New approaches in tumour immunology: Implications from mass spectrometric analyses of HLA-ligands

Neue Ansätze in der Tumorimmunologie: Implikationen aus der massenspektrometrischen Analyse von HLA-Liganden

DISSERTATION

der Fakultät für Chemie und Pharmazie
der Eberhard Karls-Universität Tübingen

zur Erlangung des Grades eines Doktors
der Naturwissenschaften

2008

vorgelegt von

Andreas Oliver Weinzierl

Tag der mündlichen Prüfung	31.01.2008
Dekan:	Prof. Dr. L. Wesemann
1. Berichterstatter	Prof. Dr. S. Stevanović
2. Berichterstatter	Prof. Dr. H.-G. Rammensee

Preface

Some chapters of this thesis have been published before. At the beginning of such chapters, it is indicated which experiments were done by the author of this thesis, which persons contributed to the publication, and in which journal the work has been published.

Contents

1. Introduction I: Tumour vs. normal tissue – Differences at various levels

In Germany each year about 400,000 new cancers are diagnosed and about half as many casualties are caused, making cancer Germany's second frequent cause of death. Tumourigenesis summarizes changes ranging from metabolic alterations within cells to physiologic changes in tissues or organs. These changes can promote tumour growth on the one hand but are principally also recognizable by the immune system and thus may provide a basis for the eradication of the tumour. The notion that the immune system could protect the body from neoplastic growth dates back to Paul Ehrlich [1] and was revived decades later as the 'cancer immunosurveillance hypothesis' [2, 3], which was at that time still heavily debated. Today it is well accepted that there are in fact substantial interactions between the immune system and tumour cells [4, 5] and that immune cells can have an active role controlling the malignancy [6, 7]. However, the amount of newly diagnosed tumours each year shows that in many cases tumours obviously find ways to escape the immune system and its surveillance due to their alterations in metabolism and physiology.

1.1. Differences from a physiological point of view

After the initial events of carcinogenesis, neoplastic growth is limited to a very restricted area if local physiological changes like angiogenesis can not be achieved. Moreover, neoplastic growth is frequently accompanied by local changes of the surrounding tissue resulting in the establishment of a complex tumour microenvironment.

1.1.1. Hypoxia and angiogenesis

Tumour tissues need their own supply with oxygen and nutrients for a growth beyond 1 mm^3 [8]. An insufficient oxygen supply can induce the expression of several angiogenic key molecules in tumour tissues. For example, the expression of the vascular endothelial growth factor (VEGF), a key regulator for neovascularization has been described as mediated by the expression of the hypoxia inducible factor 1 (HIF-1) in various tumour entities such as renal cell carcinoma (RCC) [9, 10]. However, even vascularized tumour areas differ distinctly in their

blood vessel pattern from normal vasculature. Normal vasculature is organized in a hierarchical manner, ensuring adequate nutrition and oxygen supply for all cells. In contrast to that, tumour vessels are chaotic, dilated, torturous, and often have a sluggish blood flow (reviewed in [11]). Thus, even the induction of angiogenesis in tumours does not prevent the presence of certain hypoxic and necrotic regions within the tumour. This was shown in several studies where a very heterogeneous O_2 distribution in tumours with many regions under a constant pO_2 insufficiency were detected (Table 1.1).

Table 1.1.: The partial oxygen pressure (pO_2) was measured in mmHg using an 'Eppendorf' electrode. pO_2 values are given as median; values in brackets represent the number of patients analysed. This table has been adapted from [12].

Tumour type	pO_2 tumour	pO_2 normal	References
Head and neck cancer	14.6(65)	51.2(65)	[13]
Lung cancer	7.5(17)	38.5(17)	[12]
Pancreatic cancer	2.7(7)	51.6(7)	[14]
Cervical cancer	5.0(8)	51.0(8)	[15]
Prostate cancer	2.4(59)	30.0(59)	[16]

Although hypoxia is generally an undesired state for the tumour, these areas bear a source for tumour cells with increased malignant potential. Hypoxic cells are not only hard to reach for oxygen but also for some chemotherapeutic agents [17]. Areas lacking sufficient oxygen supply show an increased level of gene expression regarding angiogenesis [9]. Furthermore hypoxia selects for cells which have lost sensitivity to p53 induced apoptosis and have an increased mutation rate [18]. Additionally, hypoxic states induce genes as P-glycoprotein, a transmembrane transporter mediating multi drug resistance [19]. Summarizing, hypoxia in tumours selects for cells with an increased malignant potential, including elevated mutation levels and increased metastatic potential [16, 20].

1.1.2. Tumour microenvironment and inflammation

Dvorak suggested the methapher 'tumours are wounds that do not heal' to describe the state in which an established tumour exists [21]. This state of inflammation is triggered by the constant disruption of homeostasis by proliferating tumour cells. Furthermore, stromal elements can also undergo alterations, promoting the inflammatory response and enhancing tumour growth e.g. by neovascularisation [22]. In the attempt to re-establish homeostasis, inflammatory cells as mast cells, granulocytes and macrophages are attracted (reviewed in [23] and [24]). However, in some cases – namely these cases, where a tumour manages to become established – these immune cells react in a way which promotes the

tumour cell survival and replication: For example, VEGF can be released in a neutrophil-dependent manner [25]. It has been shown that high levels of several cytokines and chemokines as VEGF, TGF-β, IL-6 and IL-10 enhance the immunosuppressive features of the tumour microenvironment [26]. In such an cytokine milieu, dendritic cells (DCs) tend to mature into a immunosuppressive phenotype expressing IDO [27] and co-inhibitory molecules such as B7-H1 [28]. These DCs induce arrest of the T cell cycle as well as T cell apoptosis and thus favor immune tolerance of the tumour. This tolerance is furthermore enhanced by the attraction of regulatory T cells to the tumour microenvironment respectively the differentiation of T cells into a regulatory phenotype within the tumour microenvironment [29].

1.2. Differences by losing key gene transcripts and proteins

Neoplastic changes transforming benign cells into malignant tumor cells are caused by alterations either in gene expression and/or protein levels. These changes can range from the loss of protein function regarding several key molecules to the additional expression of genes and proteins in the tumour – which will be discussed in section 1.3.2.

1.2.1. Inactivation of tumour suppressor genes

One of the most prominent examples of a tumour suppressor gene is p53, a major player involved in the regulation of cell cycle. A mutation of p53 occurs in about 50% of all human tumours [30]. The hotspot mutations (248R and 273R) in human p53 fail to bind the DNA consensus sequence, and thus lead to the loss of tumour suppressing activity of p53 [31].

An other example of a tumour suppressor, frequently mutated in tumours – especially in renal cell carcinoma – is the von Hippel-Lindau gene product (pVHL). pVHL is part of a ubiquitin ligase complex, targeting the hypoxia-inducible factor (HIF) for ubiquitinylation, provided oxygen is available [32]. Inactivation of pVHL leads to excessive VEGF production [33] due to constitutive HIF-1 [34] and HIF-2 [35] activation.

