet

resuspended in 50µl of secondary anti-mouse Alexa 488 antibody diluted 1:1000 in PBS with 0,1% BSA. Cells were incubated for 30 minutes in the dark, washed with PBS (3x5 ml) and finally resuspended in 500µl of PBS. Binding of W6/32 antibody to MHC class I molecules was analysed by FACS spectrometry, nanometres emission and absorption at the name of department. As a positive control hepatitis B virus core antigen HBVc18–27 P6Y~C (FLPSDCFPSV) was used. As a negative control T2 cells were used, to which no peptide was added.

Measured fluorescence stronger than that of the negative control was considered positive and its intensity was compared to the positive control. Results for tested peptides were expressed as the percent of the fluorescence measured for the positive control.

2.4.3 Cross presentation

The experimental approach based on cross presentation of antigen implies the introduction of antigen into the cytosol of the antigen presenting cells. As antigen presenting cells we used immortalized the mouse dendritic cell line DC2.4. We tested three modes to introduce the antigen into the cytosol: liposomes, electroporation, and incorporation of antigen into poly-lactic acid vesicles.

2.4.3.1 Culture of monocyte derived human dendritic cells

Monocyte derived dendritic cells were generated by positive selection of CD14+ monocytes of from human buffy coats on magnet assisted cell sorting system (MACS). Buffy coat is the fraction of human blood that contains mainly white cells. 50 ml buffy coat volume was separated on 4 fractions, each fraction diluted with PBS/2 mM EDTA till 30 ml and applied on 13 ml of Ficol. Upon 30 min centrifugation step on 900xg/18ºC without brake, the monocyte layer was collected; fractions merged by two, resuspended with PBS/2 mM EDTA up to 50ml and centrifuged for 10 min, 300xg, with brake on 18º. Cell pellet was resuspended in 25 ml of MACS buffer (PBS/2 mM EDTA, 0,5% BSA) and two fractions merged, centrifuged at 200xg/10 min/18ºC/brake. Pelleted cells were resuspended in 3 ml of ice-cold MACS buffer, counted and on $1x10^9$/ml 300 µl of MACS CD14 beads added. After short vortexing and 20 min incubation on 4ºC cells were diluted with additional 22 ml of MACS buffer, centrifuged (300xg/10 min, 18ºC) and the pellet resuspended in 3 ml of MACS buffer. The MACS column was attached to MIDI MACS system and washed with 2 ml of MACS buffer. Vortexed cell suspension was applied in 1 ml fractions over pre-wet MACS pre-separation filter on the column, washed 3x3 ml of MACS buffer. The column was detached from the magnet and the attached cells washed away by 3 ml of MACS buffer under the

syringe pressure. Vortexed cells were counted, centrifuges at 300xg/10 min and resuspended at $2x10^6$ cells/ml CellGro DC medium supplemented with 1xGlutamax, 800 U/ml GM-CSF and 500 U/ml IL-4. Cells were cultivated in 5 ml fractions in 6 well plate and on the 3rd and 5th day half of the media exchanged with fresh with preservation of cells. On the 7th day complete media was collected and exchanged for CelGro supplemented with GM-CSF, IL-4 as previously with addition of 20 ng/ml of TNF-alpha and liposome encapsulated antigen (100-400 µg of P1) and matured for 2-3 days. Supernatant was collected and exosomes isolated and analysed by mass spectrometry. Adherent cells were acid eluted and eluent purified by solid phase extraction and analysed by MALDI/TOF/MS.

2.4.3.2 Formation of Peptide-Loaded Liposomes

Liposomes are the smallest artificial vesicles of spherical shape that can be produced from natural non-toxic phospholipids and cholesterol. The liposomal vesicles can be used as drug carriers and loaded with a great variety of molecules, such as small drug molecules, proteins, nucleotides and even plasmids.

Liposomes were generated by hydration of lipids in round bottom glass flask rinsed with chloroform. Following lipids: cholesterol (10 µmol), POPC (15,17 µmol) and PEG-PE 2000 (1,5 µmol) (combined at 2 : 3 : 0,3 molar ratio), representing together 25 µmol of polar lipids (Ignatius R et al, 2000)) were added in 6 ml of dry chloroform : methanol solution (10 : 3 v/v). Lipids stored in chloroform solution were kept on ice to minimize oxidation and evaporation. The mixture was gently swirled between additions and sonicated briefly in bath sonicator to ensure mixing. The organic solvent was carefully evaporated using a gentle stream of dry nitrogen to produce a film on a glass. When the film became white, the flow was turned up to spread out last drops. The film was dried without thick clumps of precipitated lipid. The round bottom glass flask was dried a in vacuum desicator for 30 min and the lipid film resuspended in cyclohexane with 1-2% of ethanol to increase suspension in water. The solution was frozen on dry ice and lyophilised until dryness.

Dry lipid powder was suspended in 2ml of 150mM NaCl solution kept at 64°C (above the transition temperature of the lipids) with 10 nmol/ml of P1 peptide (400 µg) resulting in final lipid concentration of 2 mg/ml. The mixture was left to hydrate at 64°C for 2 h with slow rotation. The hydration procedure formed large multilamellar vesicles that were sized down by 4 freeze-thaw cycles (1 min liquid N2 followed by 2 h at 4°C). Formed multilamellar vesicles were converted to large unilamellar vesicles by extrusion at 64°C through the pre-

wetted 100 nm membrane for an odd number of times (more than 11) and the material deposited into a buffer rinsed glass vial.

The liposomes where separated from non-encapsulated antigen by size exclusion chromatography with 2 passages over a 15 x 1 cm Sepharose 4B column (section 2.4.3.1.1.) and the fractions containing liposomes detected by thin layer chromatography (section 2.4.3.1.2). Fractions containing liposomes were applied on $1x10^7$ human monocyte derived immature dendritic cells or on $4x10^7$ DC2.4 cells, incubated for 24 h under standard conditions, washed with PBS (3x5 ml) and acid eluted. The acidic eluent was collected, centrifuged at 3000xg for 10min and, ultrafiltrated over the 3000 Da cut-off membrane. The filtrate was purified by solid phase extraction on a C18 column, dried and analysed by MALDI/TOF/MS, all as previously described (section 2.2.4). Detected masses were analysed by FindPep software as described (section 2.2.5).

2.4.3.2.1 Size exclusion chromatography

A glass column (20 x 1 cm, 16 ml volume) was filled to the top with pre-swollen Sepharose 4B gel and packed with 50 cm/h flow of 150 mM NaCl for 30 minutes followed by 25 cm/h for 1 h.

The sample is applied and eluted by 1 ml/min flow of 150 mM NaCl. One ml fractions were collected and lipid presence was analysed by thin layer chromatography or by noticing the milky appearance of the solution.

2.4.3.2.2 Thin layer chromatography

A closed glass chamber was equilibrated for a minimum of 30 minutes by a mobile phase solution of a mixture of chloroform, methanol and water (130:50:8, v/v). Ten µl of each fraction was applied on the start line of the TLC (2 cm from the end) and air dried. As stationary phase silica plate 60WF254S precoated aluminium sheets, 20 x 20 cm were used. The plate was placed for 30 minutes in the glass chamber containing 1cm of the mobile phase, air dried and sprayed by 0,001% Primulin in 80% Acetone/Water. Detection was performed by exposure of dry plates to long-wave UV lamp (354nm).

2.4.3.2.3 Acid elution

Acid elution is performed according to a modified protocol from Storkus (Storkus WJ et al., 1993) with incubation of cells for 30-60 s in citrate-phosphate buffer:

0,131M citric acid,

0,066M Na2HPO4,

Adjusted with 0,1N NaOH to pH 3.3.

2.4.3.2.4 Solid phase extraction

Samples were purified from proteins and salts on one-way C18 columns connected to an evacuated chamber:

The column was equilibrated with 5 ml 80% acetonitrile in water with 0,1%TFA (v/v). The sample was pre-acidified by addition of TFA to pH2 and applied onto the column, and washed with 5 ml of 0,1%TFA in water (v/v). Bound peptides were eluted with 2-4ml 80% acetonitrile in water with 0,1%TFA (v/v).

2.4.3.3 Electroporation

DC2.4 cells (5x106) were grown in DMEM without phenol red, washed in PBS and resuspended in 0,8 ml PBS containing 5 mg ovalbumin. The cellular suspension was placed in an electroporation vial and incubated on ice for 10 min. Electroporation was performed by single pulse of 280 V / 960 µF for 30 sec with decay time constant 3-6 A, followed by 5 min incubation on ice. Electroporated cells were resuspended in 70 ml complete media, incubated for 24 h at 37°C and 5 % CO2, washed in PBS (3 x 5 ml) and acid eluted for 1 minute. The acid eluat was centrifuged, purified on C18 solid phase extraction as described (section 2.4.3.1.4) and analysed by MALDI/TOF/MS and by FindPep software.

The efficiency of the method was checked by pulsing 1x107 DC2.4 cells with FITC-labelled ovalbumin epitope SIINFEKL. Cells were resuspended in 800µl of PBS with 100µg SIINFEKL-FITC, and electroporated as described. After the electro-pulse the presence of peptide in the cytosol was monitored by a fluorescent microscope.

2.4.3.4 Formation of Peptide Loaded Poly-lactic Acid Vesicles:

Poly-lactic acid based vesicles were generated mixing two solutions:

Solution A: 200 nmol of P2 peptide (1 mg) resolved in 3 % cold water solution of poly-vinyl alcohol

Solution B: 5% poly-lactic acid in a mixture of ethanol and acetone (9:1, v/v)

Solution B (30 ml) was slowly added to solution A (150 ml) under constant mixing and stirring overnight at room temperature. Next day the mixture was centrifuged at 10.000xg for 10 min and washed three times with PBS (200 ml), followed by centrifugation steps as mentioned. The generated vesicles were resuspended in 330 µl of PBS and added to media of $4x10^8$ DC2.4 cells growing confluent in 9 cell culture flasks (300cm²). The cells were grown as described above (section 2.1.1), but with exosome-free FCS (section 2.4.3.3.1).

Cells were incubated with PLA-vesicles for 24 h at 37°C and 5% CO2. The cell culture supernatant was collected and exosomes were isolated as described (section 2.4.3.3.1). Isolated exosomes were resuspended in 100µl 0,1%TFA/H2O. Acidification of exosomes denatures tertiary MHC complex and releases bound peptide epitopes. The solution was shortly spinned and injected on an rpHPLC narrow-bore ODS column to chromatogram as described in section 2.2.2.1 for analytical chromatography. Fractions were collected manually, dried and analysed by MALDI/TOF/MS as described (section 2.2.4). Molecular masses of detected ions were analysed by FindPep software (section 2.2.5) and by nanospray MS/MS analyses (section 2.2.4).

2.4.3.4.1 Isolation of exosomes

Exosomes were isolated from FCS or cell culture supernatant by subsequent centrifugation steps: 10min/1200g; 30min/10000g; 30min/20.000g; 60min/100.000g. The exosomes were pelleted in the final centrifugation step. Exosomes were washed with PBS and used for further analyses. For the generation of exosome-free FCS, exosomal pellet was discarded and the supernatant sterile filtrated through 0,22 µm syringe filter.

2.4.4 Isolation of MHC bound peptides from breast cancer tumor cells

2.4.4.1 Testing the MHC Expression on breast cancer cell lines

MCF7 and T47D breast cancer tumor cell lines were analysed for the expression of MHC alleles by fluorescence-activated cell sorting (FACS) and confocal laser microscopy.

2.4.4.1.1 Confocal laser microscopy

Glass coverslips were cleaned over night by incubation in 70% ethanol with 5% HCl solution. Coverslips were placed on a parafilm covered glass plate to dry, washed with PBS and covered with cell suspension. A 300 µl of cell suspension ($1x10^6$) was applied per coverslip and cells were allowed to attach for 30 minutes. Thereafter, cells were washed with PBS from the residual medium and fixed by incubation in PBS containing 3% formaldehyde for 15 min. The fixing solution was aspirated and the coverslip washed with PBS containing few drops of 0.1M glycine/HCl, pH 8.0 (3x5 min). Fixed cells were incubated with 1µg of primary mouse monoclonal W6/32 antibody (0.1µg/ml PBS) for 30 min, washed with PBS (3x5 min) and covered with 100 µl of secondary anti-mouse Alexa 488 antibody diluted 1:1000 in PBS. Cells were incubated in the dark for 30 min, washed with PBS (3x5 min) and mounted on glass slides. For the mounting on glass slides, the coverslip was swirled once in deionised water and the

excessive water was soaked off with soft tissue paper. A drop of DAKO mounting medium was placed on the glass slide and the coverslip with cells facing the glass slide mounted avoiding formation of air bubbles. Mounted glass-slides were allowed to dry during the night on 4°C before they were analysed on fluorescent microscopy.
Visual analyses of stained cells were performed on a confocal laser scanning microscope. A 488 nm Argon laser for excitation of green flurophores of Alexa 488 antibody was used.

2.4.4.2 Isolation of MHC-Ligands from tumor cell lines by acid elution

Breast cancer cell lines T-47D and MCF7 were screened for MHC-bound MUC1 derived (glyco) peptides. The $8x10^7$ cells were grown under standard conditions in phenol red free media, washed with PBS and acid-eluted as previously described (section 2.4.3.1.3). Acid eluent was centrifuged and purified by solid phase extraction on C18 as described in section 2.4.3.1.4. The purified material was dried by vacuum centrifugation, resuspended in 50 µl of 2.5 mM phosphoric acid and analysed by coupled cation exchange /reversed phase liquid chromatography/MS/MS in the central service facility of the Center for Molecular Medicine Cologne. Alternatively, MHC molecules or MHC molecules were affinity purified by HLA-A, B, C antibody W6/32. The eluted epitopes were analysed by ESI/MS/MS analyses.

2.4.4.3 Isolation of MHC ligands from tumor cells by affinity chromatography

T-47D cells (8x107) were washed with PBS and collected by cell scraping. After centrifugation at 180xg for 5 min the cell pellet was lysed for 1 h at 4°C in 1 ml of lysis buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 1,5 % Nonidet-N40, 0.2 % Sodium azide and protease inhibitor Complete mini. The lysate was ultracentrifuged at 100.000xg for 30 min and HLA molecules purified by affinity chromatography with the W6/32 antibody on a protein-A column. The protein A column (2 ml) was packed according to the manufacturers' protocol, with PBS used as packing buffer and 0,05 M Tris-HCl, 3 M NaCl, pH 7,8 as start buffer. The column was washed with start buffer and cell lysate applied continuously during the night at 4°C with a flow rate of 0,3 ml/h using a peristaltic pump. Unbound material was washed successively by 10 ml of start buffer, 10 ml of lysis buffer, 10 ml of lysis buffer without detergent, 10 ml of lysis buffer without detergent and NaCl. The bound protein was eluted with 9,1 ml of 0,2 M acetic acid. The MHC molecules were denaturated by addition of 0,9 ml of concentrated acetic acid and incubated for 30 min on ice. The released peptide epitopes were separated from the protein fraction by ultrafiltration on a 3000 Da cut-off membrane. The filtrated peptide fraction was purified on C18 SPE as described, dried and analysed by mass spectrometry.

3 RESULTS

3.1 Biochemical studies

3.1.1 Influence of buffer composition on processing efficiency

Peptides and glycopeptides were processed by immunoproteasomes in parallel using three different buffers, as described in section 2.2.3.2, and processing efficiency was analysed by rpHPLC (Fig. 13). The most effective degradation of glycopeptides, indicated by the amount of starting substrate remaining after 24h digestion time, was achieved with buffer C (containing DTT), whereas nonglycosylated peptides were most efficiently processed in Buffer B (DTT free) as shown on figure 13. The efficient use of buffer A in *in vitro* proteasomal assays reported by Groettrup M was not confirmed for MUC1 VNTR peptides and glycopeptides (Groettrup M et al., 1995).

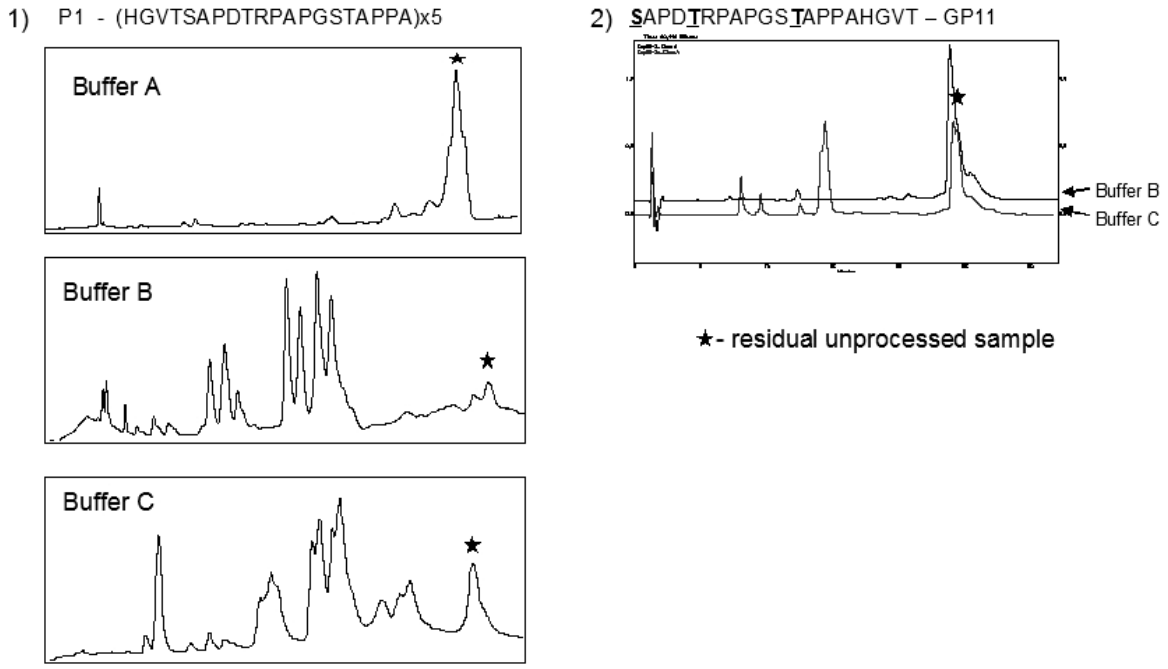

Figure 13. Influence of buffer composition on processing efficiency of MUC1 peptide and glycopeptide by immunoproteasomes. Digestion of MUC1 P1 peptide (1) and GP11 glycopeptide (2) was performed in different buffer systems and analysed by rpHPLC. (1) rpHPLC of IP digests of non-glycosylated 100mer P1 in three tested buffers shows highest amount of unprocessed substrate in Buffer A, and lowest in buffer B (without DTT). This indicates that processing of non-glycosylated substrates is most effective in HEPES buffer without DTT (Buffer B). (2) The part 2 shows two overlaid rpHPLC chromatograms of IP digests of glycosylated GP11 substrate,. The amount of nonprocessed substrate is higher under buffer B conditions, while the addition of DTT improves digestion efficiency.

3.1.2 Control of proteasome purity

To control the purity of purchased proteasome preparations, the P1 peptide as substrate was incubated with immunoproteasome in the presence of proteasome specific inhibitor. A parallel control assay without proteasome inhibitor was performed. Digests were fractionated by rpHPLC and fractions analysed by MALDI/MS/TOF/MS.

Upon addition of proteasome inhibitor no processing activity was noticed (Fig. 14). This indicates purity of the proteasomal preparation and absence of contaminating proteases.

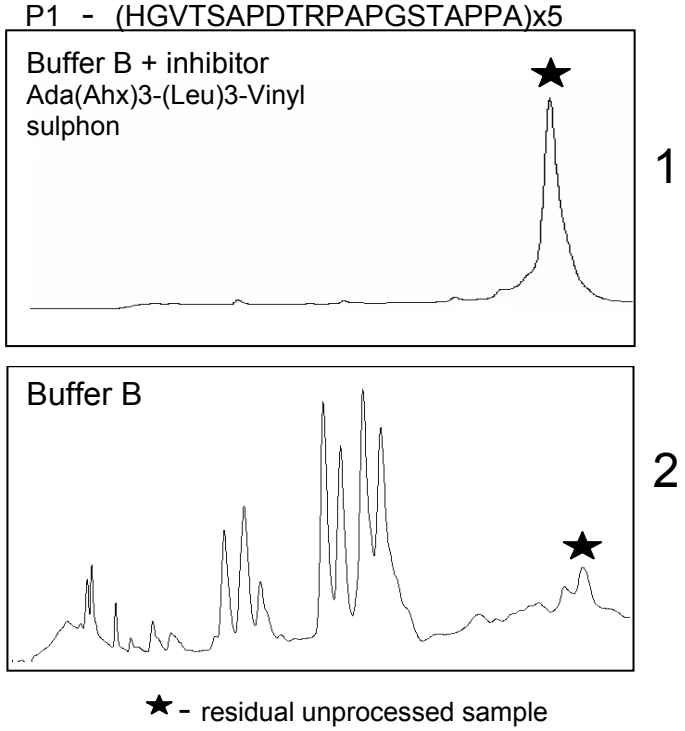

Figure 14. The control of proteasome purity by rpHPLC of P1 proteasome digest. The P1 100-mer was processed by immunoproteasomes in buffer B in the presence of proteasome inhibitor (1). The result was compared with a control experiment without addition of inhibitor (2). The single peak present in the chromatogram 2 represents the non-processed P1 peptide (marked by asterisk) and proves the complete inhibition of proteasomal activity and the absence of contaminating proteases.

3.1.3 Influence of glycosylation on proteasomal processing efficiency

To test the influence of glycosylation on proteasomal processing efficiency, the array of 29 MUC1 peptides and glycopeptides was *in vitro* processed by immunoproteasomes (Table 2, section 2.2.1). We used a labelling system by sorting of glycopeptides based on their glycan chain length. Unglycosylated peptides are named systematically starting with P followed by the number, glycopeptides with GP when glycosylated with the monosaccharide GalNAc, with GGP when glycosylated with the disaccharide Gal-GalNAc and as SGGP when carrying sialylated disaccharide. Previously used, unsystematic, names are mentioned in brackets.