1.2.2. TAP deletion

The transporter associated with antigen processing 1 and 2 (TAP1/2) is crucial for efficient transport of peptides from the cytosol into the ER where peptides are bound to major histocompatibility complex I (MHC-I) molecules. Peptide-MHC complexes are transported to the cell surface and make the cell subsequently available for immune surveillance by T cells. Thus, tumours with a reduced TAP

functionality show a decreased MHC surface expression and therefore have an alleviated T cell mediated immune pressure. Mutations have frequently been reported in various tumour types, e.g. in more than 50% of cervical carcinoma [36]. TAP mutations can have various effects on several levels ranging from altered TAP-mRNA stability [37] to functional changes in peptide transport [38].

Furthermore several viruses with oncogenic potential are known to reduce either TAP mRNA levels (Adenovirus 12, [39]) or inhibit its peptide transporting function (US6, HCMV, [40]).

1.3. Differences as observable by the adaptive immune system

The immune system is potentially able to recognize every intracellular change leading to altered mRNA translation products and/or proteins, provided the alteration can be processed and presented on MHC molecules. Defective ribosomal products (DRiPs) [41] and/or proteins are degraded within the cell by proteases such as the proteasome [42]. Thus generated peptides are transported into the ER via TAP [43] and subsequently loaded on MHC-I. MHC-peptide complexes are transported from the ER onto the cell surface and are there accessible for T lymphocytes.

1.3.1. Antigen recognition by T lymphocytes

T lymphocytes mediate adaptive cellular immunity via their $\alpha\beta$ T cell receptor (TCR), which is present on the majority of T cells, and which recognizes MHC-peptide complexes [44]. T cells expressing the co-receptor CD8 ($CD8^+$ T cells) recognize MHC-I molecules with bound 8-12mer peptides, derived mainly from intracellular source proteins. CD4 expressing T cells ($CD4^+$ T cells) recognize 9-25mer peptides in the context of MHC-II molecules. Under normal conditions, MHC-II molecules are exclusively expressed on immune cells and the MHC-II bound peptides are mainly derived from extracellular sources [45].

In order to be able to respond properly to non- or altered-self peptides, presented on MHC-I molecules, $CD8^+$ cytotoxic T lymphocytes (CTLs) have to be activated by dendritic cells. DCs are professional antigen presenting cells, which are most effective in priming CTL responses [46]. After differentiation and activation, CTLs recognizing a specific epitope are able to migrate into the periphery to exert their effector function, e.g. migration into tumour tissue and eliciting an anti-tumour response. This response can for example be apoptosis of the target cell either via the Fas pathway and/or by perforin and granzyme [47]. Furthermore, activated CTLs can produce cytokines like interferon gamma (IFN-γ) which can inhibit tumour angiogenesis and proliferation [48].

1.3.2. Tumour antigens

For several different types of tumours it has been observed that even established tumours can be overcome by the body's own immune system [49]. The principle that tumours can be controlled by the immune system is the basis for cancer immunotherapy. For any cancer immunotherapeutic approach it is crucial to identify target structures, so called tumour antigens, which have to be as specific as possible for the tumour tissue. Tumour antigens in general are proteins that are produced by the tumour itself and are exclusively expressed in the tumour tissue. Hitherto known tumour antigens can be divided into five sub-classes: cancer-testis antigens, tumour specific antigens, tumour-virus antigens, differentiation antigens, and overexpressed antigens [50]. Yet, hitherto only about 100 tumour epitopes are known [51].

Table 1.2.: Selection of tumour antigens and their T cell epitopes. CDK4, cyclin-dependent kinase 4, CEA, carcinoembryonic antigen; HPV, human papillomavirus; MAGE, melanoma antigen; MUC1, mucin 1; PSMA, prostate-specific membrane antigen. This table has been adapted from [50]. A comprehensive listing of tumour antigens can be found in [51] or under http://www.cancerimmunity.org.

Antigen class	Tumour antigen	CTL epitope (sequence)	HLA	CD4 Helper epitope known	References
Cancer-testis	MAGE1	EADPTGHSY	A*0101	+	[52]
	MAGE3	FLWGPRALV	A*0201	+	[53]
	NY-ESO-1	SLLMWITQC	A*0201	+	[54]
Tumour specific	CDK4	ACDPHSGHFV	A*0201	-	[55]
	β-Catenin	SYLDSGIHF	A*2402	-	[56]
	Caspase-8	FPSDSWCYF	B*3501	-	[57]
Tumour virus	HPV16 E7	TLGIVCPI	A*0201	-	[58]
Differentiation	Tyrosinase	SEIWRDIDF	B*4402	+	[59]
	GP100	ALLAVGATK	A*0301	+	[60]
	PSMA	ALFDIESKV	A*0201	-	[61]
Overexpressed	ERBB2	KIFGSLAFL	A*0201	-	[62]
	MUC1	LLLLTVLTV	A*0201	-	[63]
	CEA	HLFGYSWYK	A*0301	-	[64]

In order to use CTLs for the elimination of tumour cells, T cell epitopes derived from such tumour antigens have to be defined. As the expression of most tumour antigens is frequently restricted to distinct tumour types and patients, they are not ubiquitously utilizable in tumour immunotherapy. Moreover, the actual HLA alleles expressed by a specific patient limit the known tumour-antigen-derived

epitopes which can be taken into account for an individual CTL-based tumour immunotherapy.

Due to this lack of tumour specific target structures, tumour immunotherapeutic approaches frequently resort additionally to tumour associated antigens (TAAs). TAAs are proteins which are highly overexpressed in tumour tissue but can also found in at least some normal tissues [50]. Therefore, all TAA based approaches bear the risk of inducing autoimmunity, in contrast to tumour antigen based approaches. TAAs can be subdivided into differentiation antigens and overexpressed antigens. Table 1.2 lists a selection of human tumor (associated) antigens with their corresponding T cell epitopes and HLA restrictions.

1.3.2.1. Cancer-testis antigens

Cancer-testis antigens are named according to their expression pattern. Regarding benign tissues, they are expressed exclusively in spermatocytes of the testis and sometimes in placenta [65]. As these cells lack HLA expression [66], CTLs in healthy individuals are normally not tolerized against these cancer-testis antigens. In certain tumours the expression of cancer-testis antigens can be induced and thus provide a target for CTL based tumour immunotherapy. However, as no physiological function of these molecules is known, an immune escape of the tumour by downregulation of the cancer-testis antigen is well possible [67].

1.3.2.2. Tumour specific antigens

Tumour specific antigens develop by amino acid changes in the protein sequence or by posttranslational modifications which do not occur in normal tissue. Examples of posttranslational modifications are altered glycosylation [68] or altered amino acid side chains [69]. Specific changes in the protein sequence can occur at the fusion site of chromosomal translocations, which has been described for hematopoietic malignancies like chronic myelogenous leukemia (CML) [70, 71]. Moreover, DNA point mutations [56] as well as alternatively spliced mRNA may result in tumour specific antigens. Furthermore, mRNA transcription can be initiated at non classical start codons leading to cryptic translation products [72, 73]. Tumour specific CTL epitopes can be generated by posttranslational splicing by the proteasome. However, not all epitopes derived from spliced peptides are tumour specific [74]. Finally, B cell lymphomas represent a special case: As they all share a common idiotypic B cell receptor, which is – due to its generation by somatic recombination – unique and thus tumour specific.

1.3.2.3. Tumour-virus antigens

Tumour-virus antigens are proteins which are introduced into a benign cell by an oncogenic virus and thus are present only in virus-infected cells. Therefore, an immunotherapy which is targeted against these viral proteins is highly specific.