The substrates represented sections of the MUC1 tandem repeat domain starting with different motifs of the repeat (AHG or SAP). Twenty five of them were O-glycosylated with single or multiple GalNAc, Gal1-3GalNAc or NeuAc2-3Gal1-3GalNAc moieties at Ser/Thr positions. *In vitro* proteolysis with immunoproteasomes was performed until the reference substrate, a non-glycosylated 100-meric repeat peptide (P1), was completely processed and no starting peptide detectable. Generated fragments were chromatographed by rpHPLC for identification and quantification. The amino acid sequences of fragments were identified either by MALDI/MS-TOF mass spectrometry *via* their molecular mass ions or by ESI-Q TOF-MS/MS *via* their b_i or y_i fragment ions (Figure 15). Two independent assays were performed with most of the substrates and showed close similarity with respect to proteolytic fragmentation patterns and relative amounts of fragments, thereby proving the reproducibility of the assay. The relative amounts of all generated fragments were calculated based on relative HPLC peak areas and MALDI/MS ion intensities, (section 2.3.1). The rpHPLC chromatograms and tables listing digestion rates for all analysed peptides and glycopeptides are attached in appendix.

3.1.3.1 Processing of nonglycosylated peptide

Digestion of all non-glycosylated peptides occurred at high rates, resulting in quantitative degradation of substrates during 48h, independent of their sizes (21-mer to 100-mer) and irrespective of sequence variations at the immunodominant motif (DT > ES). During the same time period digestion of all glycopeptide substrates resulted in more or less reduced degradation rates that varied from <1 to 90% (Table 4).

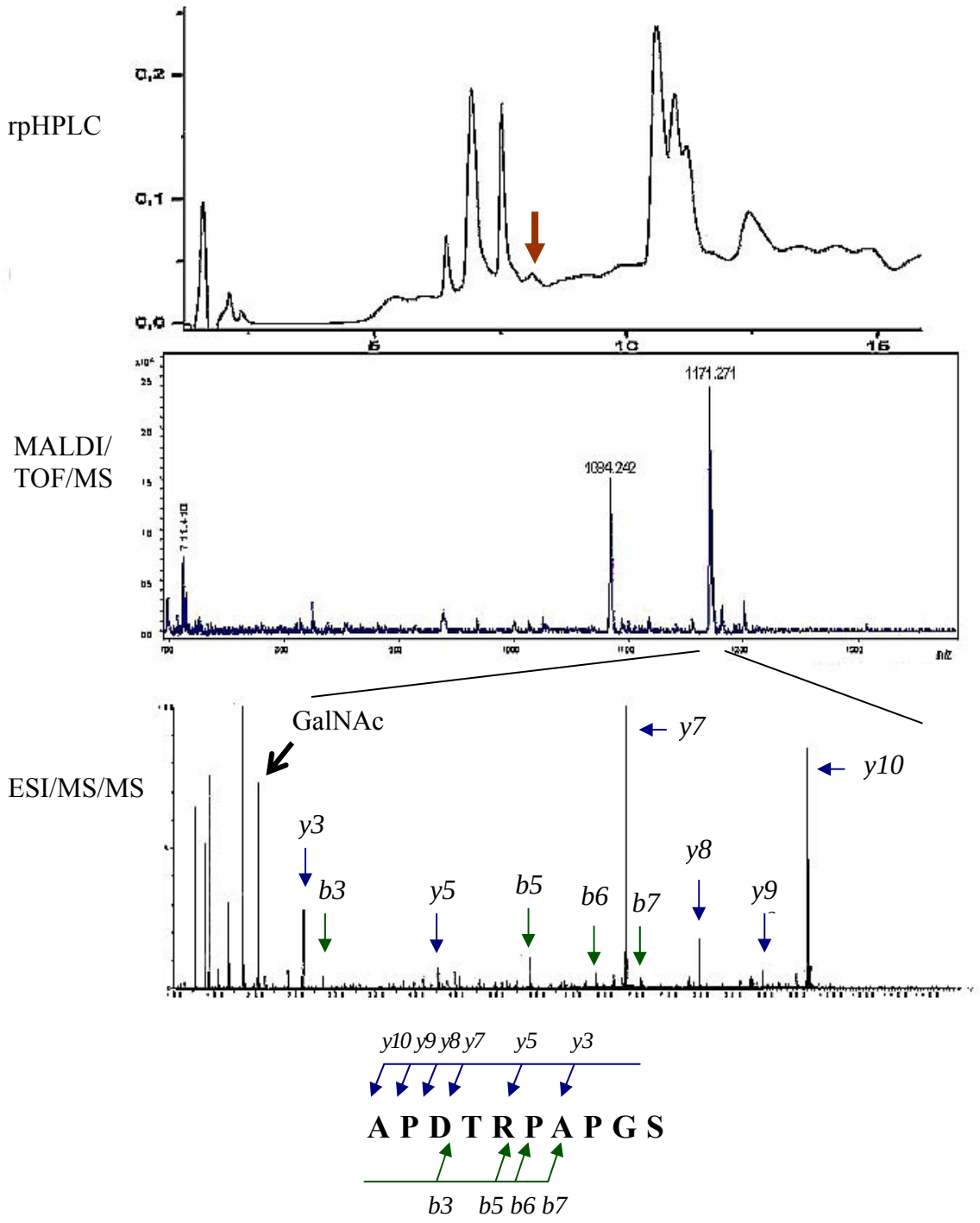

Figure 15. Analysis of digestion mixture of GP3 glycopeptide by rpHPLC, MALDI/TOF/MS and ESI/MS/MS. The HPLC fraction labelled with red arrow was measured by MALDI/TOF/MS. Ion of MW 1171 was sequenced by ESI/MS/MS and APD**T**RPAPGS-GalNAc glycopeptide identified according its b_i and y_i fragment ions

3.1.3.2 Processing of glycosylated peptides

Two series of glycopeptides with different starting motif: SAP or AHG were used in our experiments. The 20-meric glycopeptides (GP9-GP14) starting with SAP amino acid motif were degraded at 50-70% rate of the nonglycosylated peptide rate even when carrying three, two or one glycan moieties. In this series of glycopeptides no direct correlation between number of glycan chains per repeat and proteolysis efficiency was revealed, since these 20-mers were degraded at comparable rates (see section 3.1.4.2.2).

On the other hand, glycosylation of AHG series glycopeptides provoked a strong specific impact on the degradation rate. The heavily O-glycosylated 20-meric glycopeptide GP16 resembling one tandem repeat of the fully glycosylated, naturally occurring MUC1 glycoforms in tumor patients was revealed as the poorest substrate with a degradation rate of <1% (Fig. 16).

Reduction of glycosylation to three GalNAc moieties per repeat on 100-meric GP1 glycopeptide did not improve the processing efficiency. The length of the densely glycosylated substrate had no influence on the degradation rate.

Strikingly, the 20-meric glycopeptide GP15 with only two GalNAc residues was also fully resistant to IP proteolysis pointing to site-specific effects of O-glycosylation (Fig. 16).

Noticeably, the digestion of monoglycosylated peptides GP2-GP7 occurred at 60-70% rate of the nonglycosylated peptide, irrespective of sequence variation (DT>ES). No clear-cut evidence could be obtained that variation of single glycosylation sites did strongly influence the overall processing efficiency (Table 4).

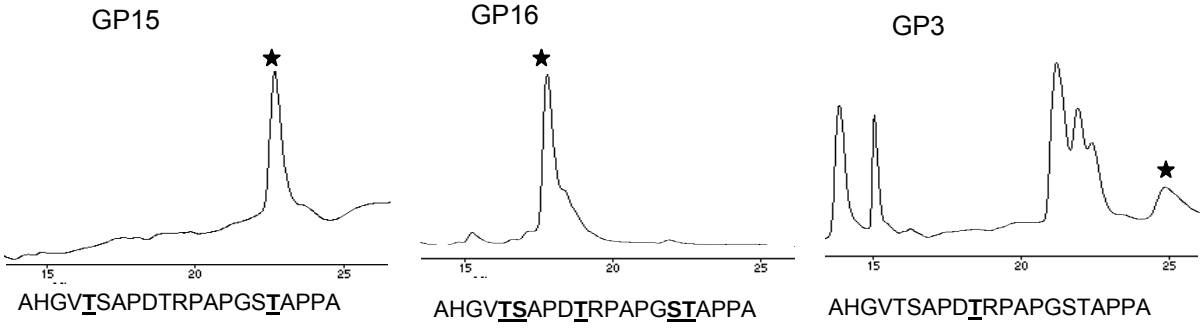

Figure 16. Influence of glycan presence on processing efficiency. Glycosylated MUC1 VNTR peptides GP15, GP16 and GP3 were processed by IPs and analysed by rpHPLC. The residual unprocessed peptide is marked by an asterisk and positions of glycosylation within amino acid sequences by bold underlined letters.

3.1.3.2.1 Influence of glycan length on processing efficiency

We addressed the question whether more elongated glycans, like the disaccharide Gal1-3GalNAc or NeuAc2-3Gal1-3GalNAc, could have negative effects on processing rates compared to substrates glycosylated with the core sugar GalNAc. Sequence identical peptides with the AHG starting motif (GP2-GP4; GGP1-GGP3; SGGP1-SGGP5) mono-substituted with any of the three glycans were assayed. Strikingly, rpHPLC analyses revealed higher degradation rates for the disaccharide substituted peptides than for complementary monosaccharide substituted peptides (Table 4). The replacement of GalNAc by Gal(1-3)GalNAc at the same peptide position increased the degradation rate by 30% (compare GGP1 *vs.* GP2; GGP3 *vs.* GP4).

On the other hand, sialylated glycopeptides were processed at very low rate, comparable to the densely glycosylated GP1 and GP16 glycopeptides.

Together these data indicate that glycopeptide antigens can be processed by immunoproteasomes, but the process is influenced by localisation and type of glycan chains. Furthermore, naturally occurring densely glycosylated MUC1 is a poor substrate for immunoproteasomes due to its abundant glycosylation. The reduction of glycosylation density or change of the peptide starting motif plays important roles in *in vitro* processing of MUC1 VNTR peptides. The influence of the peptide starting motif on processing rates could be explained by the specificity of proteasomal digestion, which together with glycan substitution influence digestion outcome. To analyse the specificity of proteasome for MUC1 repeat sequence, further studies were performed.

Peptides table

	NAME		SEQUENCE	% PROCESSED
PEPTIDES	P1(DTR100)		(HGVTSAPDTRPAPGSTAPPA)x5	100
	P2(ESR61)		A(HGVTSAPESRPAPGSTAPPA)x3	100
	P3(AHG21DTR)		AHGVTSAPDTRPAPGSTAPPA	100
	P4(AHG21ESR)		AHGVTSAPESRPAPGSTAPPA	100
GLYCOPEPTIDES	MONOSACCHARIDE	GP1(Tn100)	(HGV**TS**APDTRPAPG**ST**APPA)x5	<1
		GP2 (H11)	AHGV**T**SAPDTRPAPGSTAPPA	57
		GP3(H12)	AHGVTSAPD**T**RPAPGSTAPPA	94
		GP4(H13)	AHGVTSAPDTRPAPGS**T**APPA	63
		GP5(H14)	AHGV**T**SAPESRPAPGSTAPPA	65
		GP6(H15)	AHGVTSAPE**S**RPAPGSTAPPA	74
		GP7(H16)	AHGVTSAPESRPAPGS**T**APPA	64
		GP8(A4GALNAC)	AHGVT**S**APESRPAPGSTAPPA	70
		GP9(SAP20-1)	**S**APDTRPAPGS**T**APPAHGVT	54
		GP10(SAP20-2)	**S**APD**T**RPAPGSTAPPAHGVT	54
		GP11(SAP20-3)	**S**APD**T**RPAPGS**T**APPAHGVT	65
		GP12(SAP20-4)	SAPD**T**RPAPGSTAPPAHGVT	56
		GP13(SAP20-5)	SAPD**T**RPAPGS**T**APPAHGVT	74
		GP14(SAP20-6)	SAPDTRPAPGS**T**APPAHGVT	68
		GP15(A11)	HGV**T**SAPDTRPAPGS**T**APPA	<1
		GP16(A12)	HGV**TS**APD**T**RPAPG**ST**APPA	<1
	DISSACCH.	GGP1(A1)	AHGV**T**SAPDTRPAPGSTAPPA	91
		GGP2(A2)	AHGVTSAPD**T**RPAPGSTAPPA	89
		GGP3(A3)	AHGVTSAPDTRPAPGS**T**APPA	88
		GGP4(A4)	AHGVT**S**APDTRPAPGSTAPPA	86
	SIALYL.	SGGP1 (A4ST)	AHGV**T**SAPDTRPAPGSTAPPA	<1
		SGGP2 (A13ST)	AHGV**TS**APDTRPAPGSTAPPA	<1
		SGGP3 (A14ST)	AHGVTSAPDTRPAPG**ST**APPA	<1
		SGGP5 (A8ST)	AHGV**TS**APD**T**RPAPG**ST**APPA	<1

Table 4. List of synthetic peptides and glycopeptides used in experimental studies. Synthetic peptides and glycopeptides were incubated with 20S IPs and the products were chromatographed by rpHPLC to quantify the residual amount of non-processed peptide. Peptides and glycopeptides, named according to a systematic labelling system are grouped based on their glycosylation type. The glycosylated amino acids in the peptide sequences are labeled with bold underlined letter.

3.1.4 Influence of glycosylation on processing pattern of MUC1 VNTR sequence

3.1.4.1 Processing pattern of nonglycosylated peptide

The processing pattern of MUC1 VNTR peptides was elucidated from the frequencies of cleavage at each peptide bond within the analysed substrates. The influence of glycosylation on processing patterns was evaluated from comparative analyses of preferred cleavage sites for different glycopeptides and for peptides. Between two independent digests of the same peptide no difference in cleavage patterns was observed.

The most frequent cleavages of nonglycosylated 100meric P1 peptide occurred within the S-A-P motif contributing to 50% of all proteolytic products (Fig. 17). The high specificity of IPs for this motif is emphasized by the fact that other peptide bonds are cleaved at less than 1% rates. This confirms previous reports claiming that the relative usage of processing sites can differ substantially. (Niedermann G et al., 1996; Toes RE et al., 2001).

The second frequently used cleavage region is G-S-T releasing about 9% of generated fragments.

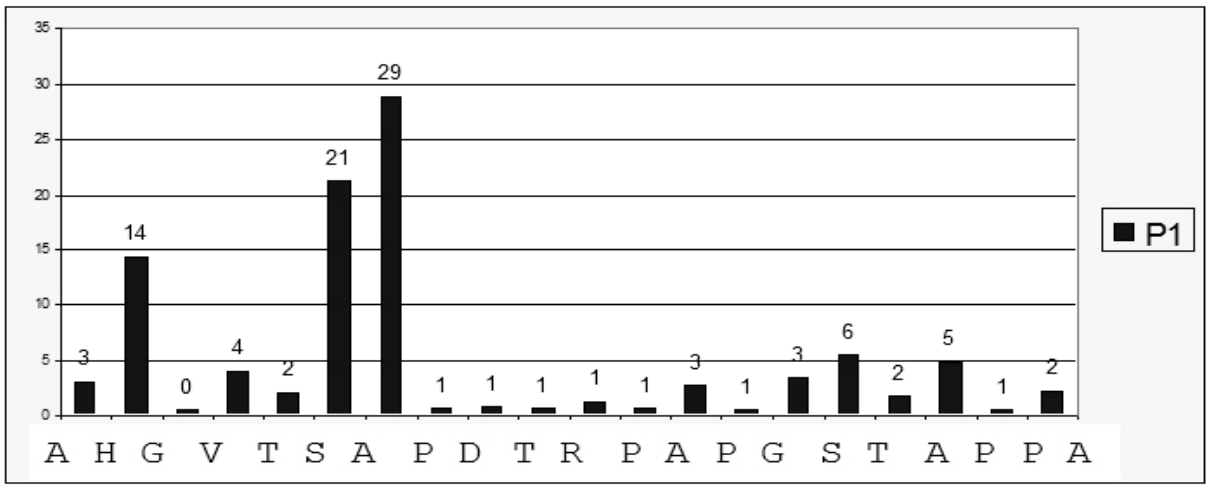

Figure 17. The digestion frequencies of peptide bonds of nonglycosylated P1 (100mer). The P1 peptide was digested by IPs and the cleavage frequencies between its amino acids calculated. The histogram bars show cleavage frequency between amino acids. The amino acid sequence is given on the x-coordinate. The numerical values of cleavage frequencies are shown by the numbers above the bars. The processing of long nonglycosylated P1 peptide occurs most frequently within the SAP motif as well as within the GST motif.

The processing pattern depicted by the digestion frequencies is reflected by the pool of major peptide fragments. Fragments generated during proteolysis of P1 peptide originate from the digestions in the S-A-P and/or G-S-T region, as shown in figure 18.

Figure 18. The most prominent peptide fragments generated by immunoproteasomal cleavage from P1 peptide. Fragments are presented as the horizontal bars and most frequent terminating positions squared by the box. The majority of generated fragments terminate in the SAP or GST regions.

3.1.4.1.1 Influence of variant amino acid sequence of VNTR region on processing patterns

The introduction of variant amino acid sequence (DT>ES) into the 61meric P2 peptide did not substantially change the processing pattern of immunoproteasomes, as presented on figure 19.

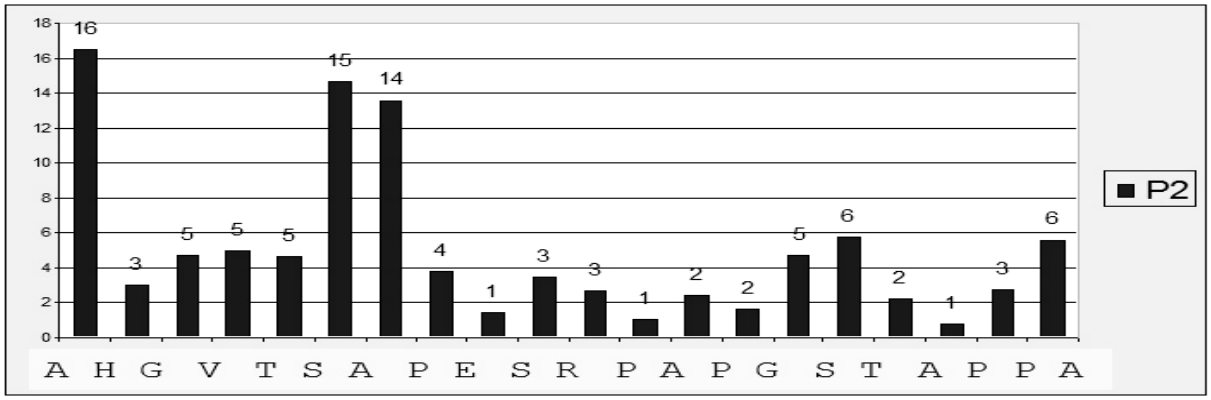

Figure **19. The digestion frequencies of peptide bonds of nonglycosylted P2 61-mer.** The 61-meric P2 peptide was digested by IPs and the processing frequencies between its amino acids calculated. The histogram bars show cleavage frequencies between two amino acids. The amino acid sequence is given on the x-coordinate. The numerical value of cleavage frequencies are shown by the numbers above the bars. The processing of long nonglycosylated sequence-variant P2 peptide occurs most frequently within the SAP motif as well as within the GST motif.

Moreover, the shorter nonglycosylated substrates, like the 21-meric peptides P3 and P4, are also processed in identical manner as longer substrates. The S-A-P and G-S-T motifs are for both peptides the major cleavage regions. These data confirm that naturally existing variation of the amino acid sequence does not considerably influence processing of the nonglycosylated MUC1 VNTR core protein (Fig. 20).

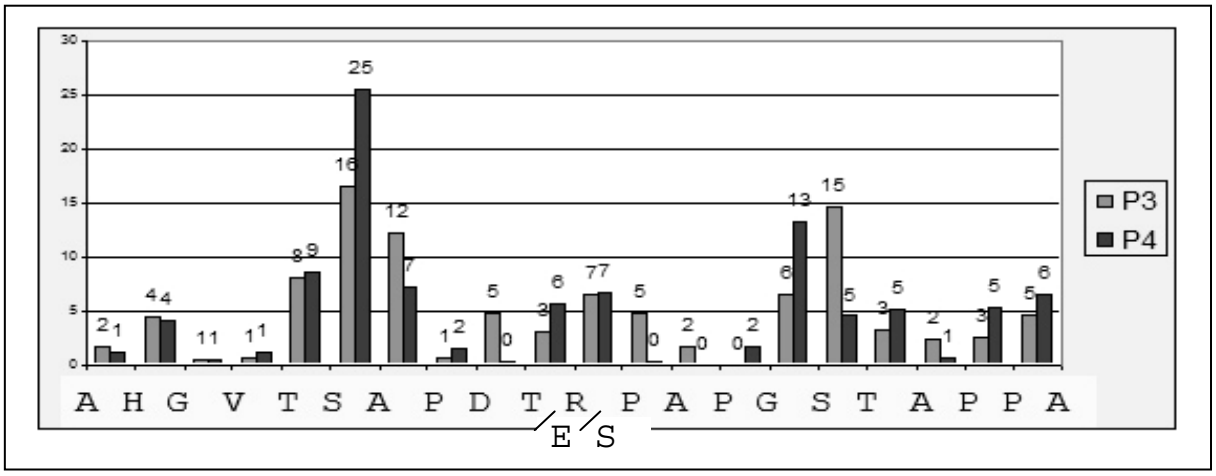

Figure 20. The cleavage frequencies of peptide bonds of nonglycosylted P3 and P4 20-mers. The 20-meric P3 and P4 peptide were digested by IPs and the cleavage frequencies between their amino acids calculated. The histogram bars show cleavage frequencies between two amino acids. The amino acid sequence is shown on the X-coordinate. The numerical value of cleavage frequencies are shown by the numbers above the bars. The processing of short nonglycosylated peptides occurs most frequently within the SAP motif as well as within the GST motif. Digestion frequencies of sequence variant peptide P4 (DT>ES) is comparable to frequencies of P3 peptide.

From these data we conclude that immunoproteasomes can digest all peptide bonds of the MUC1 VNTR sequence, but show strong preference for peptide bonds within SAP and GST motifs. This processing pattern is irrespective of the peptide length or naturally existing sequence variance (DT>ES).