It has been shown that herpes viruses (HHV) are the causative agent for 99.7% [75] of cervix carcinoma (thereof, more than 70% are caused by the types HPV-16 and HPV-18 [76]). Gardasil and Cervarix, two vaccines targeted against HPV L1, have shown to be of substantial impact regarding protective immunity against cervical cancers [77].

1.3.2.4. Differentiation antigens

Differentiation antigens are shared between the tumour and the normal tissue from which the tumour was originating. Most of these TAAs belong to the melanoma/melanocyte group and are involved in the melanin biosynthesis [51]. In order to use such TAAs in immunotherapy, a potential tolerance against these antigens has to be broken. Thus, autoimmunity against the original normal tissue can be triggered, leading for example to vitiligo in the case of melanocytes [78]. Compared to the potential therapeutic benefit, these side-effects are negligible.

An other example for differentiation antigens are prostate specific TAAs. As the prostate is frequently resected completely in prostate cancer patients, no adverse effects are to be anticipated using CTLs against prostate specific TAAs for tackling remaining prostate cancer cells.

1.3.2.5. Overexpressed antigens

This class of antigens is expressed in many normal tissues, however in lower levels compared to the tumour tissue. Although a marked T cell tolerance has to be expected, this group of TAAs is biggest in number [51]. Many of the overexpressed antigens are regulatory factors – needed e.g. for tumour survival or sustained growth – and thus can be found in various tumour types [79]. Therefore, they are applicable quite broadly as therapeutical targets. However, using overexpressed antigens for tumour immunotherapy both T cell tolerance and induction of autoimmunity are major impediments. Nevertheless, several overexpressed antigens have been used in clinical trials showing both low risk of autoimmunity and the possibility to induce specific T cells [80, 81].

1.3.3. Identification of TAAs

Although a considerable amount of TAA-derived T cell epitopes are known, the number is far from being sufficient for a comprehensive CTL-based tumour immunotherapy. This is can be partially explained by two biases in the identification process of tumour antigens and TAAs. On the one hand, many studies are performed using melanoma samples – a model cancer in tumour immunotherapy. On the other hand, studies are frequently performed with HLA-A*02 positive patients, as this is the most frequent MHC-I allele in the Caucasian population

[82]. Thus, CTL epitopes suitable for tumour immunotherapy are lacking for many tumour types and for less frequent HLA alleles.

1.3.3.1. Epitope mapping, isolation of CTLs and subsequent identification of T cell specificity

CTLs derived from tumour patients are isolated and used to screen expression libraries derived from tumour cells which are recognized by specific CTLs. Thus, the gene recognized by the tumour specific T cells can be identified. For example, by truncation of the TAA gene MAGE1 [83] and subsequent usage of overlapping 9-15 mer peptides, the A*01-restricted CTL epitope EADPTGHSY (MAGE1, aa 161-169) was identified [52].

Based on this approach, several further developments have been used for the identification of tumour-derived T cell epitopes. On the one hand, viral expression libraries can be used, which allow the enrichment of expression constructs recognized by specific CTLs [84]. On the other hand, MHC-bound peptides can be isolated directly from a cell line, recognized by tumour specific CTL lines. After separation of the isolated peptides by HPLC, fractions are tested for their potential to stimulate the tumour specific CTL lines. Using this approach, the peptide YLEPGPVTA (gp100 aa 280-288) was identified by mass spectrometry in the HPLC fraction recognized by the CTLs [85].

1.3.3.2. Reverse immunology, epitope prediction from defined antigens

As HLA molecules have an allele-specific peptide motif [86], CTL epitopes can be predicted for a combination of a tumour antigen of interest with an HLA allele of interest. This can be done using epitope prediction programs like BIMAS or SYFPEITHI. Additionally, proteasomal processing as well as binding affinity to the TAP molecule can be taken into account ([87]; PRED^{TAP}, [88]). Finally, all three prediction steps can be integrated in combined approaches (NetCTL, [89]; MHC-Pathway, [90]). For experimental validation of predicted T cell epitopes, the peptides of interest are synthesized and used to stimulate T cells derived from e.g. healthy blood donors or tumour patients. Finally, peptide specific T cells have to be tested against tumour cells expressing the antigen of interest, in order to verify that the identified peptide is processed naturally [91].

1.3.3.3. The 'Tübingen approach', combining mass spectrometry, gene expression analysis and T cell monitoring

The identification of patient specific tumor antigens or TAAs for tumor immunotherapy is still a bottleneck. Thus, the 'Tübingen approach' combines on a patient individual basis large scale identification on MHC-bound peptides from tumour tissue with gene expression analysis [92]. The peptide analysis identifies peptides actually presented on MHC, which are thus potentially recognizable by

CTLs. The gene expression analysis identifies which of the genes corresponding to the identified peptides are uniquely expressed or overexpressed in the tumour tissue. Thus, the combination of the identified peptide sequence and the gene expression data results in a patient specific list of candidate epitopes derived from TAAs. Finally, these candidate TAA epitopes are to be confirmed either by *in vitro* T cell assays [63, 93] or by immuno monitoring of *in vivo* T cell responses in patients vaccinated with candidate TAA epitopes.

1.4. Differences in tumour vs. normal tissue: Immunotherapeutic obstacles and their overcome

Although the identification of T cell epitopes derived from tumour antigens or TAAs is a prerequisite for a molecularly defined, CTL-based tumour immunotherapy, it is not sufficient for effective, long lasting anti-tumour responses *in vivo*. On the one hand, the tumour microenvironment can suppress immune responses actively by manipulating immune cells in a way that they e.g. produce immunosuppressive chemokines like TGF-β, IL-10, VEGF, and IL-6 [26] or express co-inhibitory ligands such as B7-H1 [28] and IDO [27]. On the other hand, it has been shown that $CD4^+$ T cell help is crucial for maintenance of CTL responses and the induction of $CD8^+$ memory T cells [94, 95]. Thus, an effective T cell-based immunotherapy has to be based on three major pillars:

First, the identification of as many tumour antigens and TAAs as possible. This would result in a large set of possible target structures, providing a broad arsenal from which the best candidates for an individual therapy could be chosen.

Second, the identification of the herein contained CD4 and CD8 T cell epitopes for maximal amount of HLA alleles. The combination of matching CD4 and CD8 T cell epitopes would allow a long lasting anti-tumour responses. The chance that tumour escaped the therapy could be further reduced, if several different epitope pairs would be used in parallel for a patient-specific therapy.

Third, appropriate therapeutical settings have to be found, which break the immunosuppressive features of the tumour microenvironment and thus allow an effective immune response.

1.5. References

[1] P. Ehrlich. Ueber den jetzigen Stand der Karzinomforschung. *Ned. Tijdschr. Geneeskd.*, 53:273–290, 1909.

[2] M. BURNET. Cancer; a biological approach. I. The processes of control. *Br. Med. J.*, 1(5022):779–786, 1957.

[3] F. M. Burnet. The concept of immunological surveillance. *Prog. Exp. Tumor Res.*, 13:1–27, 1970.

[4] D. Pardoll. Does the immune system see tumors as foreign or self? *Annu. Rev. Immunol.*, 21:807–839, 2003.

[5] G. P. Dunn, L. J. Old, and R. D. Schreiber. The immunobiology of cancer immunosurveillance and immunoediting. *Immunity.*, 21(2):137–148, 2004.

[6] T. A. Waldmann. Immunotherapy: past, present and future. *Nat. Med.*, 9 (3):269–277, 2003.

[7] J. N. Blattman and P. D. Greenberg. Cancer immunotherapy: a treatment for the masses. *Science*, 305(5681):200–205, 2004.

[8] A. D. Thornton, P. Ravn, M. Winslet, and K. Chester. Angiogenesis inhibition with bevacizumab and the surgical management of colorectal cancer. *Br. J. Surg.*, 93(12):1456–1463, 2006.

[9] D. Mukhopadhyay and K. Datta. Multiple regulatory pathways of vascular permeability factor/vascular endothelial growth factor (VPF/VEGF) expression in tumors. *Semin. Cancer Biol.*, 14(2):123–130, 2004.

[10] T. Shinojima, M. Oya, A. Takayanagi, R. Mizuno, N. Shimizu, and M. Murai. Renal cancer cells lacking hypoxia inducible factor (HIF)-1alpha expression maintain vascular endothelial growth factor expression through HIF-2alpha. *Carcinogenesis*, 28(3):529–536, 2007.

[11] J. M. Brown and A. J. Giaccia. The unique physiology of solid tumors: opportunities (and problems) for cancer therapy. *Cancer Res.*, 58(7):1408–1416, 1998.

[12] J. M. Brown and W. R. Wilson. Exploiting tumour hypoxia in cancer treatment. *Nat. Rev. Cancer*, 4(6):437–447, 2004.

[13] Q. T. Le, M. S. Kovacs, M. J. Dorie, A. Koong, D. J. Terris, H. A. Pinto, D. R. Goffinet, K. Nowels, D. Bloch, and J. M. Brown. Comparison of the comet assay and the oxygen microelectrode for measuring tumor oxygenation in head-and-neck cancer patients. *Int. J. Radiat. Oncol. Biol. Phys.*, 56(2): 375–383, 2003.

[14] A. C. Koong, V. K. Mehta, Q. T. Le, G. A. Fisher, D. J. Terris, J. M. Brown, A. J. Bastidas, and M. Vierra. Pancreatic tumors show high levels of hypoxia. *Int. J. Radiat. Oncol. Biol. Phys.*, 48(4):919–922, 2000.

[15] H. Lyng, K. Sundfor, and E. K. Rofstad. Oxygen tension in human tumours measured with polarographic needle electrodes and its relationship to vascular density, necrosis and hypoxia. *Radiother. Oncol.*, 44(2):163–169, 1997.

[16] D. M. Brizel, S. P. Scully, J. M. Harrelson, L. J. Layfield, J. M. Bean, L. R. Prosnitz, and M. W. Dewhirst. Tumor oxygenation predicts for the likelihood of distant metastases in human soft tissue sarcoma. *Cancer Res.*, 56(5):941–943, 1996.

[17] R. E. Durand. The influence of microenvironmental factors during cancer therapy. *In Vivo*, 8(5):691–702, 1994.

[18] T. G. Graeber, C. Osmanian, T. Jacks, D. E. Housman, C. J. Koch, S. W. Lowe, and A. J. Giaccia. Hypoxia-mediated selection of cells with diminished apoptotic potential in solid tumours. *Nature*, 379(6560):88–91, 1996.

[19] K. M. Comerford, T. J. Wallace, J. Karhausen, N. A. Louis, M. C. Montalto, and S. P. Colgan. Hypoxia-inducible factor-1-dependent regulation of the multidrug resistance (MDR1) gene. *Cancer Res.*, 62(12):3387–3394, 2002.

[20] M. Höckel, K. Schlenger, B. Aral, M. Mitze, U. Schaffer, and P. Vaupel. Association between tumor hypoxia and malignant progression in advanced cancer of the uterine cervix. *Cancer Res.*, 56(19):4509–4515, 1996.

[21] H. F. Dvorak. Angiogenesis: update 2005. *J. Thromb. Haemost.*, 3(8): 1835–1842, 2005.

[22] A. Albini and M. B. Sporn. The tumour microenvironment as a target for chemoprevention. *Nat. Rev. Cancer*, 7(2):139–147, 2007.

[23] A. Albini, F. Tosetti, R. Benelli, and D. M. Noonan. Tumor inflammatory angiogenesis and its chemoprevention. *Cancer Res.*, 65(23):10637–10641, 2005.

[24] K. E. de Visser, A. Eichten, and L. M. Coussens. Paradoxical roles of the immune system during cancer development. *Nat. Rev. Cancer*, 6(1):24–37, 2006.

[25] P. Scapini, M. Morini, C. Tecchio, S. Minghelli, E. Di Carlo, E. Tanghetti, A. Albini, C. Lowell, G. Berton, D. M. Noonan, and M. A. Cassatella. CXCL1/macrophage inflammatory protein-2-induced angiogenesis in vivo is mediated by neutrophil-derived vascular endothelial growth factor-A. *J. Immunol.*, 172(8):5034–5040, 2004.

[26] W. Zou. Immunosuppressive networks in the tumour environment and their therapeutic relevance. *Nat. Rev. Cancer*, 5(4):263–274, 2005.

[27] D. H. Munn, M. D. Sharma, J. R. Lee, K. G. Jhaver, T. S. Johnson, D. B. Keskin, B. Marshall, P. Chandler, S. J. Antonia, R. Burgess, Jr. C. L. Slingluff, and A. L. Mellor. Potential regulatory function of human dendritic cells expressing indoleamine 2,3-dioxygenase. *Science*, 297(5588):1867–1870, 2002.

[28] T. J. Curiel, S. Wei, H. Dong, X. Alvarez, P. Cheng, P. Mottram, R. Krzysiek, K. L. Knutson, B. Daniel, M. C. Zimmermann, O. David, M. Burow, A. Gordon, N. Dhurandhar, L. Myers, R. Berggren, A. Hemminki, R. D. Alvarez, D. Emilie, D. T. Curiel, L. Chen, and W. Zou. Blockade of B7-H1 improves myeloid dendritic cell-mediated antitumor immunity. *Nat. Med.*, 9(5):562–567, 2003.

[29] U. K. Liyanage, T. T. Moore, H. G. Joo, Y. Tanaka, V. Herrmann, G. Doherty, J. A. Drebin, S. M. Strasberg, T. J. Eberlein, P. S. Goedegebuure, and D. C. Linehan. Prevalence of regulatory T cells is increased in peripheral blood and tumor microenvironment of patients with pancreas or breast adenocarcinoma. *J. Immunol.*, 169(5):2756–2761, 2002.

[30] T. Soussi. p53 alterations in human cancer: more questions than answers. *Oncogene*, 26(15):2145–2156, 2007.

[31] A. Sigal and V. Rotter. Oncogenic mutations of the p53 tumor suppressor: the demons of the guardian of the genome. *Cancer Res.*, 60(24):6788–6793, 2000.

[32] Jr. W. G. Kaelin. The von Hippel-Lindau tumor suppressor protein and clear cell renal carcinoma. *Clin. Cancer Res.*, 13(2 Pt 2):680s–684s, 2007.