3.1.4.2 Processing pattern of glycosylated peptide

Abundant O-glycosylation of MUC1 is regarded as the reason for its weak immunogenicity by sterically blocking endosomal processing. We assessed whether and how the presence and position of O-glycans within MUC1 VNTR peptides influence the proteasomal processing specificity and efficiency. Previosly we identified SAP and GST motifs as the major processing regions in MUC1 VNTR peptide. Interestingly, the immunodominant peptide epitopres of MUC1 are located within the VNTR domain, which is generally heavily O-glycosyltated at 5 positions, and four of which overlap with the defined major processing sites.

As expected, the presence of glycans at these major processing regions significantly reduced the cleavage rates by IPs, as shown in figure 21. Glycosylation of Thr adjacent to the SAP motif significantly reduced the relative frequencies of cleavage within the motif to about 8% compared to app 50% observed in nonglycosylated P4 peptide (refer to glycopeptide GP5 and the nonglycosylated reference peptide P4 in figure 21. Glycosylation at this position did not significantly change cleavage frequencies at the distant GST region.

Vice versa, glycosylation of Thr within the GST motif reduced cleavage rates at this region from 8% to 3%, while the cleavage frequency within the SAP region remained largely unaffected compared to the reference peptide (refer to glycopeptide GP7 and P4 peptide). The negative influence of glycosylation is in this case less striking than the one noticed for the SAP region.

Glycosylation of the Thr within the DTR motif located proximal to both preferred cleavage sites did not change the proteasomal cleavage pattern (refer to glycopeptide GP6 versus peptide P4in figure 3).

Figure 21 further depicts major processing products of GP5, GP6 and GP7 mono-glycosylated peptides in comparison to nonglycosylated P4 peptide. It is noticeable that glycosylation of threonine adjacent to the SAP motif reduces the amount of fragments generated by digestion within the SAP region (light blue bars), whereas glycosylation of threonine within GST region reduces the amount of fragments terminating within the GST region (dark red bars). As the SAP region contributes to 50% of generated fragments, compared to maximally 8% reported for GST motif, its glycosylation has much stronger effect on processing efficiency. The glycosylation of central threonine within the DTR motif does not change proteasomal processing patterns and generates major fragments comparable to nonglycosylated reference peptide P4 (grey bars).

From these data it can be concluded that the cleavage frequencies at the major processing regions are affected by the positions of O-glycans. Precisely, presence of O-glycosylation reduces cleavage frequency at the peptide bonds in its close proximity, without affecting more distant positions. Simultaneous presence of glycans at both preferred regions of cleavage therefore prevents IP proteolysis almost completely. This result explains low digestion efficiency of glycopeptides GP15, GP16 and GP1 that are glycosylated at both (T)SAP and GST motifs as well as low cellular immunogenicity of naturally occurring cancer associated form of MUC1.

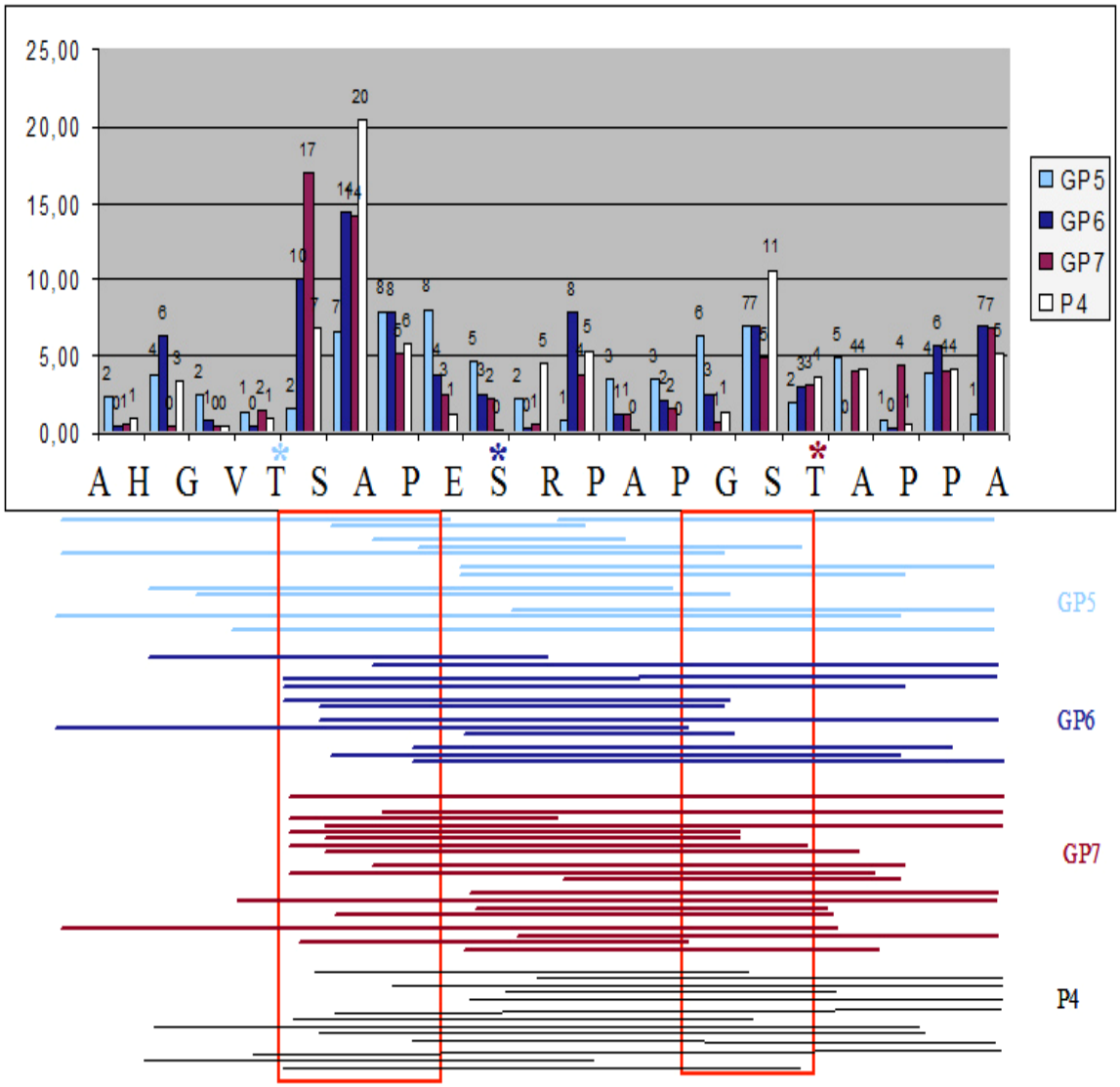

Figure 21. Frequency of digestion between each two amino acids of three differently glycosylated peptides and major generated fragments. The GP5, GP6 and GP7 glycopeptides were digested by IPs in parallel to nonglycosyalted P4 peptide and the processing frequencies between their amino acids were calculated. The histogram bars show cleavage frequencies between two amino acids. The amino acid sequence is shown on the x-coordinate. The numerical value of cleavage frequencies are shown by the numbers above the bars. Cleavage frequencies of glycopeptides were compared to values calculated for non-glycosylated P4 substrate. The frequencies of GP5 glycopeptide with GalNAc on Thr5 show significant reduction of cleavage within the SAP region (light blue bars). The glycosylation at the central ESR region does not change the pattern of immunoproteasomal activity (dark blue bars). Presence of GalNAc at Thr17 within the GST region of GP7 slightly reduces the cleavage within this region. Major processing fragments are presented as the horizontal bars in the corresponding colours and most frequent terminating positions squared by the orange box. The majority of generated fragments terminate in the SAP or GST regions.

3.1.4.2.1 Influence of glycan length on processing pattern

As notices previously, peptides glycosylated with disaccharide chains were processed at even increasing processing rate compared to GalNAc substituted peptides. To study weather the replacement of a monosaccharide (GalNAc) by a disaccharide (Gal1-3GalNAc) introduces changes in proteasomal site specificity, we performed parallel assays with peptides glycosylated at identical positions with either of these sugars. We noticed no change of cleavage specificity upon introduction of more elongated glycans. Major fractions generated from sequence identical peptides GGP2/GP3 and GP8/GGP4 mono-substituted at identical positions were terminated within SAP and /or GST regions (Fig. 22).

3.1.4.2.2 Influence of starting motif on processing pattern

The elucidation of MUC1 VNTR processing patterns presents the base for understanding the difference in processing efficiency of glycopeptides with different starting motifs, reported in section 3.1.3.2. Glycopeptides starting with the SAP motif were efficiently digested by IPs irrespective of the number of glycan chains per repeat. As the major processing region SAP is, in this case, located at the N terminus of the peptide, that major processing region was not used by IPs and therefore no block by O-glycosylation was noticed.

In this series of experiments we addressed the question wether short peptides starting with another repeat motif generate a principally different pool of fragments and if the starting motif influences proteasomal site specificity. Glycopeptides with AHG or SAP starting motifs were assayed in parallel and the released peptide products identified. As shown in figure 22 major processing products in both cases are mainly terminated in SAP or GST region indicating that proteasome cleavage site specificity is not influenced by substrate starting motif.(Fig. 22).

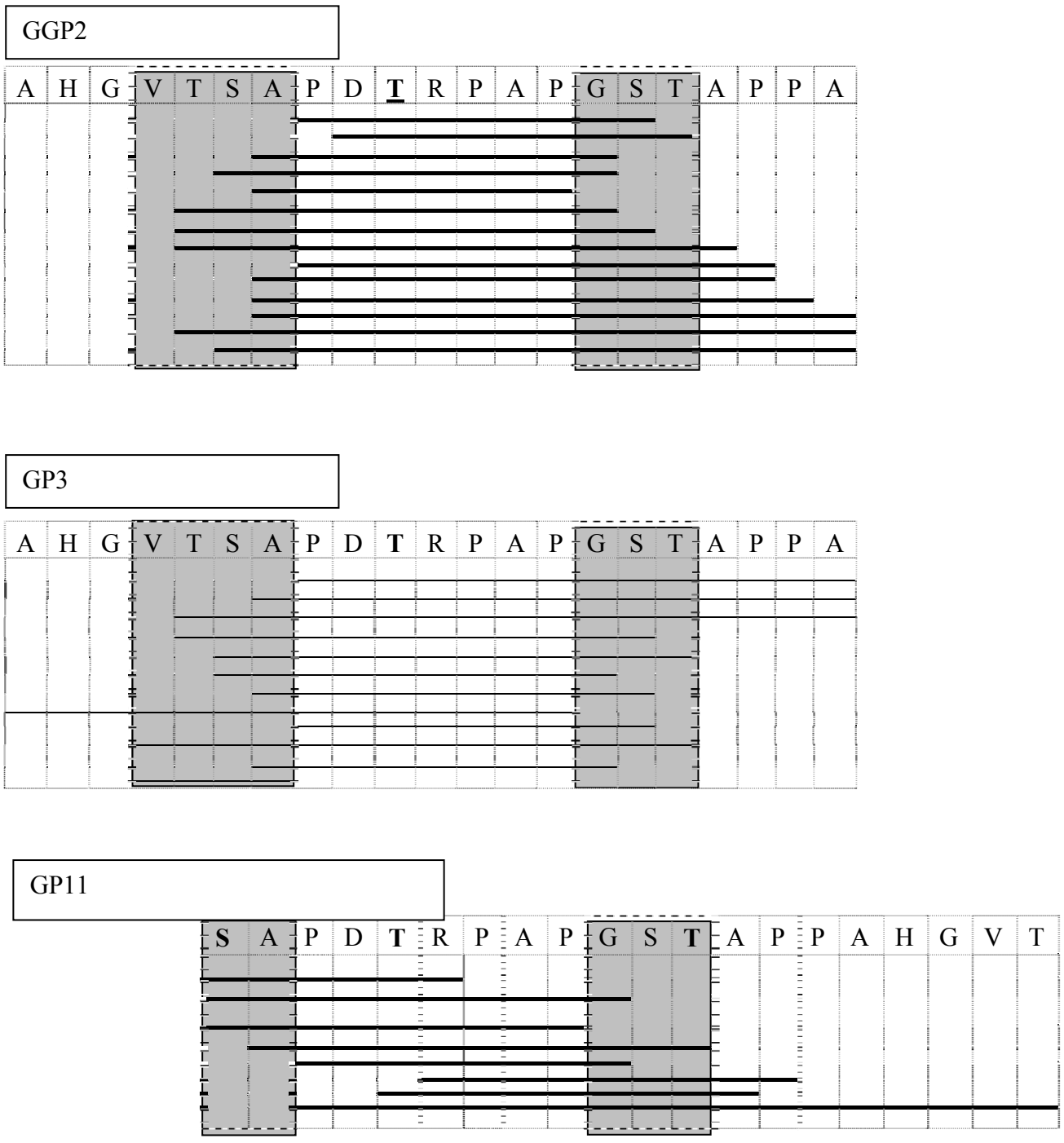

Figure 22. The most prominent peptide fragments generated by IP digestion from glycopeptides with different starting motif and glycosylation type (GGP2, GP3 and GP11). Fragments are presented as horizontal bars and most frequent terminating positions squared by the grey-shaded box. The majority of generated fragments terminate within the SAP or GST region. The type of glycan or the shifted starting motif did not cause changes in proteasome specificity. Glycosylated amino acids are marked with bold letters and underlined in the case of disaccharide substitution.

3.1.5 Analyses of generated short fragments

3.1.5.1 Average peptide length and number of different fragments

After defining major processing regions within the MUC1 VNTR sequence, explaining the influence of glycosylation on proteasome specificity and digestion outcome and identifying tolerated glycan positions and type, we approached the problem of best suited glycopeptides for eventual breast cancer vaccination strategies. These glycopeptides should resemble cancer related MUC1, should be processable by immunoproteasomes and be the source of MHC fitting glycosylated fragments/epitopes.

To determine which glycopeptides could be possible sources of MHC fitting glycosylated fragments, we analysed the total number of generated fragments as well as the average fragment length generated by IPs digestion (Table 5).

In the digest of a nonglycosylated MUC1-100mer by immunoproteasome 105 different peptides with a mean fragment length of 25,2 amino acids were identified. The digest of shorter MUC1 derived peptide AHG21 contained 30 different peptides with a mean fragment length of approx. 11 amino acids. The digestion of AHG series glycopeptides substituted with GalNAc released between 45 and 63 different peptides. The general mean fragment length in the digest of shorter substrates was 11 amino acids.

NAME		SEQUENCE	Nr of fragments	Average fragment length
PEPTIDE	P1(DTR100)	(HGVTSAPDTRPAPGSTAPPA)x5	105	25,2
	P2(ESR61)	A(HGVTSAPESRPAPGSTAPPA)x3	105	17,5
	P3(AHG21DTR)	AHGVTSAPDTRPAPGSTAPPA	29	10,8
	P4(AHG21ESR)	AHGVTSAPESRPAPGSTAPPA	31	11,6
GLYCOPEPTIDE	GP2 (H11)	AHGV**T**SAPDTRPAPGSTAPPA	63	11,7
	GP3(H12)	AHGVTSAPD**T**RPAPGSTAPPA	33	10,9
	GP4(H13)	AHGVTSAPDTRPAPGS**T**APPA	47	12,9
	GP5(H14)	AHGV**T**SAPESRPAPGSTAPPA	46	11,6
	GP6(H15)	AHGVTSAPE**S**RPAPGSTAPPA	46	10,6
	GP7(H16)	AHGVTSAPESRPAPGS**T**APPA	61	11,7
	GP8(A4GALNAC)	AHGVT**S**APESRPAPGSTAPPA	46	10,5

Table 5. Number of fragments generated by IPs and average fragment length in the digest of MUC1 VNTR peptides and glycopeptides. Nonglycosylated and glycosylated peptides with AHG starting motif were processed by IPs, and analysed by rpHPLC and MALDI/MS mass spectrometry followed by identification and quantification of fragments. We found that shortening of substrate results in shorter average fragment length. 20-meric peptides are releasing 11 amino acids long fragments in average, thereby enriching the amount of short MHC fitting epitopes.

From these results it can be concluded that average length of fragments is proportional to the substrate length. Processing of short 21-meric peptides resulted in shorter average fragment

lengths. The shorter mean fragment length increases the probability that a higher number of MHC fitting glycosylated epitopes is generated.

3.1.5.2 Amount of generated 8 to 11-meric fragments

To determine a percent of MHC fitting fragments within all generated fragments, the amounts of 8 to11-meric peptides were calculated. In the case of long nonglycosylated P1 and P2 peptides 8 to11-mers represented only 20% of all generated fragments, whereas their amount showed a tendency to increase up to 50% with shortening of the substrate (refer to line A, Table 6).

The design of breast cancer vaccines should involve glycosylated MHC epitopes, resembling the structure of cancer-associated MUC1 protein. Therefore, we analysed the amount of glycosylated 8 to11-mers within the pool of all generated fragments of the substrate. As nonglycosylated peptides can not be the source of peptides that resemble fragments of cancer MUC1, they should be considered inapplicable. The GP2 and GP5 glycopeptides glycosylated at threonine adjacent to the SAP motif released high amounts of 8 to11-mers, but only 20% of them were glycosylated. The glycosylation of Thr within the GST region resulted in generation of high amounts of 8 to11-mers (app 40% of all generated fragments), of which app 40% were glycosylated.

GP3 and GP6 glycopeptides, although releasing lower amounts of 8 to11-mers (app 25%) compared to other glycopeptides, generated almost exclusively glycosylated fragments. Therefore peptides with glycosylation at the central DTR/ESR motif can be considered as the best source of short, glycosylated MHC fitting fragments.

Table 6. Amount of 8 to 11-meric fragments (A) compared to the amount of glycosylated 8 to 11-meric fragments (B). Amount of MHC fitting fragments, 8 to 11 amino acids long was calculated and expressed as the percent of the total amount of all digestion products of one substrate. From the amount of 8 to 11-mers, the percent of glycosylated fragments was calculated.

A

B

peptide	P1	P2	P3	P4	GP2	GP3	GP4	GP5	GP6	GP7	GP8
8-11 mers (%)	13	26	52	50	55	15	48	79	35	29	35
% of glycopeptide	0	0	0	0	18	98	45	26	96	39	21

3.1.5.3 Region of origin of 8 to 11-meric peptides

To explain the difference in amount of glycosylated 8 to11-mers we analysed, from which region of peptide they were generated. The 8 to11-meric fragments from the digest of MUC1 VNTR glycopeptides were grouped based on their N-terminal amino acid in three groups and the relative amounts of fragments in each group were calculated (expressed as % of the total amount of 8 to11-mers).

The majority of 8 to 11-mers was produced by successive digestion at the SAP and GST motifs of analysed peptides. (Table 7, middle section). As GP3 and GP6 carry sugar at the central DTR/ESR localised between two major cleavage regions, most of the 8 to11-mers released are glycosylated.

	A H G V T	S A P D T R	P A P G S T A P P A
GP2	18%	42%	40%
GP3	20%	78%	2%
GP4	3%	60%	37%
GP5	26%	27%	47%
GP6	55%	42%	3%
GP7	8%	61%	31%

Table 7.Relative frequencies of 8 to 11-meric peptide fragments grouped according to their N-terminus. The 8 to11-meric fragments from the digest of MUC1 VNTR glycopeptides were grouped based on their N-terminal amino acid into three groups and the relative amount in each group was calculated (expressed as % of the total amount of 8 to 11mers). Bold grey squares mark the third of the peptide that carries glycan.

One more confirmation that DTR/ESR glycosylated peptides represent the best substrates for generation of potential MHC epitopes was found in the rpHPLC chromatograms of three different mono-glycosylated peptides: GP2, GP3 and GP4. DTR glycosylated GP3 releases the largest amount of short peptides (section 3 on Fig. 23), when compared to GP2 and GP4. GP4 releases large amounts of longer fractions, generated mainly by digestion within the SAP motif (section 2 Fig. 23). These peaks are present in low amount in the digest of GP2 glycopeptide glycosylated on Thr proximal to SAP, and the majority of substrate remains unprocessed (section 1, Fig. 23)

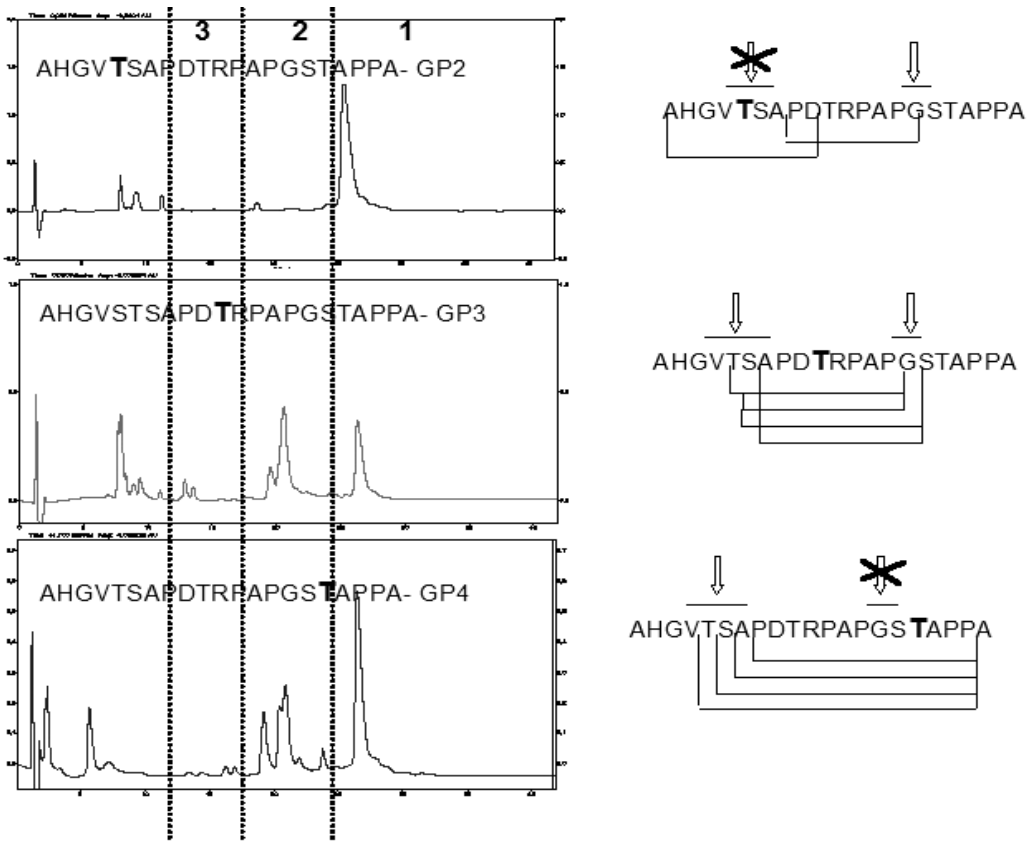

Figure 23. Influence of glycan position on the generation of 8 to11-meric peptides and processing specificity. The digests of three peptides mono-glycosylated at three different regions were analysed by reversed phase HPLC. On the chromatograms, area 1 marks time frame when non-processed peptide elutes from the column while area 2 - when peptides generated by single processing in VTSA region and with unprocessed C terminus elute. Area 3 marks the time period during which short 8 to 11-mers starting at VTSA and finishing within the GST region are eluted. The GP2 glycopeptide was weakly processed due to blocked digestion in the VTSA region. Glycosylation of the DTR motif in GP3 did not prevent the cleavage at VTSA and GST regions, but reduced processing efficiency. Glycosylation of GP4 at the GST region reduced digestion within the motif as well as overall processing efficiency.

3.1.5.4 Prediction of MHC binding activity of 8 to11-meric fragments

In the digests of all analysed 21-meric glycopeptides, 117 fragments in the size range of 8 to 11 amino acid long peptides and glycopeptides were generated (Table 8).

AHGVTSAP	GVTSAPDT	PDTRPAPGSTA	SRPAPGSTAP
AHGVTSAPDTR	GVTSAPDT	PESRPAPG	SRPAPGSTAPP
AHGVTSAPE	GVTSAPDTR	PESRPAPG	TRPAPGST
AHGVTSAPE	GVTSAPDTR	PESRPAPGS	TRPAPGSTA
AHGVTSAPES	GVTSAPDTRP	PESRPAPGST	TRPAPGSTA
AHGVTSAPESR	GVTSAPES	PESRPAPGSTA	TRPAPGSTAP
AHGVTSAPESR	GVTSAPES	PGSTAPPA	TRPAPGSTAP
APDTRPAP	GVTSAPESR	RPAPAGSTAP	TRPAPGSTAPP
APDTRPAPG	GVTSAPESRP	RPAPGSTA	TSAPDTR
APDTRPAPG	HGVTSAPD	RPAPGSTA	TSAPDTR
APDTRPAPGS	HGVTSAPD	RPAPGSTAP	TSAPDTRP
APDTRPAPGS	HGVTSAPDTR	RPAPGSTAPP	TSAPDTRPA
APDTRPAPGST	HGVTSAPES	RPAPGSTAPPA	TSAPDTRPA
APDTRPAPGST	HGVTSAPES	RPAPGSTAPPA	TSAPDTRPAPG
APESRPAP	HGVTSAPES	SAPDTRPA	TSAPDTRPAPG
APESRPAPG	HGVTSAPESR	SAPDTRPAP	TSAPDTRPAPG
APESRPAPG	HGVTSAPESR	SAPDTRPAPG	TSAPESRP
APESRPAPGS	HGVTSAPESRP	SAPDTRPAPG	TSAPESRPA
APESRPAPGS	PAPGSTAP	SAPDTRPAPGS	TSAPESRPAPG
APESRPAPGST	PAPGSTAPP	SAPESPAPG	TSAPESRPAPG
APGSTAPP	PAPGSTAPPA	SAPESRPA	VTSAPDTR
APGSTAPPA	PAPGSTAPPA	SAPESRPA	VTSAPDTR
DTRPAPGS	PDTRPAPG	SAPESRPAP	VTSAPDTRPAP
DTRPAPGST	PDTRPAPG	SAPESRPAPG	VTSAPESR
DTRPAPGSTAP	PDTRPAPGS	SAPESRPAPG	VTSAPESRPA
ESRPAPGS	PDTRPAPGS	SAPESRPAPGS	VTSAPESRPAP
ESRPAPGST	PDTRPAPGST	SAPESRPAPGS	VTSAPESRPAP
ESRPAPGST	PDTRPAPGST	SRPAPGST	
ESRPAPGSTA	PDTRPAPGST	SRPAPGSTA	
ESRPAPGSTAP	PDTRPAPGSTA	SRPAPGSTAP	

Table 8. The list of all found 8 to 11meric peptides with marked glycosylation position. Fractions found in digests of AHG series (glyco)peptides were analysed and 8 to 11 amino acids long fractions selected fro further analyses.

To identify MHC binding glycopeptides, we performed a pre-selection based on predicted ligation strength of the peptide. The ligation strength of found 8 to 11-meric fragments to MHC class I was analysed by the SYFPEITHI MHC-epitopes analysing software and experimentally tested by MHC stabilization assay (Table 9). The tested peptides and glycopeptides were different in their amino acid sequence and glycosylation position. The prediction software could not incorporate the glycosylation parameter, since its influence on MHC binding is not known. Furthermore, the software is not able to include 11 amino acids

long fragments. Therefore the information on glycosylation position was not included and just amino acid sequence of glycopeptides was analysed. Due to these restrictions, the number of software-analysed peptides was reduced to 52 8 to 10-meric peptides. Analyses concentrated on the HLA-A0201, as the most frequent MHC class I allele.

Fragment	Ligation strength	Fragment	Ligation strength	Fragment	Ligation strength
SAPESRPAP	10	GVTSAPDTRP	4	AHGVTSAPDTR	0?
SAPESRPAPG	10	GVTSAPESRP	4	AHGVTSAPESR	0?
TSAPDTRPA	10	APESRPAPG	3	APDTRPAP	0?
PGSTAPPA	9	APESRPAPGS	3	APDTRPAPGST	0?
RPAPGSTA	9	HGVTSAPES	3	APESRPAP	0?
SAPDTRPA	9	PESRPAPGST	3	APESRPAPGST	0?
SAPDTRPAP	9	SRPAPGSTAP	3	DTRPAPGS	0?
SAPDTRPAPG	9	HGVTSAPDTR	2	DTRPAPGSTAP	0?
SAPESRPA	9	HGVTSAPESR	2	ESRPAPGS	0?
SRPAPGSTA	9	PDTRPAPGST	2	ESRPAPGSTAP	0?
VTSAPESRPA	9	TRPAPGSTAP	1	HGVTSAPESRP	0?
PAPGSTAPPA	8	GVTSAPDT	0	PDTRPAPGSTA	0?
TSAPESRPA	8	GVTSAPES	0	PESRPAPGSTA	0?
APGSTAPPA	7	HGVTSAPD	0	RPAPGSTAPPA	0?
PAPGSTAPP	7	PAPGSTAP	0	SAPDTRPAPGS	0?
RPAPGSTAP	7	PDTRPAPG	0	SAPESRPAPGS	0?
TRPAPGSTA	7	PDTRPAPGS	0	SRPAPGSTAPP	0?
GVTSAPDTR	6	PESRPAPG	0	TRPAPGSTAPP	0?
GVTSAPESR	6	PESRPAPGS	0	TSAPDTRPAPG	0?
RPAPGSTAPP	6	SRPAPGST	0	TSAPESRPAPG	0?
AHGVTSAPES	5	TRPAPGST	0	VTSAPDTRPAP	0?
ESRPAPGST	5	TSAPDTRP	0	VTSAPESRPAP	0?
ESRPAPGSTA	5	TSAPESRP	0	VTSAPESR	0
AHGVTSAPE	4	VTSAPDTR	0		
APDTRPAPG	4	APGSTAPP	0		

Table 9. Binding prediction of proteasome generated fragments to the HLA-A0201 allele. 52 experimentally found 8 to 10-meric fragments were analysed by the SYFPEITHY software tool for ligation strength to the human HLA-A0201 allele. The information about glycosylation position was not included in analyses due to software restrictions.

From the list of analysed fragments, the following peptides were synthesized: GVTSAPDTR, GVTSAPESR, SAPDTRPAPG, SAPESRPAPG (Table 10). Peptides with DTR/ESR sequence in the central region were not synthesised regardless of the good binding prediction because they could not be glycosylated *in vitro* (Table 10).

The fragments of the MUC1 VNTR sequence that were not found in the digests, but predicted by SYFPEITHY to have high ligation strength were synthesized if they had possible *in vitro* glycosylation sites.

found in digests	not found in digests	found in digests, not synthesised
GVTSAPDTR GVTSAPESR SAPDTRPAPG SAPESRPAPG	STAPPAHGVT STAPPAHGV APPAHGVTSA	VTSAPESRPA SAPESRPA SAPDTRPA TSAPDTRPA TSAPESRPA SAPESRPAP APDTRPAPGS APESRPAPG APDTRPAPG DTRPAPGST

Table 10. List of peptides that bind to HLA-A0201 with strong ligation strength, as predicted by SYFPEITHI software. Peptides found in the digests of analysed MUC1 VNTR substrates and theoretical fragments of MUC1 VNTR sequence, not found in the digests, were selected for further analyses. Peptides that were found in digests of analysed substrates and predicted as good binders in non glycosylated form, but could not be glycosylated in vitro were not analysed further.

3.1.5.5 Binding efficiency of MUC1 peptides and glycopeptides to cellular HLA-A0201 molecules

The selected peptides were synthesised, glycosylated in vitro and tested for binding to empty HLA-A0201 molecules on T2 cells by MHC stabilisation assay. Tested peptides were analysed in parallel with the positive control: FLPSDCFPSV peptide. T2 cells incubated without addition of peptide were used as negative control. The binding affinities were measured by FACS and results for glycosylated peptides were compared with affinities of their non-glycosylated forms. Figure 24 presents the ligation strengths of tested peptides compared with negative control considered as 0 and positive control considered as 100%.

The strong reduction of ligation strength is noticed upon introduction of GalNAc, reducing it to the level of negative control. The exception presents the glycopeptide SAPD**T**RPAPG-GalNAc (SAP10-GalNAc) which is the only MUC1 VNTR glycopeptide binding with the same affinity in the glycosylated and the nonglycosylated form. All other peptides lost their binding abilities after glycosylation.

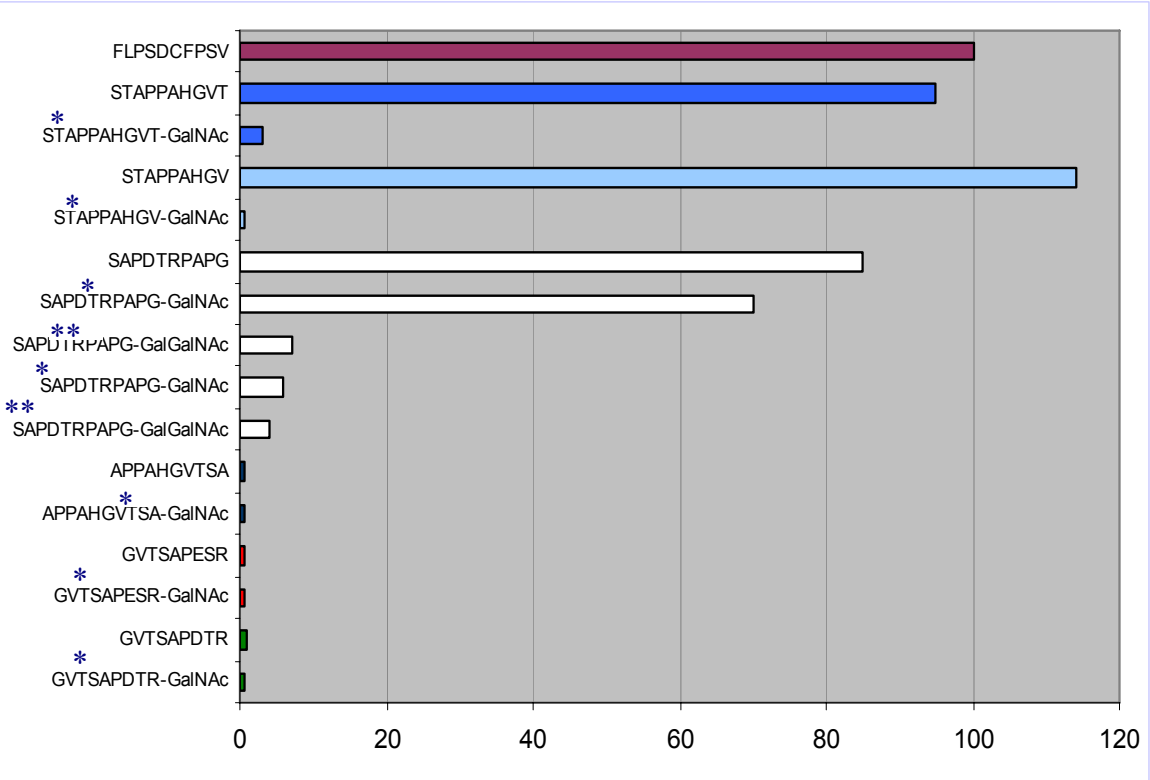

Figure 24. Ligation strength of MUC1 VNTR peptides and glycopeptides to HLA-A0201. Binding affinities of tested peptide and glycopeptides were measured by MHC stabilisation assay and expressed as % of the positive control. The negative control was considered as 0. Glycosylation position is labelled with asterisk above the amino acid (Thr or Ser); two asterisks represent disaccharide.

3.2 Cellular assays

3.2.1 Cross presentation

In another series of experiments we analysed the processing of MUC1 VNTR sequence on the cellular level. We performed cross-presentation of MUC1 peptide in the mouse dendritic cell line DC2.4 or in human monocyte derived dendritic cells. MUC1 peptides P1 or P2 were introduced into the cytosol of DCs by electroporation, loaded liposomes or poly-lactic acid vesicles. The liposomal approach as well as electroporation did not generate enough amount of MUC1 derived MHC-ligands to enable a mass spectrometric detection. The control experiment showed that electroporation results in the introduction of fluorescently labelled peptides into the cytosol of 5-10% of treated cells. The liposomal approach was confirmed successful with ELISPOT assay performed by Dr. Kondo Eisei at Molecular tumor biology and tumor immunology group, University of Cologne, but generated epitopes could not be identified by mass spectrometry.

The loading of poly-lactic vesicles with antigen resulted in the generation of high amounts of MUC1-derived MHC epitopes that could be identified for the first time on the peptidomic level by mass spectrometry. After incubation of mouse dendritic cells DC2.4 with poly-lactic acid vesicles (PLA) carrying P2 peptide, exosomes produces by the cells were collected by ultracentrifugation, acidified and fractionated by rpHPLC. Collected fractions were dried by vacuum centrifugation and analysed by MALDI/MS/MS. The molecular masses of detected ions were analysed by FindPep software application, as described in section 2.2.5.

The following MUC1-generated fragments were detected after the pulse of DC2.4 with PLA-encapsulated P2 peptide (Table 11):

(DA)	*SEQUENCE*	MW(DA)	*SEQUENCE*
823	*APESRPAP*	1264	*PESRPAPGSTAPP*
844	*TSAPESRP*	1523	*TSAPESRPAPGSTAPP*
881	*APESRPAPG*	1549	*HGVTSAPESRPAPGST*
896	*PESRPAPGS*	1629	*SAPESRPAPGSTAPPAH*
937	*STAPPAHGVT* *TAPPAHGVTS*	1692	*AHGVTSAPESRPAPGSTA* *VTSAPESRPAGSTAPPA*
968	*SAPESRPAPG* *APESRPAPGS*	1786	*SAPESRPAPGSTAPPAHGV*
999	*PESRPAPGST*	1986	*VTSAPDTRPAPGSTAPPAHGV*
1056	*SAPESRPAPGS*	2323	*TAPPAHGVTSAPESRPAPGSTAPPA*
1169	*GVTSAPESRPAP* *VTSAPESRPAPG*	2459	*TSAPESRPAPGSTAPPAHGTSAPES*

Table 11. Peptides generated during cross-presentation of P2 peptide in DC2.4 mouse dendritic cells. To masses measured on MALDI/MS the corresponding sequence of P2 was assigned.

3.2.2 Analyses of MHC bound MUC1 glycopeptides on breast cancer cell lines

Presence of MHC molecules on the breast cancer cell line T-47D was confirmed by confocal microscopy (Figure 25) and flow cytometry. We attempted to detect MUC1 derived (glyco)peptides bound to MHC class I molecules of T-47D cells. Cells were acid-eluted or MHC molecules were affinity purified by HLA-A,B,C antibody W6/32. The eluted epitopes were analysed by ESI/MS/MS analyses. To the ionisation spectra the corresponding amino acid sequences of human proteins were assigned and the reliability of the result presented as homology score. The majority of fragments corresponded to house-keeping proteins of the cell. Table 12 presents the selection of 8-11meric peptides connected to cancer phenotype and the normal house keeping cellular proteins.

We did not succeed in the identification of MHC-bound MUC1-derived peptides. The eluted mixture contained a group of 14 different peptides corresponding to the CA125 protein from mucin gene family (MUC16). In table 12 only a nonameric fragment is listed. Beside MUC16, a tumor necrosis factor derived peptide, a cervical cancer suppressor derived peptide and the macrophage migration inhibitory factor were identified.

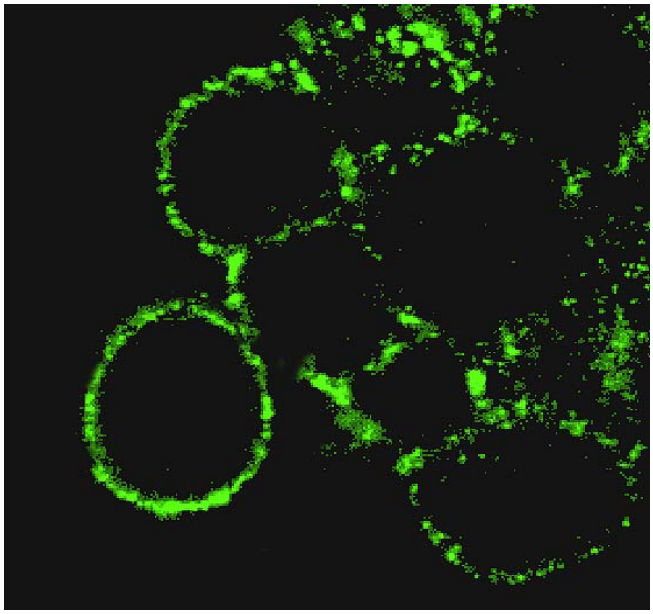

Figure 25. Fluorescent microscopy of T-47D cells stained with MHC class I specific antibody and FITC labelled secondary antibody.

Suggested sequence	Homology score	MW	Name
DQGITSLTN	9	947,5	gi|24419041 CA125 ovarian cancer related tumor marker
KPGMVVTFA *FSDYPPLGRF*	13 24	948,56 1197,61	gi|24210508 Cervical cancer suppressor 3
LFIMLENSQMR	19	1380.72	gi|339992, Tumor Necrosis Factor
LTIGSKLQDA *ALIYNEALKG* *LDRNKDQEVNF*	32 32 30	1044,58 1090,61 1376,65	gi|30582769 Calcyclin
FVQMMTAK	54	954,56	gi|16974826 Calcium calmodulin
QPGMVVTFA	13	948,56	gi|12006049 EF1a-like protein
FADKVPKTAEN *ILKHTGPGILS*	35 23	1218.65 1134.68	gi|3659980 Cyclophilin A
RISPDRVY *RLRISPDRVY*	27 23	1004,54 1273,73	gi|188556 (Macrophage) migration inhibitory factor
PVGAANFREA *FRSGKYDLD*	41 15	1030,50 1099,54	gi|2661039 Alpha enolase
QLALETIDINK	27	1256.84	gi|409219, Spliceosomal protein
EDQDLLKLVK	14	1199,54	gi|14488774, Human Mica In Complex With Natural Killer Cell Receptor Nkg2d
FYYLGPFK	12	1033.55	gi|5514625, cadherin-10
VIQKKSGK	12	1263.67	gi|2257764, polymerase
MAMDSSLQAR	19	1108.58	gi|14250610, Kinesin family member 2C
YTFEIFDGK	11	1118.67	gi|31418212, Myomesin 2
FTMTIDAHTN	24	1149.72	gi|28780, apo-B100 precursor
PPGTFGQN	19	816.46	gi|14017779, MEGF11 protein
LLKDVQCS	17	904.68	gi|15826732, Chain B, Human S-Adenosylmethionine Decarboxylase
MRYVASYLL	38	1114.66	gi|4506671, ribosomal protein P2

Table 12. The list of ESI/MS/MS sequenced peptides from acid eluent of T-47D breast cancer cell line.

4 DISCUSSION

4.1 Tumor antigen MUC1 and the immune system

MUC1 is a human tumor antigen used in tumor diagnostics serum tests (CA15-3) and involved in numerous attempts to develop tumor specific vaccines for both preventive and therapeutic purposes. Unlike the majority of mucins that are secreted from cells, MUC1 is expressed in a transmembrane and secreted form. Though encoded as a single protein, it is expressed as a type I transmembrane heterodimer. The major part of the MUC1 ectodomain is composed of tandemly repeated 20 amino acid sequences, GVTSAPDTRPAPGSTAPPAH, referred to as VNTR domain. The 20-mer can be repeated up to 120 times in a single MUC1 molecule and is densely O-glycosylated on its five serine and threonine residues. During the lactation phase in women and breast cancer development secreted and shed forms of MUC1 can be detected in serum and bod fluids. Lactation-associated protein is less densely O-glycosylated by long polylactosamine-type chains on its VNTR domain. The extensive shortening of the carbohydrates together with their increased density and a higher substitution with sialic acid are the hallmarks of cancer-associated MUC1 glycoforms. This shortening of carbohydrates exposes the peptide core residue to antibodies, protein receptors and digestive enzymes (Liu X et al., 1995; Grinstead JS et al., 2002; Vlad AM et al., 2004).

The secreted or shed MUC1 in sera of cancer patients reaches lymph nodes and other immunogenic hot-spots. Antigen-presenting cells residing there can take up the circulating MUC1 and possibly present its peptides on MHC class I and class II molecules stimulating thereby the cellular immune response.