[33] G. Siemeister, K. Weindel, K. Mohrs, B. Barleon, G. Martiny-Baron, and D. Marme. Reversion of deregulated expression of vascular endothelial growth factor in human renal carcinoma cells by von Hippel-Lindau tumor suppressor protein. *Cancer Res.*, 56(10):2299–2301, 1996.

[34] P. H. Maxwell, M. S. Wiesener, G. W. Chang, S. C. Clifford, E. C. Vaux, M. E. Cockman, C. C. Wykoff, C. W. Pugh, E. R. Maher, and P. J. Ratcliffe. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature*, 399(6733):271–275, 1999.

[35] V. A. Carroll and M. Ashcroft. Role of hypoxia-inducible factor (HIF)-1alpha versus HIF-2alpha in the regulation of HIF target genes in response to hypoxia, insulin-like growth factor-I, or loss of von Hippel-Lindau function: implications for targeting the HIF pathway. *Cancer Res.*, 66(12):6264–6270, 2006.

[36] N. L. Fowler and I. H. Frazer. Mutations in TAP genes are common in cervical carcinomas. *Gynecol. Oncol.*, 92(3):914–921, 2004.

[37] T. Yang, B. A. McNally, S. Ferrone, Y. Liu, and P. Zheng. A single-nucleotide deletion leads to rapid degradation of TAP-1 mRNA in a melanoma cell line. *J. Biol. Chem.*, 278(17):15291–15296, 2003.

[38] U. Ritz, F. Momburg, H. P. Pircher, D. Strand, C. Huber, and B. Seliger. Identification of sequences in the human peptide transporter subunit TAP1 required for transporter associated with antigen processing (TAP) function. *Int. Immunol.*, 13(1):31–41, 2001.

[39] R. Rotem-Yehudar, S. Winograd, S. Sela, J. E. Coligan, and R. Ehrlich. Downregulation of peptide transporter genes in cell lines transformed with the highly oncogenic adenovirus 12. *J. Exp. Med.*, 180(2):477–488, 1994.

[40] P. J. Lehner, J. T. Karttunen, G. W. Wilkinson, and P. Cresswell. The human cytomegalovirus US6 glycoprotein inhibits transporter associated with antigen processing-dependent peptide translocation. *Proc. Natl. Acad. Sci. U. S. A*, 94(13):6904–6909, 1997.

[41] J. W. Yewdell, L. C. Anton, and J. R. Bennink. Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? *J. Immunol.*, 157(5):1823–1826, 1996.

[42] W. Baumeister, J. Walz, F. Zuhl, and E. Seemuller. The proteasome: paradigm of a self-compartmentalizing protease. *Cell*, 92(3):367–380, 1998.

[43] J. C. Shepherd, T. N. Schumacher, P. G. Ashton-Rickardt, S. Imaeda, H. L. Ploegh, Jr. C. A. Janeway, and S. Tonegawa. TAP1-dependent peptide translocation in vitro is ATP dependent and peptide selective. *Cell*, 74(3): 577–584, 1993.

[44] H. G. Rammensee, K. Falk, and O. Rötzschke. MHC molecules as peptide receptors. *Curr. Opin. Immunol.*, 5(1):35–44, 1993.

[45] B. Mach, V. Steimle, E. Martinez-Soria, and W. Reith. Regulation of MHC class II genes: lessons from a disease. *Annu. Rev. Immunol.*, 14:301–331, 1996.

[46] J. Banchereau and R. M. Steinman. Dendritic cells and the control of immunity. *Nature*, 392(6673):245–252, 1998.

[47] J. T. Harty and V. P. Badovinac. Influence of effector molecules on the CD8(+) T cell response to infection. *Curr. Opin. Immunol.*, 14(3):360–365, 2002.

[48] Z. Qin, J. Schwartzkopff, F. Pradera, T. Kammertoens, B. Seliger, H. Pircher, and T. Blankenstein. A critical requirement of interferon gamma-mediated angiostasis for tumor rejection by CD8+ T cells. *Cancer Res.*, 63 (14):4095–4100, 2003.

[49] M. Marchand, P. Weynants, E. Rankin, F. Arienti, F. Belli, G. Parmiani, N. Cascinelli, A. Bourlond, R. Vanwijck, and Y. Humblet. Tumor regression responses in melanoma patients treated with a peptide encoded by gene MAGE-3. *Int. J. Cancer*, 63(6):883–885, 1995.

[50] S. Stevanovic. Identification of tumour-associated T-cell epitopes for vaccine development. *Nat. Rev. Cancer*, 2(7):514–520, 2002.

[51] L. Novellino, C. Castelli, and G. Parmiani. A listing of human tumor antigens recognized by T cells: March 2004 update. *Cancer Immunol. Immunother.*, 54(3):187–207, 2005.

[52] C. Traversari, P. Van der Bruggen, I. F. Luescher, C. Lurquin, P. Chomez, A. Van Pel, E. De Plaen, A. Amar-Costesec, and T. Boon. A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic T lymphocytes directed against tumor antigen MZ2-E. *J. Exp. Med.*, 176(5): 1453–1457, 1992.

[53] P. Van der Bruggen, J. Bastin, T. Gajewski, P. G. Coulie, P. Boel, C. De Smet, C. Traversari, A. Townsend, and T. Boon. A peptide encoded by human gene MAGE-3 and presented by HLA-A2 induces cytolytic T lymphocytes that recognize tumor cells expressing MAGE-3. *Eur. J. Immunol.*, 24(12):3038–3043, 1994.

[54] E. Jäger, Y. T. Chen, J. W. Drijfhout, J. Karbach, M. Ringhoffer, D. Jäger, M. Arand, H. Wada, Y. Noguchi, E. Stockert, L. J. Old, and A. Knuth. Simultaneous humoral and cellular immune response against cancer-testis antigen NY-ESO-1: definition of human histocompatibility leukocyte antigen (HLA)-A2-binding peptide epitopes. *J. Exp. Med.*, 187(2):265–270, 1998.

[55] T. Wölfel, M. Hauer, J. Schneider, M. Serrano, C. Wölfel, E. Klehmann-Hieb, E. De Plaen, T. Hankeln, K. H. Meyer zum Buschenfelde, and D. Beach. A p16INK4a-insensitive CDK4 mutant targeted by cytolytic T lymphocytes in a human melanoma. *Science*, 269(5228):1281–1284, 1995.

[56] P. F. Robbins, M. El Gamil, Y. F. Li, Y. Kawakami, D. Loftus, E. Appella, and S. A. Rosenberg. A mutated beta-catenin gene encodes a melanoma-specific antigen recognized by tumor infiltrating lymphocytes. *J. Exp. Med.*, 183(3):1185–1192, 1996.

[57] S. Mandruzzato, V. Stroobant, N. Demotte, and Bruggen P. van der. A human CTL recognizes a caspase-8-derived peptide on autologous HLA-B*3503 molecules and two unrelated peptides on allogeneic HLA-B*3501 molecules. *J. Immunol.*, 164(8):4130–4134, 2000.