Vlad et al (Vlad AM et al., 2002), and Apostolopoulos et al (Apostolopoulos V et al., 2003) showed that MUC1 glycoprotein can be presented to T cells without prior removal of sugar chains at MHC class II or class I complexes of antigen presenting cells. In a previous study published by Hanisch et al (Hanisch FG et al., 2003), we showed that cathepsin L, the major enzyme involved in generation of MHC class II epitopes, can process glycosylated MUC1 peptides, with restrictions introduced by glycosylation positions and type. The study identified T-S-A-P, G-ST and H-GV peptide bonds as main processing regions within the MUC1 VNTR sequence. Moreover, it showed that glycans located in the proximity of the cleavage sites caused inhibitory effects on enzyme activity, whereas their localisation on amino acids distant from the digestion site did not have effects on processing efficiency. The MHC class II bound MUC1 glycopeptides can stimulate T helper cells as shown by Vlad et al (Vlad A et al., 2002). Although the stimulation of T helper cells is of importance for long-lasting anti-cancer

immune responses and the induction of B-cell responses, the activation of cytotoxic T-cells is necessary for tumor eradication.

4.2 Processing of nonglycosylated MUC1 sequence by immunoproteasomes

The stimulation of cytotoxic T-cells requires the presentation of MUC1 derived glycopeptides on MHC class I molecules of antigen presenting cells. The generation of MHC class I bound peptides is performed by a special type of proteasomes, although the involvement of cytoplasmic endopeptidases or even cathepsins have been reported (Seifert U et al., 2003; Shen L et al., 2004).

According to results described in this work, isolated proteasomes process the MUC1 VNTR sequence most frequently within two major digestion regions of the tandem repeat peptide-S-A-P and G-S-T, in contrast to some peptide bonds that are rarely used and that release only 1% of fragments. Some processing is also noticed at RP and AP(PA) regions.

Processing and presentation of peptides in cellular assays showed a more specific cleavage pattern restricted almost exclusively to S-A-P and G-S-T motifs. No peptides that were eluted from mouse H2Kb MHC class I allele started within minor processing regions. The high similarity of results we achieved by *in vitro* proteasome digestion and in cellular assays supports the use of *in vitro* system for the analysis of antigen processing and presentation pathways.

The frequent proteasomal cleavage within S-A-P and G-S-T motifs correlates with the prediction assay at NetChop 3.0 server and with known proteasomal processing patterns (Fig. 26) (Nielsen M et al., 2005).

According to the proteasomal cleavage site prediction tool, the main digestion region within MUC1 VNTR peptides is the VTSAP motif, but processing can occur also at RP, GS and AP regions (Fig. 26). The results of *in vitro* experiments indeed revealed that proteasomes preferentially process MUC1 VNTR nonglycosylated peptides within S-A-P and G-S-T regions with some digestion of lower frequency at RP or AP motifs (Fig. 26B). Cleavage frequencies at low level could be noticed at all peptide bonds of the sequence. The variation of amino acid sequence from DTR to ESR did not influence the processing specificity of proteasomes. All non-glycosylated MUC1 VNTR peptide sequences were processed by immunoproteasome at high rate (100%) with no residual substrate detected. This shows that

the MUC1 VNTR amino acid sequence is processed in a highly specific manner and contains no peptide bonds resistant to proteasomal processing.

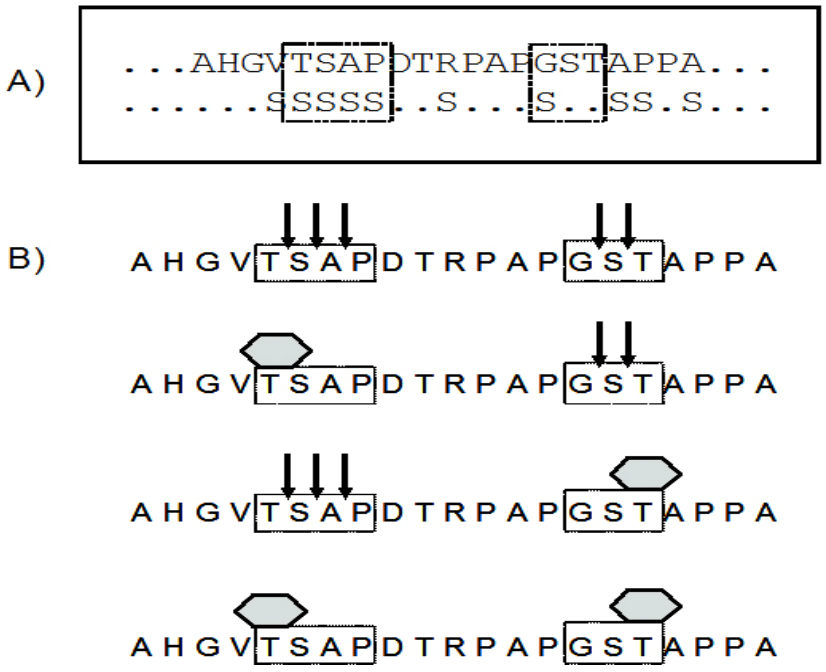

Figure 26. The processing pattern of MUC1 VNTR sequence predicted by NetChop 3.0 (A) and experimentally shown. (B). According to prediction of cleavage sites performed by NetChop 3.0, major cleavage regions of non-glycosylated MUC1 VNTR are located within S-A-P and G-S-T motif. Digestion of non-glycosylated and glycosylated MUC1 VNTR peptides by immunoproteasomes revealed S-A-P and G-S-T as the main digestion regions which are blocked by glycosylation in their close proximity

Our experimental result of processing specificities correlates with known proteasomal preferences (Fig. 27) (Emmerich NP et al., 2000; Toes RE et al., 2001). The proteasome is a multi-catalytic protease complex containing three different active sites with chymotrypsin-, trypsin and caspase like activity. But the proteasome cannot be considered a simple collection of such enzymes. In fact it is the structural architecture of the proteolytic chamber that determines specificity. The local structure around the enzyme active site structurally selects the peptide substrates fitting to the substrate binding clefts to be lysed. Due to its substrate specificity, the proteasome digests preferentially between specific amino acids (labelled P1 and P1'), but more distant acids also influence the process.

The major cleavage sites of MUC1 VNTR, the S-A-P and G-S-T regions, correspond to the frequent Ser and Gly, as well as Thr and Ala occurrence at P1 position (N terminaly of the cleaved peptide bond) (Fig. 27). Digestion at RP region corresponds to tryptic like activity of proteasome with preference for Arg at P1 position. Digestions occurring within DTR/ESR motifs can be attributed to caspase like activity with preference for Glu or Asp at P1 position.

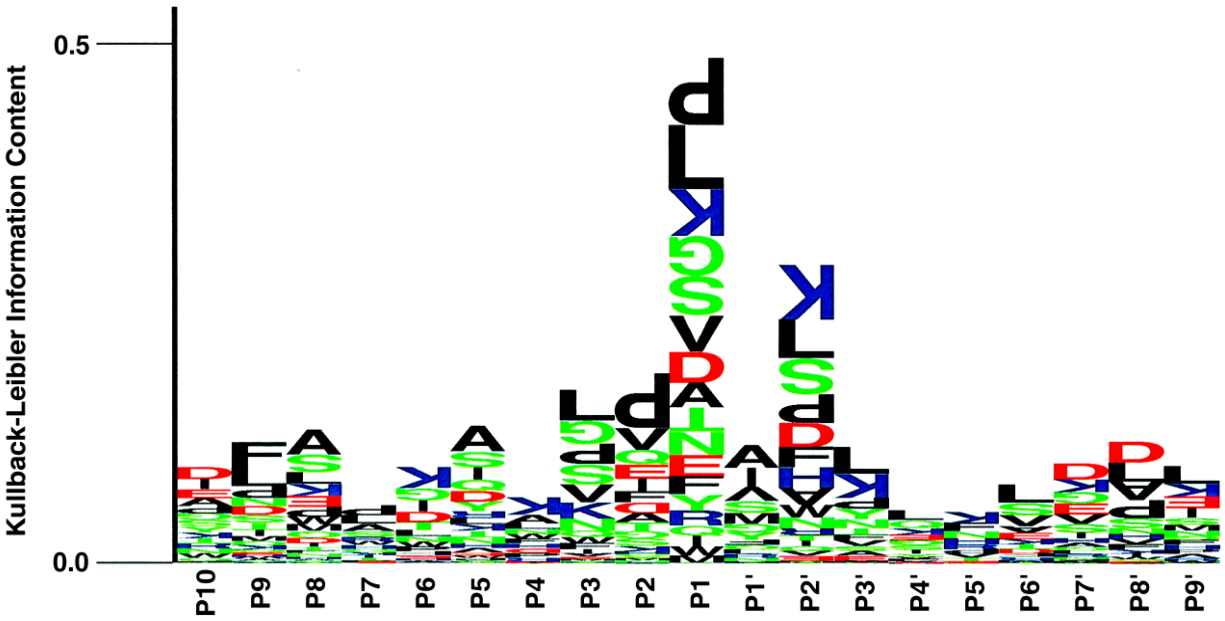

Figure 27. Sequence logo generated using *in vitro* data on digestion of enolase and ß-casein by human 20S constitutive proteasome. Prefered amino acids for proteasomal digestion on the place of processing (P1 and P1') and distant from it. Amino acids are colour coded according to their physicochemical characteristics. Neutral and polar, green; basic, blue; acidic, red; neutral and hydrophobic, black. Upside-down letters show the amino acids that are under-represented compared to the background distribution. Proteins digestion statistic generated on enolase (Toes *et al*., 2001) and ß-casein (Emmerich *et al*., 2000). Picture from Kesmir C et al.,2002.

4.3 Processing of glycosylated MUC1 by immunoproteasomes

Up to now, no proteasomal processing of glycosylated peptides was described. Although it was shown that MHC molecules can bind glycopeptides and that glycopeptide specific T-cells exist, it is not known how the glycosylation of a substrate influences the proteasome specificity. The glycosylation was reported to prevent proteasomal degradation of Sp1 transcription factor (Han I et al., 1997). The question whether glycan position and type could control processing was up to now not addressed.

We performed for the first time the *in vitro* digestion of glycosylated MUC1 VNTR peptides and analysed the influences of glycan position and type on the processing outcome. The digested MUC1 glycopeptides were fractionated by rpHPLC and analysed by MALDI-TOF/MS or ESI/MS/MS. Based on relative peak areas in the chromatograms and MALDI/MS ion intensities we quantified the relative amounts of generated fragments, the percent of degraded substrate and the frequencies of cleavages at single peptide bonds. This quantification of processing efficiencies in a relative manner (%) allowed us to compare data obtained for different glycopeptides as well as to correct for eventual variations in absolute amounts of starting substrate.

We observed a general reduction in digestion rates of glycosylated peptides compared to nonglycosylated substrates, resulting in digestion efficiencies below 100%. We determined that activity and specificity of proteasomes is influenced by glycan structure and its position within the peptide sequence.

- **Monosaccharide glycopeptides**

  Glycosylation of AHG series of peptides, even by GalNAc monosaccharide**,** modulated proteasomal efficiency. The proteasomes retain their site specificities of cleavage in case of glycopeptides, but the actual proteolysis is directed by glycan presence and thereby the digestion outcome is modulated. Glycosylation of major processing motifs significantly reduces processing within the motif. This reduction causes overall reduction of processing rate of the glycosylated substrate (Fig. 26B).

  The effect of glycosylation is stronger in the case of glycosylation of Thr adjacent to the SAP motif, than within the GST motif, agreeing with the observation that digestion within the GST motif contributes less to substrate degradation than within SAP. The glycosylation of DTR reduces the overall digestion rate in comparison to nonglycosylated peptides, but does not specifically influence the digestion outcome. Substrates glycosylated on both (T)SAP and GST positions are not processed due to the cumulative effect of glycosylation-induced blocking on both major processing motifs (Fig. 26B). In such cases poor substrate positions are used at low frequencies. As the immunoproteasome generates MHC binding peptides, the outcome of antigen processing and presentation will therefore be dependent on its glycosylation. We could not point out any universal pattern describing this low-specificity proteolysis.

- **Disaccharide-glycopeptides**

  The cancer-associated MUC1 is often glycosylated by **T antigen (Gal-GalNAc**). Surprisingly, *in vitro* processing of disaccharide-glycosylated peptides resulted in higher digestion rate than that of monosaccharide-glycosylated peptides, although a decrease should be expected due to sterical hindrance. The analyses of processing products revealed that introduction of disaccharide increased the digestion at the AP(PA) motif proximate to the carboxy terminus and thereby resulted in a considerable amount of long, almost full-length fragments. This phenomenon was not observed in monosaccharide glycosylated substrates. The disaccharide substituent possibly introduces a shift in substrate conformation and alters its interaction with the proteasomal active site.

However, the digestion in S-A-P and G-S-T was reduced by adjacent glycosylation as much as in the case of monosaccharide glycosylated peptides.

- **Sialylated-glycopeptides**

The most common form of glycan chains of tumor-associated MUC1 is sialylated T antigen (**NeuAc-Gal-GalNAc**). The secretory MUC1 found in sera of cancer patients is densely sialylated and proteasomes of antigen presenting cells in cancer patients possibly come into contact with this form of antigen. Sialyl-Gal-GalNAc modifications of peptides completely prevented substrate degradation by immunoproteasomes. Substrates with one sialylated disaccharide even outside of the main processing region were not processed.

The possible explanation for this could be found in the size of the glycan chain, preventing introduction of the glycopeptide into the 1,3 nm wide entrance of 20S proteasome. As the average size of monosaccharide unit is 0,5nm, the trisaccharide can be bigger than hollow diameter of the proteasomal entrance. The 20S proteasome used in these *in vitro* studies lack the regulatory and activator subunits. The studies have shown that hydrophilic substrates (and MUC1 VNTR sequence belongs to these) can enter the 20S proteasome without opening of the entrance by proteasomal 19S regulatory subunit or proteasomal activator PA28 (Kohler A, Finley D, Mol Cell 2001).

The other reason for unsuccessful processing of NeuAc-Gal-GalNAc glycosylated peptides can be the strong charge of sialic acids. Their negative charge could interfere with polar amino acids of the proteasomal active centre and disturb the proteolytic mechanism. The activation of proteasomal proteolyses depends on fine interaction of electro-negativities and charges of amino acids within the active site of the proteasome. The presence of a strong negative charge of sialic acid could probably interact with positively charged Lys and therefore block the stabilisation of Thr–O^- and shift of its H to NH_2, and with that the complete proteolytic process.

4.4 Influence of peptide starting motif

In contrast to these results, which are acquired by processing of peptides with the AHG starting motif, glycosylation of peptides with the SAP starting motif revealed less clear interference with proteasome specificities. In this case the main cleavage site (SAP) is localised N terminally and therefore no processing at that position takes place. Therefore, the glycosylation of Ser within SAP did not considerably change proteasomal specificity although

it reduced the processing rate. The GalNAc located on Thr of DTR and on Thr of G-S-T reduced the processing efficiency in a similar way as reported in case of AHG starting peptides.

4.5 The buffer influence

MUC1 VNTR peptides and glycopeptides were processed in three different buffer systems and optimal degradation was achieved with HEPES buffer containing DTT. The digestion of glycopeptides was remarkably increased under these reducing conditions. The effect was not visible with nonglycosylated substrate. The function of DTT could be correlated with reduction of the S-S bridges within proteasomes that can form by oxidation processes under *in vitro* conditions. The proteasomes are reported to be oxidised with aging as do most other proteins, and this oxidation is reducing their proteolytic functions (Ferrington DA et al., 2005). A similar process could take place *in vitro* and addition of DTT could explain reversion of oxidative deactivation. Why the same phenomenon is not observed in the case of non-glycosylated peptides is not clear. The processing of nonglycosylated peptides occurs at higher rates and can be completed before proteasomal deactivation. In cellular system reducing metabolites could take over the role of DTT, keeping the proteasome active for longer periods of time. Another possible role of DTT could be its interaction with OH groups of glycans, thereby preventing them to displace the Thr OH group in interaction with Lys from the active site of the proteasome.

4.6 Low immunity of cancer related MUC1 glycoforms

The VNTR domain of MUC1 is O-glycosylated at Thr and Ser and primarily at the TSAP and GST motifs. From the processing results the overlapping of major digestion and glycosylation regions: S-A-P/G-S-T is obvious. This location of glycans on major processing regions protects native, cancer related MUC1 from digestion by proteasomes and thereby prevents the formation of MHC class I ligands. This can explain the low immunogenicity of the protein. Design of a peptide vaccine with glycans located at tolerated positions could stimulate glycopeptide specific immune responses. An additional physiological role of glycosylation could be the protection of the core peptide from proteolytic attack by enzymes of the host and pathogen. Therefore, the function and immunogenicity of MUC1 could be dependent on this overlapping of regions.

Independent of the glycan influence on processing efficiency, we can conclude that the digestion of O-glycosylated MUC1 peptides can take place **without prior removal of glycans** and the glycopeptides can be suitable substrates for proteasomal digestion. Previously it was reported that glycosylated proteins of the ER are retrotranslocated into the cytoplasm in a glycosylated state (Hiller et al., 1996; Wiertz et al., 1996). The N-glycosylated proteins are than de-glycosylated prior to proteasomal degradation by cytoplasmic *N*-glycanase and the de-*N*-glycosylated intermediates were discovered in the cytosol upon inhibition of the proteasome. Recently, a highly conserved cytoplasmic N-glycanase (Png1p), which forms a complex with proteasome over RAD23 proteins, has been identified in yeast. It efficiently deglycosylates the substrate prior to degradation (Suzuki et al., 2001). Besides this, a proteasomal role in de-N-glycosylation was reported and suggested that they have a direct contact with the cytosolic side of the ER (Kraivanova RK et al., 2000)

In difference to N-glycosylated proteins, the O-glycoproteins were shown to be presented in MHC class I and II pathway without prior removal of glycan residues (Vlad AM et al., 2002).

The question whether O-glycosylated proteins are really processed by proteasomes could be raised, since other ways of generation of MHC bound peptides were reported (Jondal M et al., 1996; Rock KL et al., 1996). Moreover, it was shown that Sp1 transcription factor cannot be processed by proteasomes when glycosylated by single O-GlcNAc and this process is correlated with optimal nutritive conditions (Han I et al., 1997). Our results presented in this work showed that O-glycosylated MUC1 peptides are suitable substrates for immunoproteasomes and glycan location on major processing regions exerted specific inhibitory effects on digestion efficiency.

4.6.1 Sterical influence of glycans on processing outcome

The inhibitory influence of carbohydrates can have a sterical background. The likelihood of substrate cleavage depends on the mean residence time at the proteolysis sites, which is maximal if all binding sites are occupied (Dick TP et al., 1998; Nussbaum AK et al., 1998) The distance between proteolytic active sites is 2.8 nm, which corresponds to an extended structure of hepta or octapeptides (Baumeister W et al., 1998). The introduction of the glycan chain can introduce a sterical hindrance for interaction of peptide with proteasomal active sites and the incomplete occupancy of all binding sites can occur. This would result in reduced processing (Kinarsky L et al., 2003).

Although glycan presence *in vitro* had potential to completely prevent processing, the situation in cellular assays could have different outcome. The cellular proteasomes contain 19S regulatory and PA28 activator subunits, besides a catalytic 20S core, which was used in our study. The 19S particle has several functions. It opens gates of the 20S proteasome (13Å wide in isolated 20S proteasome, Lowe J et al., 1995), binds ubiquitinated substrates, de-ubiqutinates them and destabilizes their tertiary and/or quaternary structure and translocates the unfolded protein chain through the proteasomal open ends to the central chamber where catalytic sites for peptide-bond hydrolysis are located. The binding of the 19S subunit to the proteasome strongly activates the hydrolysis of short peptide substrates, probably by increasing the access of these substrates to the catalytic sites (Ma CP et al., 1993). In the processing of MUC1 by dendritic cells in *in vivo*, the 19S subunit could open the alpha rings of the proteasome and increase the annulus diameter, allowing peptides with longer glycan chains, such as the sialylated disaccharide, to enter the proteolytic chamber. Therefore the MUC1 glycopeptides that were not be processed under *in vitro* conditions could be the substrates of cellular proteasomes. The detailed molecular explanation for 19S induced proteasome activation is lacking due to the absence of a crystal structure of the 26S proteasome. Additionally, the presence of proteasome activator, PA28 could increase the processing of glycosylated peptides, as it is reported to activate the proteasomal hydrolysis of short peptides. The interaction between PA28 and proteasome entrance rearranges the annulus occluding amino acids thereby opening a pore through which substrate may pass (Whitby FG et al., 2000).

4.7 Identification of the best glycopeptide for MHC class I ligands generation

To identify the glycopeptide that could serve as the best proteasomal substrate and source of high-affinity MHC ligands, we analysed the diversity of derived fragments (number, length, glycosylation state).

Generally, the cleavage products of the proteasome vary in length between 3 and 25 amino acids with an average length of 7-8 amino acids. Several models have been suggested for the "molecular ruler" that determines the fragment length. Groll M and Huber R (Groll M et al., 1999) suggest that this is determined by the substrate binding clefts designed for peptides 7-9 amino acids long.

We have identified that digestion of longer substrates resulted in a higher number of diverse fragments and digestion of shorter substrates in a lower number of fragments. The reduction

in fragment diversity observed in the digest of shorter substrates could be based on missing generation of longer fragments from G-S-T or AP digestion regions which are found in digest of P1.

Although none of the glycosylated 21-meric peptides was processed completely, their digestion yielded a higher number of diverse fragments than the corresponding nonglycosylated peptides. This could be due to reduced specificity of proteasomes caused by glycosylation of the substrate, preventing specific processing and increasing cleavage rates at less favourable positions.

To study how efficient the generation of glycopeptide derived MHC class I epitopes is, we calculated the relative amount of 8- to 11-meric peptides (size fitting to MHC class I groove). The average length of generated peptides varied dependent on the substrate length and glycosylation state.

Longer substrates released in average longer fragments, whereas shorter substrates released in average shorter fragments. Therefore, 21-meric substrates generated substantially higher amounts of potential MHC class I ligands (50%) and because of that could be better applicable for the peptide vaccine design. It was reported that about 15% of the proteasomal products fall within the size range (8-9 residues) of typical MHC-I ligands (reviewed in Baumeister W et al., 1998) which correlates with our result in case of long peptide substrates, but is in discrepancy with amounts detected in digests of 20-meric peptides (50% of 8-11meric fragments). This result points out that shorter substrates, although decreasing the diversity of generated fragments actually generate higher amounts of MHC fitting peptides, what is of highest importance for vaccine design.