[58] M. E. Ressing, A. Sette, R. M. Brandt, J. Ruppert, P. A. Wentworth, M. Hartman, C. Oseroff, H. M. Grey, C. J. Melief, and W. M. Kast. Human CTL epitopes encoded by human papillomavirus type 16 E6 and E7 identified through in vivo and in vitro immunogenicity studies of HLA-A*0201-binding peptides. *J. Immunol.*, 154(11):5934–5943, 1995.

[59] V. G. Brichard, J. Herman, A. Van Pel, C. Wildmann, B. Gaugler, T. Wölfel, T. Boon, and B. Lethe. A tyrosinase nonapeptide presented by HLA-B44 is recognized on a human melanoma by autologous cytolytic T lymphocytes. *Eur. J. Immunol.*, 26(1):224–230, 1996.

[60] J. C. Skipper, D. J. Kittlesen, R. C. Hendrickson, D. D. Deacon, N. L. Harthun, S. N. Wagner, D. F. Hunt, V. H. Engelhard, and Jr. C. L. Slingluff. Shared epitopes for HLA-A3-restricted melanoma-reactive human CTL include a naturally processed epitope from Pmel-17/gp100. *J. Immunol.*, 157 (11):5027–5033, 1996.

[61] M. L. Salgaller, B. A. Tjoa, P. A. Lodge, H. Ragde, G. Kenny, A. Boynton, and G. P. Murphy. Dendritic cell-based immunotherapy of prostate cancer. *Crit Rev. Immunol.*, 18(1-2):109–119, 1998.

[62] B. Fisk, T. L. Blevins, J. T. Wharton, and C. G. Ioannides. Identification of an immunodominant peptide of HER-2/neu protooncogene recognized by ovarian tumor-specific cytotoxic T lymphocyte lines. *J. Exp. Med.*, 181(6): 2109–2117, 1995.

[63] P. Brossart, K. S. Heinrich, G. Stuhler, L. Behnke, V. L. Reichardt, S. Stevanovic, A. Muhm, H. G. Rammensee, L. Kanz, and W. Brugger. Identification of HLA-A2-restricted T-cell epitopes derived from the MUC1 tumor antigen for broadly applicable vaccine therapies. *Blood*, 93(12):4309–4317, 1999.

[64] I. Kawashima, V. Tsai, S. Southwood, K. Takesako, A. Sette, and E. Celis. Identification of HLA-A3-restricted cytotoxic T lymphocyte epitopes from carcinoembryonic antigen and HER-2/neu by primary in vitro immunization with peptide-pulsed dendritic cells. *Cancer Res.*, 59(2):431–435, 1999.

[65] E. De Plaen, K. Arden, C. Traversari, J. J. Gaforio, J. P. Szikora, C. De Smet, F. Brasseur, Bruggen P. van der, B. Lethe, and C. Lurquin. Structure, chromosomal localization, and expression of 12 genes of the MAGE family. *Immunogenetics*, 40(5):360–369, 1994.

[66] A. Jassim, W. Ollier, A. Payne, A. Biro, R. T. Oliver, and H. Festenstein. Analysis of HLA antigens on germ cells in human semen. *Eur. J. Immunol.*, 19(7):1215–1220, 1989.

[67] L. J. Old. Cancer/testis (CT) antigens - a new link between gametogenesis and cancer. *Cancer Immun.*, 1:1, 2001.

[68] A. M. Vlad, S. Muller, M. Cudic, H. Paulsen, Jr. L. Otvos, F. G. Hanisch, and O. J. Finn. Complex carbohydrates are not removed during processing of glycoproteins by dendritic cells: processing of tumor antigen MUC1 glycopeptides for presentation to major histocompatibility complex class II-restricted T cells. *J. Exp. Med.*, 196(11):1435–1446, 2002.

[69] J. C. Skipper, R. C. Hendrickson, P. H. Gulden, V. Brichard, A. Van Pel, Y. Chen, J. Shabanowitz, T. Wolfel, Jr. C. L. Slingluff, T. Boon, D. F. Hunt, and V. H. Engelhard. An HLA-A2-restricted tyrosinase antigen on melanoma cells results from posttranslational modification and suggests a novel pathway for processing of membrane proteins. *J. Exp. Med.*, 183(2): 527–534, 1996.

[70] P. Yotnda, H. Firat, F. Garcia-Pons, Z. Garcia, G. Gourru, J. P. Vernant, F. A. Lemonnier, V. Leblond, and P. Langlade-Demoyen. Cytotoxic T cell response against the chimeric p210 BCR-ABL protein in patients with chronic myelogenous leukemia. *J. Clin. Invest*, 101(10):2290–2296, 1998.

[71] W. M. Wagner, Q. Ouyang, and G. Pawelec. The abl/bcr gene product as a novel leukemia-specific antigen: peptides spanning the fusion region of abl/bcr can be recognized by both CD4+ and CD8+ T lymphocytes. *Cancer Immunol. Immunother.*, 52(2):89–96, 2003.

[72] S. R. Schwab, J. A. Shugart, T. Horng, S. Malarkannan, and N. Shastri. Unanticipated antigens: translation initiation at CUG with leucine. *PLoS. Biol.*, 2(11):e366, 2004.

[73] H. Torikai, Y. Akatsuka, M. Miyazaki, III E. H. Warren, T. Oba, K. Tsujimura, K. Motoyoshi, Y. Morishima, Y. Kodera, K. Kuzushima, and T. Takahashi. A novel HLA-A*3303-restricted minor histocompatibility antigen encoded by an unconventional open reading frame of human TMSB4Y gene. *J. Immunol.*, 173(11):7046–7054, 2004.

[74] E. H. Warren, N. J. Vigneron, M. A. Gavin, P. G. Coulie, V. Stroobant, A. Dalet, S. S. Tykodi, S. M. Xuereb, J. K. Mito, S. R. Riddell, and B. J. Van den Eynde. An antigen produced by splicing of noncontiguous peptides in the reverse order. *Science*, 313(5792):1444–1447, 2006.

[75] J. M. Walboomers, M. V. Jacobs, M. M. Manos, F. X. Bosch, J. A. Kummer, K. V. Shah, P. J. Snijders, J. Peto, C. J. Meijer, and N. Munoz. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J. Pathol.*, 189(1):12–19, 1999.

[76] F. X. Bosch, A. Lorincz, N. Munoz, C. J. Meijer, and K. V. Shah. The causal relation between human papillomavirus and cervical cancer. *J. Clin. Pathol.*, 55(4):244–265, 2002.

[77] E. Hanna and G. Bachmann. HPV vaccination with Gardasil: a breakthrough in women's health. *Expert. Opin. Biol. Ther.*, 6(11):1223–1227, 2006.

[78] V. H. Engelhard, T. N. Bullock, T. A. Colella, S. L. Sheasley, and D. W. Mullins. Antigens derived from melanocyte differentiation proteins: self-tolerance, autoimmunity, and use for cancer immunotherapy. *Immunol. Rev.*, 188:136–146, 2002.

[79] D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 100(1): 57–70, 2000.

[80] S. H. van der Burg, A. G. Menon, A. Redeker, M. C. Bonnet, J. W. Drijfhout, R. A. Tollenaar, C. J. van de Velde, P. Moingeon, P. J. Kuppen, R. Offringa, and C. J. Melief. Induction of p53-specific immune responses in colorectal cancer patients receiving a recombinant ALVAC-p53 candidate vaccine. *Clin. Cancer Res.*, 8(5):1019–1027, 2002.