Additionally, we found out that glycosylation influences the efficient generation of MHC ligands. Glycosylated substrates varied in their generation of 8- to 11-meric fragments. Although in average 11 amino acids long peptides are generated by proteasomal digestion, glycosylation of the central DTR noticeably reduces the percent of generated 8- to 11-mers and slightly the average fragment length. This feature would disfavour DTR/ESR glycosylated peptides for the vaccine design, however almost all released 8- to 11-mers are glycosylated (98%) and actually derived from the central region of the substrate. In case of other glycopeptides with (**T**)S-A-P or G-S-**T** glycosylation the amount of glycosylated 8- to 11-mers is significantly lower (about 20-40%), and it is higher for G-S-T than for (T)SAP glycosylated peptides. The reason for this could be the use of main processing sites. As the main processing site is located within the S-A-P region, glycosylation of Thr in the proximity

of (T)S-A-P is near the cleavage site. At the same time processing is reduced by glycosylation and consequently low amounts of glycosylated 8- to 11-mers will be released. The G-S-T glycosylation is located downstream of the major processing sites S-A-P and the relatively often used RP motif, so the glycosylated Thr can be included in 8- to 11-meric peptides. The glycosylation of DTR does not block processing in major digestion motifs and allows the generation of DTR containing glycosylated 8- to 11-meric peptides by simultaneous digestion at these two sites.

Therefore, the glycopeptides with Thr glycosylation within DTR would represent the optimal antigen for a vaccine formulation: although they do reduce the processing efficiency, they do not disturb proteasomal specificity, they release almost only glycosylated 8- to 11-mers which is important for glycopeptide-based vaccine approaches, and the released glycopeptides are generated by simultaneous cleavages in proximity or within the S-A-P and G-S-T regions. The (T)S-A-P glycosylated substrates could be considered as the least appropriate vaccine-model, as the glycosylation of major cleavage sites dramatically reduces processing efficiency and proteasomal specificity, and the level of released glycopeptides is low. The G-S-T glycosylated peptides could represent a second, less favourable basis for glycopeptide-vaccines, as they do not exert so strong negative effects on proteasomal specificity as (T)SAP glycosylated peptides and are found to release relatively high amounts of glycosylated 8- to 11-mers.

4.7.1 Identification of MHC class I binding MUC1 glyco-epitope

The generated fragments were tested for binding to human MHC class I allele HLA-A0201. We identified the SAPD**T**RPAPG-GalNAc glycopeptide as a strong ligand for this allele. The binding of glycosylated peptide was comparable, but slightly lower than that of the nonglycosylated SAP10, indicating the influence of sugar on binding affinity of peptide. The disaccharide glycosylated SAP10 was not binding active on HLA-A0201.

The SAP10 glycopeptide can be derived from DTR glycosylated MUC1 VNTR peptides with either AHG or SAP starting motif. The mass signals corresponding to SAP10 and APD10 glycopeptides was found in the digests of AHG glycopeptide GP3 (AHGVTS-A-PD**T**RPAPG-S-TAPPA) as well as in the SAP glycopeptide GP13 (S-A-PD**T**RPAPG-S-TAPPAHGVT). The amount of fragment peptides found in AHG starting GP3 digest was lower (3 % of all substrate converted into SAP/APD10) than one found in SAP starting GP13 digest (8 % of all substrate converted into SAP/APD10). Although from this perspective SAP glycopeptides

could appear as the better choice for vaccine development, they yielded a restricted panel of fragments (20 different peptides compared to 33 from GP3) due to the absence of processing within major processing sites and long stretches of amino acids C terminally of the GST motif that are weakly processed. This could prevent the generation of other glycosylated MHC class I ligands that were not identified in our study.

The ESI/MS/MS analysis of the putative SAP10 or APD10 glycopeptides derived by proteasomal cleavage from GP3 glycopeptide, confirmed the presence of APD10 sequence.

Previously it was reported by Apostolopoulos et al (2003) that SAP8 glycopeptide (SAPD**T**RPA-GalNAc) was binding the MHC peptide groove of mouse H2Kb allele stronger than its nonglycosylated form, but was recognized by same CD8 T-cell clone. They suggested that the sugar of the glycopeptide was directed toward the MHC peptide groove (F pocket) and used as an anchor residue. The glycosylation of SAP8 by a disaccharide failed to stimulate development of T-cell clones, possibly because it prevented binding to the MHC molecule.

The MHC mediated presentation of glycopeptides was shown for diverse glycosylated molecules. O-GlcNAc glycosylated epitopes were detected after elution from MHC class I molecules (Haurum JS et al., 1994). It was described by Haurum et al. that binding of O-GalNAc glycopeptides to MHC class I is possible if the glycan is not attached to the amino acid that serves as an anchor residue for binding to the MHC groove (Haurum JS et al., 1995). When crystal structures of H2Kb in complex with non-MUC1 RGY8-6H-Gal2 glycopeptide and of H2Db in complex with two synthetic O-GlcNAc glycopeptides were analysed, it became clear that the sugar chain, usually not longer that two saccharide units, is solvent exposed and that it is protruding from the MHC groove toward outside.

4.8 Cross presentation

On analysis of potential MHC class I ligands from non-glycosylated MUC1 VNTR peptides Apostolopoulos et al (1997, Eur J Immunol) identified two groups of peptides: SAP9, APD9, TSA9 (containing the DTR motif) and APG9, PAP9 and RPA9 (containing the GST motif) with binding activity to the mouse H2Kb allele. These peptides are in agreement with processing products of nonglycosylated MUC1 VNTR peptide reported in this work and are possibly generated by digestions within the SAP, GST and RP regions. The presentation of endocytosed antigen by dendritic cells in the MHC class I pathway proceeds by cross-presentation pathways, that include the processing of extracellular antigen by proteasomes or

other peptidases and loading of epitopes onto MHC class I molecules. In the cellular cross presentation experiment we confirmed the presence of APD9 and TSA9 peptides presented by MHC molecule of a mouse dendritic cells line ($H2K^b$ positive). Due to the lower sensitivity of mass spectrometry-based experimental set-up compared to immunological assays, we cannot exclude that other nonameric peptides reported by Apostolopoulos were generated and eluted from exosomal MHC molecules (Apostolopoulos V et al., 1997). No peptides containing the GST motif (i.e. APG9, PAP9 and RPA9 reported by Apostoloulos) were found. Among the 8- to 11-meric peptides possibly derived by cross-presentation only peptides starting within the (T)SAP region were detected after elution from the mouse $H2K^b$ allele. The MHC bound material eluted from exosomes contained also a fraction of longer MUC1 VNTR fragments (15-25 amino acids) possibly possibly eluted from MHC class II molecules.

The cellular processing products were in full agreement with those generated by immunoproteasomal *in vitro* digestion. Therefore, MUC1 MHC class I binding peptides can be generated by proteasomal digestion without the help of additional proteases in the cytosol. In the past years several studies reported the involvement of non-proteasomal enzymes involved in generation of MHC class I peptides (Del Val M et al., 2002; reviewed in Rock KL et al., 2004). Various cytosolic amino- and endopeptidases have been identified with the capacity to either degrade proteasome products into shorter fragments or to exert limited trimming at the amino-terminal end of the peptide. It was reported that many proteasomal products are intermediate products with optimally fitting C terminus of the definitive MHC I ligand, but an extended N terminal part of varying length (Niedermann G et al., 1996). Other groups confirmed that proteasomes can generate subsets of peptides *in vitro* that do not need any additional processing for MHC I binding (Dick TP et al., 1998; Lucchiari-Hartz M 2000). It was noted that about 15% of proteasomal products fall within the size range (8-9 residues) of typical MHC-I ligands, another 15-20% are larger and thus require additional trimming, and the remaining two thirds are too short and present the degradation products of the substrate. Our cross-presentation assay showed that MHC class I binding peptides reported by Apostolopoulos can be generated solely by proteasomal degradation of the antigen and that the outcome of *in vitro* digestion reflects the situation *in vivo*.

The MUC1 VNTR peptide sequence does not contain any Lys residues and therefore can not be ubiquitinated and targeted to proteasomal degradation. But the processing of non-ubiquitinated substrates is a naturally existing phenomenon (Murakami Y et al.,1992; Benaroudj N et al.,

2001). Thereby, we can assume that MUC1 fragments found in the cellular processing assay could be derived by proteasomal degradation.

But the possibility that other enzymes are involved in our cross-presentation assay cannot be excluded. A striking similarity is observed between MUC1 cleavage patterns of proteasomal and cathepsin L catalyzed proteolysis (Hanisch FG et al., 2003). Cathepsin L is the main enzyme involved in generation of MHC-II binding peptides in late endosomes. The processing pattern of MUC1 VNTR peptides in the endosomal pathway overlaps largely with the proteasomal pattern described in this work. The loading of MHC molecules normally occurs via the ER/TAP/proteasome pathway, where peptides generated by proteasomes bind to empty MHC molecules. But also re-endocytosed MHC class I molecules from the plasmamembrane can loose their peptides under the low pH conditions in late endosomes and can bind new fragments generated by cathepsin L cleavage of exogenous antigens, like MUC1. The similar cleavage specificities of cathepsin L and proteasomes provide the basis for generation of similar peptide pools. The biochemical cause for this overlap could be in the caspase like activity of both cathepsin and proteasome, but this can not explain the processing pattern in complete. Also, the restrictions included by glycosylation of MUC1 sequence are comparable for proteasome and cathepsin L. The glycosylation of Thr within the TSA motif prevented the digestion by cathepsin L completely at the major cleavage site T-S, and drastically reduced it by proteasomes. The other main processing site H-GV reported for cathepsin L digestion of MUC1 VNTR was not strongly used by proteasome, but the molecular mass corresponding to GVT12 or VTS12 was detectable in the cross-presentation assay.

4.9 MHC class I bound peptides of a breast cancer cell line

The stimulation of MHC–restricted MUC1 glycopeptide-specific T-cell responses can be achieved only if both, tumor cells and antigen presenting cells, would be able to process glycosylated forms of MUC1. Our analyses of epitopes bound to breast cancer cell MHC class I did not reveal any MUC1-derived glycopeptides. This could be due to the sensitivity restrictions of the used mass spectrometry-based methods.

A number of other MHC-presented peptides involved in tumorigenesis was identified, including a fragment derived from another mucin protein (MUC16), from macrophage migration inhibitory factor, and from tumor necrosis factor. The possible therapeutic use of these epitopes has to be evaluated.

The presentation of glycosylated MUC1 by tumor cells theoretically could be performed via the endocytotic pathway similar to the cross-presentation over endosomal compartments or by retrograde transport of glycosylated MUC1 from the Golgi to the ER (reviewed in Lee MC et al., 2004). MUC1 is known to be initially expressed in a non-mature glycoform on the plasmamembrane and to recycle for further glycosylation, probably via the trans-Golgi network or the trans-Golgi, but the re-entry into the secretory pathway via other compartments, like the ER or ERGIC, is not excluded. The proteins can be translocated from the ER to the cytosol (Fiebiger E et al., 2002; Ye Y et al., 2004) and targeted for degradation by the ERAD pathway. Also, glycoprotein specific ubiquitin ligases were reported recently (Yoshida Y 2005). The transport of N-glycosylated proteins from the ER to cytosolic proteasomes involves specific proteins that interact with glycoproteins by targeting them to ERAD (Kim W et al., 2005). The presence of similar mechanisms for O-glycosylated proteins cannot be excluded.

4.10 Previous immunological studies

Studies dedicated to the development of MUC1-based cancer vaccines mainly were oriented toward non-glycosylated peptides.

Several groups have identified the STAPPAHGV peptide in cellular systems as a ligand for HLA-A11 and –A2 restricted CTLs. (Apostolopoulos V et al., 1997b; Domenech n et al., 1995). The recognition of MUC1 tandem repeat epitopes by CTLs from breast cancer patients was reported to occur in an MHC-unrestricted manner (Jerome KR et al., 1991). Clustering of T-cell receptors that bind to the repetitive peptide of MUC1 apparently triggers the activation of T-cells.

Murine transgenic models were used as an attractive model system to study immunogenicity of MUC1 and to perform tumor rejection studies. Apostolopolus V identified (1997) several MUC1 VNTR-derived peptides containing the DTR or GST motifs as recognized by mouse CTLs and confirmed the binding of SAPD**T**RPAP(GalNAc) glycopeptide and peptide to the mouse H2Kb allele. The majority of studies on MUC1 immunogenicity, performed in animals, used wild-type mice or rats as recipients of various MUC1-based vaccines. Since human MUC1 is xenogenic to these animals, results cannot be directly correlated to human models.

Human HLA-class I- as well as human MUC1-transgenic mice were developed. Several MUC1-based vaccines were evaluated in the mouse systems (Apostolopoulos V et al., 1996,

1994, Acress B et al., 2000). In a study by Soares MM (2001) tumor rejection in MUC1 Tg mice immunised with MUC1 VNTR peptide was reported. The vaccination of MUC1 transgenic mice with MUC1-bearing syngeneic tumors failed to develop effective anti-tumor response (Rowse GJ et al., 1998), possibly due to the inability to process fully glycosylated MUC1 antigen.

From the non-VNTR part of the protein several MHC class I binding peptides were identified, demonstrating that immunogenicity of MUC1 is not restricted to the tandem repeat region. In some cases stimulation of MUC1-specific CTLs from healthy donors was shown (Brossart P et al., 1999; Carmon L et al., 2000).

Several vaccination studies in humans were initiated based on the peptide approach (Apostolopoulos V et al., 2006; Karanikas V et al., 1997, 2001, Gilewski T et al., 2000, Brossart P et al., 2000). The vaccines succeeded in activation of humoral immune responses and in some cases weak cellular responses were developed. The cytotoxic T-cells were often of low frequency and low potency.

4.11 MUC1 breast cancer vaccine - outlook

The design of MUC1-based cancer vaccines is still confronted with the dilemma whether glycopeptides or peptides should be used for vaccination strategies. The former can be tested with respect to their immunogenicity in transgenic animals, but not with respect to their therapeutic effects. The reason for that is found in MUC1 glycosylation, which expectedly differs between humans and mice. Hence the immunotarget in hMUC1-tg mice is not authentic.

4.11.1 Peptide-based vaccination strategy

Available data on MUC1-based vaccination studies are extremely diverse due to variations of the experimental approach: mice, human cells, transgenic mice for human MUC1, transgenic mice for human MHC alleles, immunisations with glycopeptides, peptides, complete protein. The reason for this diversity lies in lacking information about the MUC1-MHC presentation in normal and tumor cells and the possibly existing tolerance to MUC1. Whether normal cells present MUC1 peptides on MHC molecules is not known. This problem relates to the use of nonglycosylated MUC1 in vaccination strategies and raises the following questions: (1) If normal cells are presenting MUC1 peptides in MHC complex, why no auto-immunity is noticed in vaccine trials? (2) Is the number of MHC/MUC1 peptide complexes a critical

parameter for T cell activation? If so, normal cells can be tolerated because they have a low level of MHC/MUC1 complexes and cancer cells are recognized due to higher abundance of the same complex originating from increased expression of MUC1? It was shown that MUC1 peptides from the non-VNTR region are able to induce CTL activation indicating that non-VNTR peptide specific T-cells are not depleted and are recognizing the normally expressed MUC1 (Pietersz GA et al., 2000). Although no auto-immunity is detected toward epithelial cells upon MUC1 peptide vaccination, the possibility exists that low-level immunity against self-MUC1 is given and boosting the immune response could increase it. (3) Is the presence of co-stimulatory molecules the crucial factor for T-cell activation? T-cells are reported to be activated only in the presence of co-stimulatory molecules and danger signals (Acres B et al., 2000; Matzinger P et al., 1994) that are not present on normal cells. All these problems are still not addressed and their solving will help clear up the peptide/glycopeptide dilemma.

4.11.2 Glycopeptide-based vaccination strategies

The tolerance and auto-immunity problem may be avoided by glycopeptide based vaccination and induction of glycopeptide-specific T-cells. As the glycopeptide form is the authentic target on cancer cells, and it was shown that glycopeptide specific T-cells exist, the recognition of tumor-related forms of MUC1 possibly can be achieved. During embryonic development no cancer-related glycosylation should be presented to T-cells and therefore no depletion of cancer-associated MUC1-specific T-cells should be expected. With this approach the problem of tolerance as well as the problem of auto-immunity would be solved. In this case the question if tumor cells are able to present MUC1 glycopeptide on MHC molecules can be raised and up to now is not answered. We did neither succeed in the detection of MUC1 glycopeptides on MHC class I, nor in some other GalNAc glycosylated peptides. Due to the sensitivity restrictions of our experimental set-up, we cannot exclude that such glycopeptides are actually presented. Tumor cells could process their own MUC1 over cross-presentation mechanisms or by retrograde Golgi-ER transport and delivery to proteasomes, where glycosylated MUC1 could enter the digestion machinery. The next problem to look for is if such a retrograde transport to proteasomes actually takes place and which form of MUC1 is involved.

4.11.3 Stimulation of MHC-unrestricted MUC1-specific recognition of tumor cells

Even if no MUC1 glycopeptides are presented on tumor cell MHC complexes, glycopeptide-specific T-cells (stimulated by vaccination) could recognize tumor-associated MUC1 in MHC-unrestricted manner and trigger tumor cell lysis. Such MHC-unrestricted T-cell response was reported to take place in breast cancer patients (Jerome KR et al., 1991). Their low frequency could be increased by MUC1 glycopeptide vaccination.

4.12 Concluding remarks

MUC1 tandem repeat domain-derived peptides and glycopeptides were found to be processed by immunoproteasomes in specific manner. We identified major proteasome processing sites in the MUC1 tandem repeat peptide: SAP and GST, but a lower level of processing is observed at all other peptide bonds in the sequence. Glycosylation of the substrate generally reduced the proteasome digestion rate in a higher or lower degree. Moreover, it influences the substrate processing in specific manner: Glycan localisation in close proximity or directly on the major processing sites of the sequence prevents the proteasomal cleavage within the site. The major cleavage sites overlapped with the major glycosylation sites of the sequence.

Hence, we could identify peptides glycosylated at the central DTR as the best substrate for processing and presentation. These glycopeptides yielded high amounts of glycosylated fragments with the potential to bind MHC class I. Moreover, these glycopeptides have a potential to release the SAPD**T**RPAPG-GalNAc glycopeptide which we identified as a good ligand for the HLA-A0201 allele of MHC class I.

The low immunity of native endogenous MUC1 originates from the overlapping of glycosylation and digestion sites. The naturally occurring MUC1 is glycosylated at major processing regions and therefore is not a suitable antigen for presentation by dendritic cells and the stimulation of glycopeptide-specific cytotoxic T-cells.

Summary

MUC1 is a human transmembrane and secreted glycoprotein expressed in numerous epithelial tissues; in breast cancer tissues MUC1 is aberrantly glycosylated and a soluble form in the sera of cancer patients' triggers weak immune responses. The stimulation of cellular immune responses toward tumor-associated MUC1 glycoprotein is a primary goal in immunotherapeutic anti-cancer strategies.

For an effective activation of tumor-MUC1 specific T-cell repertoires the extracellular MUC1 antigen should be presented by host dendritic cells in a MHC-dependent way. The generation of MHC class I epitopes is mainly performed by immunoproteasomes of dendritic cells. We performed *in vitro* studies to address the questions whether glycosylated MUC1 peptides are cleaved by IPs, and in which way the position and structure of *O*-linked glycans control the site specificity of peptide cleavage.

In our work $S^{\downarrow}A^{\downarrow}P^{\downarrow}$ (N-terminal) and G↑S↑T↑ (C-terminal) were identified as the major processing sites within the non-glycosylated MUC1. These sequences are parts of the MUC1 tandem repeat sequence (HGV**TS**APD**T**RPAPG**ST**APPA) which can be *O*-glycosylated on Thr and Ser residues. Glycosylation of the substrate in the close proximity of processing sites significantly reduced processing efficiency. Proteasomes were not able to digest glycopeptides carrying the Sialyl-Gal-GalNAc glycan, often found in the tumor associated form on MUC1 glycoprotein. The best processed glycopeptide carried glycosylation on Thr of the central DTR motif, distant from major processing regions. The highest amount of glycosylated fragments with a potential to bind to the MHC class I groove was found in the digest of glycopeptides with glycosylation of Thr within the central DTR motif.

Furthermore, the peptide SAPD**T**RPAPG glycosylated with GalNAc on Thr was identified in a cellular assay as an epitope with high affinity for the human HLA-A0201 allele. To confirm the processing pattern of MUC1 tandem repeat sequence obtained *in vitro*, cross-presentation of the MUC1 peptide on dendritic cells was performed, and eluted peptides from cell-derived exosomes corresponded to sequences found *in vitro*.

These results indicate that design of a MUC1 glycopeptide-based cancer vaccine should include glycopeptides with GalNAc on Thr of DTR motif, because of their efficient degradation by immunoproteasome, production of high amount of glycosylated 8- to 11-mers and potential generation of MHC class I binding SAPD**T**RPAPG-GalNAc epitope.

For future work, immunisation of the MUC1 transgenic mice with DTR glycosylated peptides is in progress.

Zusammenfassung

MUC1 ist ein humanes Transmembranprotein, das auf vielen epithelialen Geweben exprimiert wird. Die Glykosylierung von MUC1 auf Mammakarzinomzellen weicht von derjenigen auf Epithelzellen ab, und eine lösliche Form des Proteins im Serum der Patientin verursacht eine schwache Immunantwort des Körpers. Die Stimulierung einer zellulären Immunantwort gegen tumorassoziiertes MUC1 ist ein primäres Ziel immunotherapeutischer Strategien gegen Krebs.
Zur effektiven Aktivierung eines für das tumorassoziierte MUC1 spezifischen T-Zell-Repertoires muß das Antigen von dendritischen Zellen in einer MHC-abhängigen Weise präsentiert werden; MHC-Klasse I-Epitope werden hierbei durch Immunoproteasome (IP) erzeugt. In der vorliegenden Arbeit wurde *in vitro* untersucht, ob MUC1-abgeleitete Peptide durch IPs abgebaut werden, und inwieweit dies von Position und Struktur von *O*-Glykanen beeinflusst wird.
Die Sequenzen $S^{\downarrow}A^{\downarrow}P^{\downarrow}$ (N-terminal) und G↑S↑T↑ (C-terminal) wurden als die Hauptprozessierungsregionen innerhalb der MUC1-Aminosäuresequenz identifiziert. Diese Sequenzen sind Bestandteile der *tandem repeat*-Sequenz (HGV**TS**APD**T**RPAPG**ST**APPA) innerhalb von MUC1, deren Serin- und Threoninreste potentielle Glykosylierungsziele sein können. Glykosylierung des Substrates in Nähe der Hauptprozessierungsregionen verminderte die Effizienz des Verdaus signifikant. Die Proteasomen konnten keine Glykopeptide verdauen, die das Glykan Sia-Gal-GalNAc enthielten; solche Glykopeptide werden häufig in der tumorassoziierten Form von MUC1 gefunden. Glykopeptide, die am Threonin des zentralen DTR-Motivs – weit entfernt von den Hauptprozessierungsregionen – glykosyliert waren, wurden am besten verdaut. Zugleich resultierte aus solchen Glykopeptiden die höchste Menge an glykosylierten Prozessierungsprodukten, die aufgrund einer Länge von 8-11 Aminosäureresten am besten an die *groove* des MHC-Klasse I-Moleküls banden. Darüber hinaus ergab sich in einem zellulären Ansatz, daß das am Threonin mit GalNAc glykosylierte SAPD**T**RPAPG-Peptid eine hohe Affinität zum humanen HLA-A0201-Allel hatte. Um die *in vitro* gewonnenen Daten zur MUC1-Prozessierung zu bestätigen, wurde eine Kreuzpräsentation des MUC1-Peptides auf dendritischen Zellen durchgeführt. Die anschließend aus zellulären Exosomen eluierten Peptide entsprachen den Sequenzen, die *in vitro* gefunden worden waren.
Aufgrund dieser Resultate sollten beim Design einer auf MUC1 basierten Krebsvakzine Glykopeptide verwendet werden, die am Threonin des DTR-Motivs mit einem GalNAc-Molekül glykosyliert wurden. Gründe hierfür sind der effiziente Verdau dieser Peptide durch Immunproteasome, die Produktion einer großen Menge glykosylierter Peptide mit einer Länge von 8 - 11 Aminosäureresten und die mögliche Generierung von SAPD**T**RPAPG-GalNAc-Epitopen, die an MHC-Klasse I-Moleküle binden.
Im Hinblick auf zukünftige Arbeit erfolgt zur Zeit die Immunisierung transgener Mäuse mit Peptiden, die an der DTR-Sequenz glykosyliert sind.

Bibliography

Abe M., Kufe D. Characterization of *cis*-acting elements regulating transcription of the human DF3 breast carcinoma-associated antigen (MUC1) gene. Proc. Natl. Acad. Sci. USA, *90:* 282-286, 1993.

Acres B, Apostolopoulos V, Balloul JM, Wreschner D, Xing PX, Ali-Hadji D, Bizouarne N, Kieny MP, McKenzie IF. MUC1-specific immune responses in human MUC1 transgenic mice immunized with various human MUC1 vaccines. Cancer Immunol Immunother. 2000 Jan;48(10):588-94.

Agrawal B, Krantz MJ, Parker J, Longenecker BM. Expression of MUC1 mucin on activated human T cells: implications for a role of MUC1 in normal immune regulation. Cancer Res. 1998 Sep 15;58(18):4079-81.

Apostolopoulos V, Xing PX, McKenzie IF. Murine immune response to cells transfected with human MUC1: immunization with cellular and synthetic antigens. Cancer Res. 1994 Oct 1;54(19):5186-93.

Apostolopoulos V, Pietersz GA, McKenzie IF. Cell-mediated immune responses to MUC1 fusion protein coupled to mannan. Vaccine. 1996 Jun;14(9):930-8.

Apostolopoulos V, Haurum JS, McKenzie IF. MUC1 peptide epitopes associated with five different H-2 class I molecules. Eur J Immunol. 1997 Oct;27(10):2579-87.

Apostolopoulos V, Karanikas V, Haurum JS, McKenzie IF. Induction of HLA-A2-restricted CTLs to the mucin 1 human breast cancer antigen. J Immunol. 1997 Dec 1;159(11):5211-8. (b)

Apostolopoulos V, Yuriev E, Ramsland PA, Halton J, Osinski C, Li W, Plebanski M, Paulsen H, McKenzie IF. A glycopeptide in complex with MHC class I uses the GalNAc residue as an anchor. Proc Natl Acad Sci U S A. 2003 Dec 9;100(25):15029-34.

Apostolopoulos V, Pietersz GA, Tsibanis A, Tsikkinis A, Drakaki H, Loveland BE, Piddlesden SJ, Plebanski M, Pouniotis DS, Alexis MN, McKenzie IF, Vassilaros S. Pilot phase III immunotherapy study in early-stage breast cancer patients using oxidized mannan-MUC1 [ISRCTN71711835]. Breast Cancer Res. 2006;8(3):R27.

Baruch A, Hartmann M, Zrihan-Licht S, Greenstein S, Burstein M, Keydar I, Weiss M, Smorodinsky N, Wreschner DH. Preferential expression of novel MUC1 tumor antigen isoforms in human epithelial tumors and their tumor-potentiating function. Int J Cancer. 1997 May 29;71(5):741-9.

Baruch A, Hartmann M, Yoeli M, Adereth Y, Greenstein S, Stadler Y, Skornik Y, Zaretsky J, Smorodinsky NI, Keydar I, Wreschner DH. The breast cancer-associated MUC1 gene

generates both a receptor and its cognate binding protein. Cancer Res. 1999 Apr 1;59(7):1552-61.

Baumeister W, Walz J, Zuhl F, Seemuller E. The proteasome: paradigm of a self-compartmentalizing protease. Cell. 1998 Feb 6;92(3):367-80. Review.

Benaroudj N, Tarcsa E, Cascio P, Goldberg AL. The unfolding of substrates and ubiquitin-independent protein degradation by proteasomes. Biochimie. 2001 Mar-Apr;83(3-4):311-8. Review.

Bennett EP, Hassan H, Mandel U, Mirgorodskaya E, Roepstorff P, Burchell J, Taylor-Papadimitriou J, Hollingsworth MA, Merkx G, van Kessel AG, Eiberg H, Steffensen R, Clausen H. Cloning of a human UDP-N-acetyl-alpha-D-Galactosamine:polypeptide N-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. J Biol Chem. 1998 Nov 13;273(46):30472-81.

Berg JW, Hutter RVP. Breast Cancer. Cancer. 1995;75:257-269.

Boes M, Ploegh HL. Translating cell biology in vitro to immunity in vivo. Nature. 2004 Jul 8;430(6996):264-71. Review.

Bramwell ME, Wiseman G, Shotton DM. Electron-microscopic studies of the CA antigen, epitectin. J Cell Sci. 1986 Dec;86:249-61.

Brayman M, Thathiah A, Carson DD. MUC1: a multifunctional cell surface component of reproductive tissue epithelia.Reprod Biol Endocrinol. 2004 Jan 7;2:4. Review.

Brockhausen I, Yang JM, Burchell J, Whitehouse C, Taylor-Papadimitriou J. Mechanisms underlying aberrant glycosylation of MUC1 mucin in breast cancer cells. Eur J Biochem. 1995 Oct 15;233(2):607-17.

Brossart P, Heinrich KS, Stuhler G, Behnke L, Reichardt VL, Stevanovic S, Muhm A, Rammensee HG, Kanz L, Brugger W. Identification of HLA-A2-restricted T-cell epitopes derived from the MUC1 tumor antigen for broadly applicable vaccine therapies. Blood. 1999 Jun 15;93(12):4309-17.

Brossart P, Wirths S, Stuhler G, Reichardt VL, Kanz L, Brugger W. Induction of cytotoxic T-lymphocyte responses in vivo after vaccinations with peptide-pulsed dendritic cells. Blood. 2000 Nov 1;96(9):3102-8.

Carmon L, El-Shami KM, Paz A, Pascolo S, Tzehoval E, Tirosh B, Koren R, Feldman M, Fridkin M, Lemonnier FA, Eisenbach L. Novel breast-tumor-associated MUC1-derived peptides: characterization in Db-/-x beta2 microglobulin (beta2m) null mice transgenic for a chimeric HLA-A2.1/Db-beta2 microglobulin single chain. Int J Cancer. 2000 Feb 1;85(3):391-7.

Chen L, Jondal M. Endolysosomal processing of exogenous antigen into major histocompatibility complex class I-binding peptides. Scand J Immunol. 2004 Jun;59(6):545-52.

Chen P and Hochstrasser M. Biogenesis, structure and function of the yeast 20S proteasome. *EMBO J*, 1995, **14**, 2620–2630.

Clausen H, Bennett EP. A family of UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferases control the initiation of mucin-type O-linked glycosylation. Glycobiology. 1996 Sep;6(6):635-46

Cloosen S, Thio M, Vanclee A, van Leeuwen EB, Senden-Gijsbers BL, Oving EB, Germeraad WT, Bos GM. Mucin-1 is expressed on dendritic cells, both in vitro and in vivo. Int Immunol. 2004 Nov;16(11):1561-71. Epub 2004 Sep 20.

Dalziel M, Whitehouse C, McFarlane I, Brockhausen I, Gschmeissner S, Schwientek T, Clausen H, Burchell JM, Taylor-Papadimitriou J. The relative activities of the C2GnT1 and ST3Gal-I glycosyltransferases determine O-glycan structure and expression of a tumor-associated epitope on MUC1. J Biol Chem. 2001 Apr 6;276(14):11007-15. Epub 2000 Dec 15.

Del-Val M, Lopez D. Multiple proteases process viral antigens for presentation by MHC class I molecules to CD8(+) T lymphocytes. Mol Immunol. 2002 Oct;39(3-4):235-47. Review.

Dick TP, Nussbaum AK, Deeg M, Heinemeyer W, Groll M, Schirle M, Keilholz W, Stevanovic S, Wolf DH, Huber R, Rammensee HG, Schild H. Contribution of proteasomal beta-subunits to the cleavage of peptide substrates analyzed with yeast mutants. J Biol Chem. 1998 Oct 2;273(40):25637-46.

Domenech N, Henderson RA, Finn OJ. Identification of an HLA-A11-restricted epitope from the tandem repeat domain of the epithelial tumor antigen mucin. J Immunol. 1995 Nov 15;155(10):4766-74.

Elhammer AP, Kezdy FJ, Kurosaka A. The acceptor specificity of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. Glycoconj J. 1999 Feb;16(2):171-80. Review.

Emmerich NP, Nussbaum AK, Stevanovic S, Priemer M, Toes RE, Rammensee HG, Schild H. The human 26 S and 20 S proteasomes generate overlapping but different sets of peptide fragments from a model protein substrate. J Biol Chem. 2000 Jul 14;275(28):21140-8.

Ferrington DA, Husom AD, Thompson LV. Altered proteasome structure, function, and oxidation in aged muscle. FASEB J. 2005 Apr;19(6):644-6.

Fiebiger E, Story C, Ploegh HL, Tortorella D. Visualization of the ER-to-cytosol dislocation reaction of a type I membrane protein. EMBO J. 2002 Mar 1;21(5):1041-53.

Fowler J, Vinall L, Swallow D. Polymorphism of the human muc genes. Front Biosci. 2001 Oct 1;6:D1207-15. Review.

Gendler S J, Lancaster C A, Taylor-Papadimitriou J, Duhig T, Peat N, Burcheli J, Pemberton L, Lalani EN, Wilson D. Molecular cloning and expression of the human tumor associated polymorphic epithelial mucin PEM. J. Biol. Chem., 1990, *265:* 15286-15293,.

Gendler SJ, Spicer AP. Epithelial mucin genes. Annu Rev Physiol. 1995;**57**:607–634.

Gilewski T, Adluri S, Ragupathi G, Zhang S, Yao TJ, Panageas K, Moynahan M, Houghton A, Norton L, Livingston PO. Vaccination of high-risk breast cancer patients with mucin-1 (MUC1) keyhole limpet hemocyanin conjugate plus QS-21.Clin Cancer Res. 2000 May;6(5):1693-701.

Grinstead JS, Koganty RR, Krantz MJ, Longenecker BM, Campbell AP. Effect of glycosylation on MUC1 humoral immune recognition: NMR studies of MUC1 glycopeptide-antibody interactions. Biochemistry. 2002 Aug 6;41(31):9946-61.

Groettrup M, Ruppert T, Kuehn L, Seeger M, Standera S, Koszinowski U, Kloetzel PM. The interferon-gamma-inducible 11 S regulator (PA28) and the LMP2/LMP7 subunits govern the peptide production by the 20 S proteasome in vitro. J Biol Chem. 1995 Oct 6;270(40):23808-15.

Groll M, Heinemeyer W, Jager S, Ullrich T, Bochtler M, Wolf DH, Huber R. The catalytic sites of 20S proteasomes and their role in subunit maturation: a mutational and crystallographic study. Proc Natl Acad Sci U S A. 1999 Sep 28;96(20):10976-83

Han I, Kudlow JE. Reduced O glycosylation of Sp1 is associated with increased proteasome susceptibility. Mol Cell Biol. 1997 May;17(5):2550-8.

Hanisch FG, Muller S, Hassan H, Clausen H, Zachara N, Gooley AA, Paulsen H, Alving K, Peter-Katalinic J. Dynamic epigenetic regulation of initial O-glycosylation by UDP-N-Acetylgalactosamine:Peptide N-acetylgalactosaminyltransferases. site-specific glycosylation of MUC1 repeat peptide influences the substrate qualities at adjacent or distant Ser/Thr positions. J Biol Chem. 1999 Apr 9;274(15):9946-54.

Hanisch FG, Muller S. MUC1: the polymorphic appearance of a human mucin. Glycobiology. 2000 May;10(5):439-49. Review.´

Hanisch FG, Schwientek T, Von Bergwelt-Baildon MS, Schultze JL, Finn O. O-Linked glycans control glycoprotein processing by antigen-presenting cells: a biochemical approach to the molecular aspects of MUC1 processing by dendritic cells. Eur J Immunol. 2003 Dec;33(12):3242-54.

Hartman M., Tsarfaty I., Hareuveni M., Keydar I., Wreschner D. H. Expression of breast cancer mucin, MUC1, alters cell adhesion and a MUC1 variant may act as a receptor

molecule on breast cancer cells Ioanndou-Mouzaka L. Agnantis N. J. Karydas I. eds. . Senology: Proceedings of the 7th International Congress on Senology, *:* 379-383, Excerpta Medica Publishers Amsterdam 1992.

Hartman, M., Baruch, A., Ron, I., Adereth, Y., Yoeli, M., Sagi-Assif, O., Greenstein, S., Stadler, Y., Weiss, M., Harness, E., Yaakubovits, M., Keydar, I., Smorodinsky, N., and Wreschner D. H. The MUC1 isoform specific monoclonal antibody 6E6/2 detects preferential expression of the novel MUC1/Y protein in breast cancer and other epithelial malignancies. Int. J. Cancer, 1999.

Haurum JS, Arsequell G, Lellouch AC, Wong SY, Dwek RA, McMichael AJ, Elliott T. Recognition of carbohydrate by major histocompatibility complex class I-restricted, glycopeptide-specific cytotoxic T lymphocytes. J Exp Med. 1994 Aug 1;180(2):739-44.

Haurum JS, Tan L, Arsequell G, Frodsham P, Lellouch AC, Moss PA, Dwek RA, McMichael AJ, Elliott T. Peptide anchor residue glycosylation: effect on class I major histocompatibility complex binding and cytotoxic T lymphocyte recognition. Eur J Immunol. 1995 Dec;25(12):3270-6.

Hayashi T, Takahashi T, Motoya S, Ishida T, Itoh F, Adachi M, Hinoda Y, Imai K. MUC1 mucin core protein binds to the domain 1 of ICAM-1. Digestion. 2001;63 Suppl 1:87-92.

Hearn AR, de Haan L, Pemberton AJ, Hirst TR, Rivett AJ. Trafficking of exogenous peptides into proteasome-dependent major histocompatibility complex class I pathway following enterotoxin B subunit-mediated delivery. J Biol Chem. 2004 Dec 3;279(49):51315-22.

Heuser C, Ganser M, Hombach A, Brand H, Denton G, Hanisch FG, Abken H. An anti-MUC1-antibody-interleukin-2 fusion protein that activates resting NK cells to lysis of MUC1-positive tumour cells. Br J Cancer. 2003 Sep 15;89(6):1130-9.

Hilkens J, Kroezen V, Bonfrer JM, De Jong-Bakker M, Bruning PF. MAM-6 antigen, a new serum marker for breast cancer monitoring. Cancer Res. 1986 May;46(5):2582-7.

Hilkens J, Ligtenberg MJ, Vos HL, Litvinov SV. Cell membrane-associated mucins and their adhesion-modulating property. Trends Biochem Sci. 1992 Sep;17(9):359-63. Review

Hiller MM, Finger A, Schweiger M, Wolf DH. ER degradation of a misfolded luminal protein by the cytosolic ubiquitin-proteasome pathway. Science. 1996 Sep 20;273(5282):1725-8.

Hiltbold EM, Ciborowski P, Finn OJ. Naturally processed class II epitope from the tumor antigen MUC1 primes human CD4+ T cells. Cancer Res. 1998 Nov 15;58(22):5066-70.

Hiltbold EM, Alter MD, Ciborowski P, Finn OJ. Presentation of MUC1 tumor antigen by class I MHC and CTL function correlate with the glycosylation state of the protein taken Up by dendritic cells. Cell Immunol. 1999 Jun 15;194(2):143-9.

Hiltbold EM, Vlad AM, Ciborowski P, Watkins SC, Finn OJ. The mechanism of unresponsiveness to circulating tumor antigen MUC1 is a block in intracellular sorting and processing by dendritic cells.J Immunol. 2000 Oct 1;165(7):3730-41.

Ignatius R, Mahnke K, Rivera M, Hong K, Isdell F, Steinman RM, Pope M, Stamatatos L. Presentation of proteins encapsulated in sterically stabilized liposomes by dendritic cells initiates CD8(+) T-cell responses in vivo. Blood. 2000 Nov 15;96(10):3505-13.

Jerome KR, Barnd DL, Bendt KM, Boyer CM, Taylor-Papadimitriou J, McKenzie IF, Bast RC Jr, Finn OJ. Cytotoxic T-lymphocytes derived from patients with breast adenocarcinoma recognize an epitope present on the protein core of a mucin molecule preferentially expressed by malignant cells. Cancer Res. 1991 Jun 1;51(11):2908-16.

Jondal M, Schirmbeck R, Reimann J. MHC class I-restricted CTL responses to exogenous antigens.Immunity. 1996 Oct;5(4):295-302. Review.

Karaivanova VK, Spiro RG. Effect of proteasome inhibitors on the release into the cytosol of free polymannose oligosaccharides from glycoproteins. Glycobiology. 2000 Jul;10(7):727-35.

Karanikas V, Hwang LA, Pearson J, Ong CS, Apostolopoulos V, Vaughan H, Xing PX, Jamieson G, Pietersz G, Tait B, Broadbent R, Thynne G, McKenzie IF. Antibody and T cell responses of patients with adenocarcinoma immunized with mannan-MUC1 fusion protein. J Clin Invest. 1997 Dec 1;100(11):2783-92.

Karanikas V, Colau D, Baurain JF, Chiari R, Thonnard J, Gutierrez-Roelens I, Goffinet C, Van Schaftingen EV, Weynants P, Boon T, Coulie PG. High frequency of cytolytic T lymphocytes directed against a tumor-specific mutated antigen detectable with HLA tetramers in the blood of a lung carcinoma patient with long survival. Cancer Res. 2001 May 1;61(9):3718-24.

Karsten U, Diotel C, Klich G, Paulsen H, Goletz S, Muller S, Hanisch FG. Enhanced binding of antibodies to the DTR motif of MUC1 tandem repeat peptide is mediated by site-specific glycosylation. Cancer Res. 1998 Jun 15;58(12):2541-9.

Karsten U, Serttas N, Paulsen H, Danielczyk A, Goletz S. Binding patterns of DTR-specific antibodies reveal a glycosylation-conditioned tumor-specific epitope of the epithelial mucin (MUC1). Glycobiology. 2004 Aug;14(8):681-92.

Karsten U, von Mensdorff-Pouilly S, Goletz S. What makes MUC1 a tumor antigen? Tumour Biol. 2005 Jul-Aug;26(4):217-20. Review.

Kesmir C, Nussbaum AK, Schild H, Detours V, Brunak S. Prediction of proteasome cleavage motifs by neural networks. Protein Eng. 2002 Apr;15(4):287-96.

Kim W, Spear ED, Ng DT. Yos9p detects and targets misfolded glycoproteins for ER-associated degradation. Mol Cell. 2005 Sep 16;19(6):753-64.

Kinarsky L, Suryanarayanan G, Prakash O, Paulsen H, Clausen H, Hanisch FG, Hollingsworth MA, Sherman S. Conformational studies on the MUC1 tandem repeat glycopeptides: implication for the enzymatic O-glycosylation of the mucin protein core. Glycobiology. 2003 Dec;13(12):929-39.

Kohlgraf KG, Gawron AJ, Higashi M, VanLith ML, Shen X, Caffrey TC, Anderson JM, Hollingsworth MA. Tumor-specific immunity in MUC1.Tg mice induced by immunization with peptide vaccines from the cytoplasmic tail of CD227 (MUC1). Cancer Immunol Immunother. 2004 Dec;53(12):1068-84.

Kotera Y, Fontenot JD, Pecher G, Metzgar RS, Finn OJ. Humoral immunity against a tandem repeat epitope of human mucin MUC-1 in sera from breast, pancreatic, and colon cancer patients. Cancer Res. 1994 Jun 1;54(11):2856-60.

Lagow E, DeSouza MM, Carson DD. Mammalian reproductive tract mucins. Hum Reprod Update. 1999;**5**:280–292.