[81] R. K. Ramanathan, K. M. Lee, J. McKolanis, E. Hitbold, W. Schraut, A. J. Moser, E. Warnick, T. Whiteside, J. Osborne, H. Kim, R. Day, M. Troetschel, and O. J. Finn. 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.*, 54 (3):254–264, 2005.

[82] K. Cao, J. Hollenbach, X. Shi, W. Shi, M. Chopek, and M. A. Fernandez-Vina. Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Hum. Immunol.*, 62(9):1009–1030, 2001.

[83] P. Van der Bruggen, C. Traversari, P. Chomez, C. Lurquin, E. De Plaen, Eynde B. Van den, A. Knuth, and T. Boon. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science*, 254 (5038):1643–1647, 1991.

[84] E. S. Smith, A. Mandokhot, E. E. Evans, L. Mueller, M. A. Borrello, D. M. Sahasrabudhe, and M. Zauderer. Lethality-based selection of recombinant genes in mammalian cells: application to identifying tumor antigens. *Nat. Med.*, 7(8):967–972, 2001.

[85] A. L. Cox, J. Skipper, Y. Chen, R. A. Henderson, T. L. Darrow, J. Shabanowitz, V. H. Engelhard, D. F. Hunt, and Jr. C. L. Slingluff. Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. *Science*, 264(5159):716–719, 1994.

[86] K. Falk, O. Rötzschke, S. Stevanovic, G. Jung, and H. G. Rammensee. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*, 351(6324):290–296, 1991.

[87] B. Peters, S. Bulik, R. Tampe, P. M. Van Endert, and H. G. Holzhütter. Identifying MHC class I epitopes by predicting the TAP transport efficiency of epitope precursors. *J. Immunol.*, 171(4):1741–1749, 2003.

[88] G. L. Zhang, N. Petrovsky, C. K. Kwoh, J. T. August, and V. Brusic. PRED(TAP): a system for prediction of peptide binding to the human transporter associated with antigen processing. *Immunome Res.*, 2:3, 2006.

[89] M. V. Larsen, C. Lundegaard, K. Lamberth, S. Buus, S. Brunak, O. Lund, and M. Nielsen. An integrative approach to CTL epitope prediction: a combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions. *Eur. J. Immunol.*, 35(8):2295–2303, 2005.

[90] S. Tenzer, B. Peters, S. Bulik, O. Schoor, C. Lemmel, M. M. Schatz, P. M. Kloetzel, H. G. Rammensee, H. Schild, and H. G. Holzhütter. Modeling the MHC class I pathway by combining predictions of proteasomal cleavage, TAP transport and MHC class I binding. *Cell Mol. Life Sci.*, 62(9):1025–1037, 2005.

[91] S. Pascolo, M. Schirle, B. Gückel, T. Dumrese, S. Stumm, S. Kayser, A. Moris, D. Wallwiener, H. G. Rammensee, and S. Stevanovic. A MAGE-A1 HLA-A A*0201 epitope identified by mass spectrometry. *Cancer Res.*, 61(10):4072–4077, 2001.

[92] H. Singh-Jasuja, N. P. Emmerich, and H. G. Rammensee. The Tübingen approach: identification, selection, and validation of tumor-associated HLA peptides for cancer therapy. *Cancer Immunol. Immunother.*, 53(3):187–195, 2004.

[93] S. M. Schmidt, K. Schag, M. R. Müller, T. Weinschenk, S. Appel, O. Schoor, M. M. Weck, F. Grünebach, L. Kanz, S. Stevanovic, H. G. Rammensee, and

P. Brossart. Induction of adipophilin-specific cytotoxic T lymphocytes using a novel HLA-A2-binding peptide that mediates tumor cell lysis. *Cancer Res.*, 64(3):1164–1170, 2004.

[94] E. M. Janssen, E. E. Lemmens, T. Wolfe, U. Christen, M. G. von Herrath, and S. P. Schoenberger. CD4+ T cells are required for secondary expansion and memory in CD8+ T lymphocytes. *Nature*, 421(6925):852–856, 2003.

[95] D. J. Shedlock and H. Shen. Requirement for CD4 T cell help in generating functional CD8 T cell memory. *Science*, 300(5617):337–339, 2003.

2. Introduction II: LC-MS based protein and peptide quantification using stable isotope labels

This chapter is part of a review article published in *Biotechnology and Genetic Engeneering Reviews*, 23:21-39, 2006 by:
Andreas O. Weinzierl and Stefan Stevanović.

The author of this thesis was majorly responsible for writing the article and designed Figures 2.1 and 2.2.

2.1. Introduction

Detailed quantitative comparisons of different cellular states are a key prerequisite for the profound understanding of various complex cellular pathways on a molecular basis. These molecular mechanisms of interest can range from regulatory consequences of a triggered signal cascade to a deep understanding of malignant cellular behaviors like in tumors. Different levels of quantitative comparisons of cellular states became a feasible technique in the last years.

Regarding the transcriptome, quantification has been well established for many years [1]. Single gene products can be quantified using qRT-PCR and even the complete entity of different transcriptomes are quantitatively comparable in one single gene chip experiment. But due to poor correlation between the transcriptome and the proteome [2, 3] no general statements can be deduced from quantitative mRNA results in respect to changes in protein expression.

In addition to the lack of correlation between the amount of a mRNA species and its corresponding protein, mRNA expression analyses are blind for post-transcriptional modifications of proteins [4]. Thus up to now a molecular analysis of a cellular state has to be carried out on the proteome level. The higher chemical complexity of protein mixtures makes the proteome not as easily quantifiable as the transcriptome, but within the last decade several approaches made quantitative comparisons between proteomes possible in a high throughput manner [5].

In general all these approaches dealing with protein quantification have to solve several general problems. At first, the complexity of the proteome has to be reduced to an analyzable size, at best without losing valuable information and sensitivity. Second the information gained in the experiment has to be interpretable in a high-throughput way and at last the method has to be reproducible to be able to detect also minor changes in protein levels.