Lee MC, Miller EA, Goldberg J, Orci L, Schekman R. Bi-directional protein transport between the ER and Golgi. Annu Rev Cell Dev Biol. 2004;20:87-123. Review.

Ligtenberg M. J. L., Vos H. L., Genissen A. M. C., Hilkens J. Episialin, a carcinoma associated mucin, is generated by a polymorphic gene encoding splice variants with alternative amino termini. J. Biol. Chem., *265:* 15573-15578, 1990.

Ligtenberg MJ, Buijs F, Vos HL, Hilkens J. Suppression of cellular aggregation by high levels of episialin. Cancer Res. 1992 Apr 15;52(8):2318-24.

Ligtenberg M. J. L., Kruijshaar L., Buijs F., Van Meijer M., Litvinov S. V., Hilkens J. Cell associated episialin is a complex containing two proteins derived from a common precursor. J. Biol. Chem., 1992, *267:* 6171-6177.

Ligtenberg,M.J., Kruijshaar,L., Buijs,F., van Meijer,M., Litvinov,S.V. and Hilkens,J. Cell-associated episialin is a complex containing two proteins derived from a common precursor., J. Biol. Chem. 1992, 267 (9), 6171-6177.

Liu X, Sejbal J, Kotovych G, Koganty RR, Reddish MA, Jackson L, Gandhi SS, Mendonca AJ, Longenecker BM. Structurally defined synthetic cancer vaccines: analysis of structure, glycosylation and recognition of cancer associated mucin, MUC-1 derived peptides. Glycoconj J. 1995 Oct;12(5):607-17.

Loveland BE, Zhao A, White S, Gan H, Hamilton K, Xing PX, Pietersz GA, Apostolopoulos V, Vaughan H, Karanikas V, Kyriakou P, McKenzie IF, Mitchell PL. Mannan-MUC1-pulsed dendritic cell immunotherapy: a phase I trial in patients with adenocarcinoma. Clin Cancer Res. 2006 Feb 1;12(3 Pt 1):869-77.

Lowe J, Stock D, Jap B, Zwickl P, Baumeister W, Huber R. Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. Science. 1995 Apr 28;268(5210):533-9.

Lucchiari-Hartz M, van Endert PM, Lauvau G, Maier R, Meyerhans A, Mann D, Eichmann K, Niedermann G. Cytotoxic T lymphocyte epitopes of HIV-1 Nef: Generation of multiple definitive major histocompatibility complex class I ligands by proteasomes. J Exp Med. 2000 Jan 17;191(2):239-52.

Ma CP, Willy PJ, Slaughter CA, DeMartino GN. PA28, an activator of the 20 S proteasome, is inactivated by proteolytic modification at its carboxyl terminus. J Biol Chem. 1993 Oct 25;268(30):22514-9.

Matzinger P. Tolerance, danger, and the extended family.Annu Rev Immunol. 1994;12:991-1045. Review.

Muller S, Goletz S, Packer N, Gooley A, Lawson AM, Hanisch FG. Localization of O-glycosylation sites on glycopeptide fragments from lactation-associated MUC1. All putative sites within the tandem repeat are glycosylation targets in vivo. J Biol Chem. 1997 Oct 3;272(40):24780-93.

Muller S, Alving K, Peter-Katalinic J, Zachara N, Gooley AA, Hanisch FG. High density O-glycosylation on tandem repeat peptide from secretory MUC1 of T47D breast cancer cells. J Biol Chem. 1999 Jun 25;274(26):18165-72.

Murakami Y, Matsufuji S, Kameji T, Hayashi S, Igarashi K, Tamura T, Tanaka K, Ichihara A. Ornithine decarboxylase is degraded by the 26S proteasome without ubiquitination. Nature. 1992 Dec 10;360(6404):597-9.

Niedermann G, King G, Butz S, Birsner U, Grimm R, Shabanowitz J, Hunt DF, Eichmann K. The proteolytic fragments generated by vertebrate proteasomes: structural relationships to major histocompatibility complex class I binding peptides. Proc Natl Acad Sci U S A. 1996 Aug 6;93(16):8572-7.

Nielsen M, Lundegaard C, Lund O, Kesmir C. The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. Immunogenetics. 2005 Apr;57(1-2):33-41.

Nussbaum AK, Dick TP, Keilholz W, Schirle M, Stevanovic S, Dietz K, Heinemeyer W, Groll M, Wolf DH, Huber R, Rammensee HG, Schild H. Cleavage motifs of the yeast 20S proteasome beta subunits deduced from digests of enolase 1. Proc Natl Acad Sci U S A. 1998 Oct 13;95(21):12504-9.

Oosterkamp HM, Scheiner L, Stefanova MC, Lloyd KO, Finstad CL. Comparison of MUC-1 mucin expression in epithelial and non-epithelial cancer cell lines and demonstration of a new short variant form (MUC-1/Z). Int J Cancer. 1997 Jul 3;72(1):87-94.

Palmer M, Parker J, Modi S, Butts C, Smylie M, Meikle A, Kehoe M, MacLean G, Longenecker M. Phase I study of the BLP25 (MUC1 peptide) liposomal vaccine for active specific immunotherapy in stage IIIB/IV non-small-cell lung cancer. Clin Lung Cancer. 2001 Aug;3(1):49-57;

Pandey P, Kharbanda S, Kufe D. Association of the DF3/MUC1 breast cancer antigen with Grb2 and the Sos/Ras exchange protein. Cancer Res. 1995 Sep 15;55(18):4000-3.

Parkin DM. Global cancer statistics in the year 2000. Lancet Oncol. 2001 Sep;2(9):533-43. Review

Perez-Vilar J, Hill RL. The structure and assembly of secreted mucins.J Biol Chem. 1999 Nov 5;274(45):31751-4. Review

Pietersz GA, Li W, Osinski C, Apostolopoulos V, McKenzie IF. Definition of MHC-restricted CTL epitopes from non-variable number of tandem repeat sequence of MUC1. Vaccine. 2000 Apr 3;18(19):2059-71.

Ramanathan RK, Lee KM, McKolanis J, Hitbold E, Schraut W, Moser AJ, Warnick E, Whiteside T, Osborne J, Kim H, Day R, Troetschel M, Finn OJ. Phase I study of a MUC1 vaccine composed of different doses of MUC1 peptide with SB-AS2 adjuvant in resected and locally advanced pancreatic cancer. Cancer Immunol Immunother. 2005 Mar;54(3):254-64.

Rock KL. A new foreign policy: MHC class I molecules monitor the outside world. Immunol Today. 1996 Mar;17(3):131-7. Review.

Rock KL, York IA, Goldberg AL. Post-proteasomal antigen processing for major histocompatibility complex class I presentation. Nat Immunol. 2004 Jul;5(7):670-7. Review.

Rowse GJ, Tempero RM, VanLith ML, Hollingsworth MA, Gendler SJ. Tolerance and immunity to MUC1 in a human MUC1 transgenic murine model. Cancer Res. 1998 Jan 15;58(2):315-21.

Schmidtke G, Kraft R, Kostka S, Henklein P, Frömmel C, Löwe J, Huber R, Kloetzel PM and Schmidt M. Analysis of mammalian 20S proteasome biogenesis: the maturation of β-subunits is an ordered two-step mechanism involving autocatalysis. *EMBO J*, 1996, **15**, 6887–6898.

Scholl SM, Balloul JM, Le Goc G, Bizouarne N, Schatz C, Kieny MP, von Mensdorff-Pouilly S, Vincent-Salomon A, Deneux L, Tartour E, Fridman W, Pouillart P, Acres B. Recombinant vaccinia virus encoding human MUC1 and IL2 as immunotherapy in patients with breast cancer. J Immunother. 2000 Sep-Oct;23(5):570-80.

Schroeder JA, Thompson MC, Gardner MM, Gendler SJ. Transgenic MUC1 interacts with epidermal growth factor receptor and correlates with mitogen-activated protein kinase activation in the mouse mammary gland. J Biol Chem. 2001 Apr 20;276(16):13057-64.

Seemuller E, Lupas A, Stock D, Lowe J, Huber R, Baumeister W. Proteasome from Thermoplasma acidophilum: a threonine protease. Science. 1995 Apr 28;268(5210):579-82.

Seemüller E, Lupas A and Baumeister W. Autocatalytic processing of the 20S proteasome. *Nature*, 1996, **382**, 468–470.

Seifert U, Maranon C, Shmueli A, Desoutter JF, Wesoloski L, Janek K, Henklein P, Diescher S, Andrieu M, de la Salle H, Weinschenk T, Schild H, Laderach D, Galy A, Haas G, Kloetzel PM, Reiss Y, Hosmalin A. An essential role for tripeptidyl peptidase in the generation of an MHC class I epitope. Nat Immunol. 2003 Apr;4(4):375-9.

Shen L, Sigal LJ, Boes M, Rock KL. Important role of cathepsin S in generating peptides for TAP-independent MHC class I crosspresentation in vivo. Immunity. 2004 Aug;21(2):155-65.

Shimizu Y, Shaw S. Cell adhesion. Mucins in the mainstream. Nature. 1993 Dec 16;366(6456):630-1.

Smorodinsky N, Weiss M, Hartmann ML, Baruch A, Harness E, Yaakobovitz M, Keydar I, Wreschner DH. Detection of a secreted MUC1/SEC protein by MUC1 isoform specific monoclonal antibodies. Biochem Biophys Res Commun. 1996 Nov 1;228(1):115-21.

Snyder LA, Goletz TJ, Gunn GR, Shi FF, Harris MC, Cochlin K, McCauley C, McCarthy SG, Branigan PJ, Knight DM. A MUC1/IL-18 DNA vaccine induces anti-tumor immunity and increased survival in MUC1 transgenic mice. Vaccine. 2006 Apr 12;24(16):3340-52.

Soares MM, Mehta V, Finn OJ. Three different vaccines based on the 140-amino acid MUC1 peptide with seven tandemly repeated tumor-specific epitopes elicit distinct immune effector mechanisms in wild-type versus MUC1-transgenic mice with different potential for tumor rejection. J Immunol. 2001 Jun 1;166(11):6555-63.

Spicer A. P., Rowse G. J., Lidner T-K., Gendler S. J. Delayed mammary tumor progression in MUC-1 null mice. J. Biol. Chem., *270:* 30093-30101, 1995

Stepensky D, Tzehoval E, Vadai E, Eisenbach L. O-glycosylated versus non-glycosylated MUC1-derived peptides as potential targets for cytotoxic immunotherapy of carcinoma. Clin Exp Immunol. 2006 Jan;143(1):139-49.

Storkus WJ, Zeh HJ 3rd, Salter RD, Lotze MT. Identification of T-cell epitopes: rapid isolation of class I-presented peptides from viable cells by mild acid elution. J Immunother. 1993 Aug;14(2):94-103.

Suzuki T, Park H, Kwofie MA, Lennarz WJ. Rad23 provides a link between the Png1 deglycosylating enzyme and the 26 S proteasome in yeast. J Biol Chem. 2001 Jun 15;276(24):21601-7.

Tanaka Y, Koido S, Ohana M, Liu C, Gong J. Induction of impaired antitumor immunity by fusion of MHC class II-deficient dendritic cells with tumor cells. J Immunol. 2005 Feb 1;174(3):1274-80.

Ten Hagen KG, Tetaert D, Hagen FK, Richet C, Beres TM, Gagnon J, Balys MM, VanWuyckhuyse B, Bedi GS, Degand P, Tabak LA. Characterization of a UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase that displays glycopeptide N-acetylgalactosaminyltransferase activity. J Biol Chem. 1999 Sep 24;274(39):27867-74.

Toes RE, Nussbaum AK, Degermann S, Schirle M, Emmerich NP, Kraft M, Laplace C, Zwinderman A, Dick TP, Muller J, Schonfisch B, Schmid C, Fehling HJ, Stevanovic S, Rammensee HG, Schild H. Discrete cleavage motifs of constitutive and immunoproteasomes revealed by quantitative analysis of cleavage products. J Exp Med. 2001 Jul 2;194(1):1-12.

Unno M, Mizushima T, Morimoto Y, Tomisugi Y, Tanaka K, Yasuoka N, Tsukihara T. The structure of the mammalian 20S proteasome at 2.75 A resolution. Structure. 2002 May;10(5):609-18.

Vlad AM, Muller S, Cudic M, Paulsen H, Otvos L Jr, Hanisch FG, Finn OJ. Complex carbohydrates are not removed during processing of glycoproteins by dendritic cells: processing of tumor antigen MUC1 glycopeptides for presentation to major histocompatibility complex class II-restricted T cells. J Exp Med. 2002 Dec 2;196(11):1435-46.

Vlad AM, Kettel JC, Alajez NM, Carlos CA, Finn OJ. MUC1 immunobiology: from discovery to clinical applications. Adv Immunol. 2004;82:249-93. Review.

von Mensdorff-Pouilly S, Petrakou E, Kenemans P, van Uffelen K, Verstraeten AA, Snijdewint FG, van Kamp GJ, Schol DJ, Reis CA, Price MR, Livingston PO, Hilgers J. Reactivity of natural and induced human antibodies to MUC1 mucin with MUC1 peptides and n-acetylgalactosamine (GalNAc) peptides. Int J Cancer. 2000 Jun 1;86(5):702-12.

Walz G, Aruffo A, Kolanus W, Bevilacqua M, Seed B. Recognition by ELAM-1 of the sialyl-Lex determinant on myeloid and tumor cells. Science. 1990 Nov 23;250(4984):1132-5.

Wandall HH, Hassan H, Mirgorodskaya E, Kristensen AK, Roepstorff P, Bennett EP, Nielsen PA, Hollingsworth MA, Burchell J, Taylor-Papadimitriou J, Clausen H. Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3. J Biol Chem. 1997 Sep 19;272(38):23503-14.

Wesseling J, van der Valk SW, Vos HL, Sonnenberg A, Hilkens J. Episialin (MUC1) overexpression inhibits integrin-mediated cell adhesion to extracellular matrix components. J Cell Biol. 1995;129:255–265.

Wesseling J, van der Valk SW, Hilkens J. A mechanism for inhibition of E-cadherin-mediated cell-cell adhesion by the membrane-associated mucin episialin/MUC1. Mol Biol Cell. 1996 Apr;7(4):565-77.

Whitby FG, Masters EI, Kramer L, Knowlton JR, Yao Y, Wang CC, Hill CP. Structural basis for the activation of 20S proteasomes by 11S regulators. Nature. 2000 Nov 2;408(6808):115-20.

Whitehouse C, Burchell J, Gschmeissner S, Brockhausen I, Lloyd KO, Taylor-Papadimitriou J. A transfected sialyltransferase that is elevated in breast cancer and localizes to the medial/trans-Golgi apparatus inhibits the development of core-2-based O-glycans. J Cell Biol. 1997 Jun 16;137(6):1229-41.

Wierecky J, Mueller M, Brossart P. Dendritic cell-based cancer immunotherapy targeting MUC-1. Cancer Immunol Immunother. 2006 Jan;55(1):63-7.

Wierecky J, Muller MR, Wirths S, Halder-Oehler E, Dorfel D, Schmidt SM, Hantschel M, Brugger W, Schroder S, Horger MS, Kanz L, Brossart P. Immunologic and clinical responses after vaccinations with peptide-pulsed dendritic cells in metastatic renal cancer patients. Cancer Res. 2006 Jun 1;66(11):5910-8.

Wiertz EJ, Tortorella D, Bogyo M, Yu J, Mothes W, Jones TR, Rapoport TA, Ploegh HL. Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction. Nature. 1996 Dec 5;384(6608):432-8.

Wilkinson RW, Ross EL, Lee-MacAry AE, Laylor R, Burchell J, Taylor-Papadimitriou J, Snary D. A transgenic mouse model for tumour immunotherapy: induction of an anti-idiotype response to human MUC1. Br J Cancer. 2000 Nov;83(9):1202-8.

Wreschner D. H., Hareuveni M., Tsarfaty I., Smorodinsky N., Horev J., Zaretsky J., Kotkes P., Weiss M., Lathe R., Dion A. S., Keydar I. Human epithelial tumor antigen cDNA sequences—differential splicing may generate multiple protein forms. Eur. J. Biochem., *189:* 463-473, 1990.[Abstract]

Wreschner DH, Zrihan-Licht S, Baruch A, Sagiv D, Hartman ML, Smorodinsky N, Keydar I. Does a novel form of the breast cancer marker protein, MUC1, act as a receptor molecule that modulates signal transduction? Adv Exp Med Biol. 1994;353:17-26.

Yamamoto M, Bharti A, Li Y, Kufe D. Interaction of the DF3/MUC1 breast carcinoma-associated antigen and beta-catenin in cell adhesion. J Biol Chem. 1997 May 9;272(19):12492-4.

Yamamoto K, Ueno T, Kawaoka T, Hazama S, Fukui M, Suehiro Y, Hamanaka Y, Ikematsu Y, Imai K, Oka M, Hinoda Y. MUC1 peptide vaccination in patients with advanced pancreas or biliary tract cancer.Anticancer Res. 2005 Sep-Oct;25(5):3575-9.

Ye Y, Shibata Y, Yun C, Ron D, Rapoport TA. A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. Nature. 2004 Jun 24;429(6994):841-7.

Yoshida Y. Expression and assay of glycoprotein-specific ubiquitin ligases. Methods Enzymol. 2005;398:159-69.

Yuen CT, Lawson AM, Chai W, Larkin M, Stoll MS, Stuart AC, Sullivan FX, Ahern TJ, Feizi T. Novel sulfated ligands for the cell adhesion molecule E-selectin revealed by the neoglycolipid technology among O-linked oligosaccharides on an ovarian cystadenoma glycoprotein. Biochemistry. 1992 Sep 29;31(38):9126-31.

Zhang K, Baeckstrom D, Brevinge H, Hansson GC. Secreted MUC1 mucins lacking their cytoplasmic part and carrying sialyl-Lewis a and x epitopes from a tumor cell line and sera of colon carcinoma patients can inhibit HL-60 leukocyte adhesion to E-selectin-expressing endothelial cells. J Cell Biochem. 1996 Mar 15;60(4):538-49.

Zotter S, Hageman PC, Lossnitzer A, Mooi WJ, Hilgers J. Tissue and tumor distribution of human polymorphic epithelial mucin. In Hilgers J, Zotter S. (eds), Polymorphic Epithelial Mucin and CA125-bearing Glycoprotein As Tumor-Associated Antigens. Cancer Rev., Vol 11/12, pp. 55-101

Zrihan-Licht S, Baruch A, Elroy-Stein O, Keydar I, Wreschner DH. Tyrosine phosphorylation of the MUC1 breast cancer membrane proteins. Cytokine receptor-like molecules. FEBS Lett. 1994 Dec 12;356(1):130-6.

Zrihan-Licht S, Vos HL, Baruch A, Elroy-Stein O, Sagiv D, Keydar I, Hilkens J, Wreschner DH. Characterization and molecular cloning of a novel MUC1 protein, devoid of tandem repeats, expressed in human breast cancer tissue. Eur J Biochem. 1994 Sep 1;224(2):787-95.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen und Abbildungen -, die anderen Werke im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie noch nicht veröffentlicht ist, sowie, dass ich eine Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Herrn Prof. Dr. Franz-Georg Hanisch betreut worden.

Köln, den 20.08.2006.
Tanja Ninković

Publications:

Hanisch FG, Ninkovic T. Immunology of O-glycosylated Proteins: Approaches to the design of a MUC1 Glycopeptide-Based Tumor Vaccine, Current Protein and Peptide Science, 2006, 7, 307-315

Ninkovic T, Pisarev V, Hanisch FG. Processing of MUC1 tumor antigen in MHC class I pathway, Mucin in Health and Disease, 8th International Workshop on Carcinoma-associated Mucins, Robinson College, Cambridge (UK), 8 – 13 July 2005; Oral and poster presentation

Ninkovic T, Hansich FG. Processing of O-glycosylated MUC1 glycopeptides by immunoproteasomed for presentation in the MHC class I pathway, Joint meeting of the Glycobiology Society, Wageningen, Netherlands, Nov 2004. Poster and oral presentation

Lebenslauf/Curriculum Vitae

Persönliche Daten:

Name,Vorname:	Ninković, Tanja
Anschrift:	Kapitelstr 21
	51103 Köln
Geburtsdatum:	03.09.1977.
Geburtsort:	Jugoslawien/Serbien
Staatsangehörigkeit:	Serbin
Familienstand:	ledig
Eltern:	Radovan Ninković und Vesna Ninković

Schulausbildung:

1984-1992	Grundschule, Belgrad
1992-1996	Naturwissenschaftliches Gymnasium, Belgrad
Studium 1996-2002	Biologische Fakultät, Universität Belgrad, Institut für Molekularbiologie und Physiologie Hauptstudiengang: Experimentelle Biomedizin Abschluss mit der Durchschnittsnote 9,31 (auf einer Skala von 6 bis 10) Diplomprüfung mit der Note 10 Diplomarbeit: “Die Rolle des *c-myc* Onkogens in der malignen Transformation des oralen Mucus“
Feb 2003 -Okt 2006	Dissertation bei Prof. Dr. Franz-Georg Hanisch am Institut für Biochemie der Medizinischen Fakultät der Universität zu Köln mit dem Ziel derPromotion an der Math. Nat. Fakultät der Universität zu Köln

Köln den 20.08.06,

Tanja Ninković

Curriculum Vitae

Personal details

Name, Surname:	Ninković, Tanja
Adress:	Kapitelstr 21
	51103 Köln
Date of birth:	03.09.1977.
Place of birth:	Yugoslavia/Serbia
Citizenship:	Serbian
Family state:	Not married
Parents:	Radovan Ninković and Vesna Ninković

School education:

1984-1992	Primary School
1992-1996	Grammar school, section“ Natural Sciences“, Belgrade

Studies:

1996-2002	Faculty of Biology, University of Belgrade,
	Institute for Molecular Biology and Physiology
	Orientation: Experimenthal Biomedicine
	Finnished with average mark 9,31 (on the scale from 6 to 10)
	Master theses marked with 10
	Title of the Master thesis: “The role of *c-myc* oncogene in malignant transformation of the oral mucus”

Doctoral studies:

Feb 2003 -Okt 2006	PhD studies in the group of Prof. Dr. Franz-Georg Hanisch at Center for Biochemistry, Medical Faculty, University of Cologne

Köln 20. Aug 2006,

Tanja Ninković