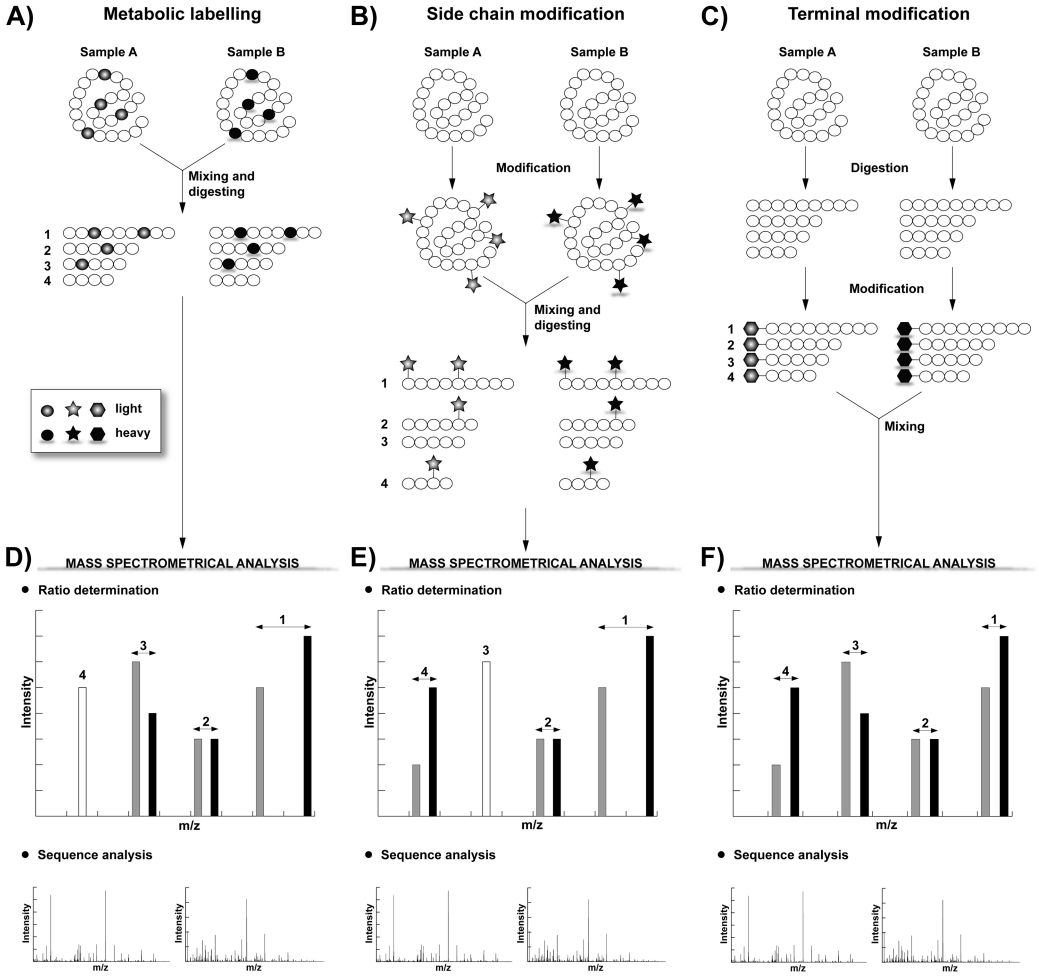

Figure 2.1.: Schematic representation of stable isotope labeling sites for protein labeling in quantitative proteomics. Light isotope marks are shaded in grey, heavy isotope marks in black and unmarked amino acids are white. A and D: Metabolic labeling: Proteins are labeled metabolically by incorporation of ^{15}N, ^{13}C or isotope labeled amino acids present in the cell culture medium. B and E: Side chain modification: Proteins are labeled at specific side chains with isotope containing reagents. The reagents can also contain affinity tags for selective enrichment of tagged proteins. C and F: Terminal modification: Proteins are labeled C- or N-terminally using e.g. enzymatic cleavage in $H_2{}^{18}O$ or chemical modification at the N-terminus with isotope containing reagents. Samples are digested after isotope labeling respectively for terminal labeling are labeled after digestion. Afterwards samples are quantitatively analyzed by mass spectrometry (D-F, ratio determination). For metabolic labeling and side chain modification various differences in the distance between light and heavy modified peptides (D and E) as well as unmodified peptides (D, peptide 4 and E, peptide 3) can be observed. Terminal modification shows a more uniform peptide pattern as no unmodified peptides emerge and distances between light and heavy modified peptides are equal (F). In a final step peptides are sequenced by tandem MS/MS experiments (D-F, sequence analysis).

2.2. 2D-Gel electrophoresis based quantification techniques

The fist attempt in protein quantification which met many of the prerequisites given above was two-dimensional gel electrophoresis (2D-GE) [6, 7]. Here the complexity of a protein mixture is reduced by isoelectric focusing and subsequent size separation of the proteins. After separation of the proteins they are stained in gel, isolated, enzymatically digested and identified by MALDI-TOF based peptide mapping [8, 9]. Although 2D-GE coupled to MS-analysis is in general a valuable approach in quantitative proteome analysis, the classic quantitative 2D-GE has some major impediments. The lack of reproducibility between different gels hampers the distinction between variations of the sample and the system. Sensitivity of the system is limited e.g. due to the used protein dye for protein visualization and to the protein recovery rate from the gel [10]. A major improvement of quantitative 2D-GE was achieved by Unlu and co-workers [11]. They introduced the difference gel electrophoresis (DIGE) allowing analysis of several samples to be compared in one single 2D gel with increased sensitivity and reproducibility. In this approach up to three samples are prelabeled with spectrally different fluorescence dyes, afterwards the samples are mixed and separated in the same gel. The consecutive spot detection is carried out with a scanner after laser excitation of the fluorescence dye. Protein identification is still carried out after spot isolation and enzymatic digest by MALDI-TOF. Yet one last major issue in 2D-GE proteomics is high throughput analysis mainly exacerbated by protein spot picking from the gel. At least for coomassie stained gels this can be automated using robots [12] yet accuracy and sensitivity of this automation is still technically elaborated.

2.3. LC-MS based quantification techniques

In order to circumvent difficulties resulting from 2D gel electrophoretic separation of proteins alternative quantitative proteomics techniques have been developed. Common to all of these techniques is the replacement of gel electrophoresis by liquid chromatography (LC) for protein separation and complexity reduction. In general in all techniques the proteins are enzymatically digested prior to separation, peptides are separated chromatographically and directly analyzed with an online coupled mass spectrometer (LC-MS). At some point in time of that workflow, depending on the respective approach, the peptide mixture is differentially labeled with isotope-coded tags. This approach is advantageous in several respects. On the one hand peptide separation and data acquisition are automated per se and therefore high throughput analysis is easily possible. On the other hand the method is highly reproducible and therefore sensitive even for minor changes in protein levels.

The fundamental obstacle in quantification by mass spectrometry is the fact that the signal a peptide yields in an MS-experiment does not correlate to its total amount and is not accurately reproducible in consecutive experiments. This is true for MS experiments irrespective of the ionization technique or the used MS device [13]. The ionization efficiency of a specific peptide depends on the peptide sequence and on the nature and amount of co-eluting peptides which are not necessarily constant especially when analyzing quantitative differences between proteomes. This obstacle is circumvented by using isotope-coded labels which differ in mass but not in their physio-chemical behavior on the LC-column [14]. Therefore both species of a differentially labeled peptide elute at the same retention time from the chromatography column but are distinguishable in the recorded mass spectrum due to their different isotope-coded label. In this case, the ionization efficiency for both peptide species is equal and relative quantification is possible. This approach can principally also be used for absolute quantification (AQUA, [15]) provided that one isotopically labeled peptide species is a spiked peptide of known amount.

This basic principle depicted above comes in several different flavors differing mainly in kind of the label itself and in place of label introduction (Fig. 2.1-A to C). In the following section various isotope-coded modification reagents developed in recent years are discussed. Detailed examples of quantitative results using the differential N-terminal isotope coding (dNIC) strategy are given in the last section.

2.3.1. Metabolic labeling

Metabolic labeling of proteins occurs during protein biosynthesis within the cell (Fig. 2.1-A). Therefore cells have to be grown in media enriched in the stable isotopes ^{15}N [16, 17] or ^{13}C [18] or which are supplemented with isotope-incorporated amino acids like D_3-leucine [19] and arginine containing various numbers of ^{13}C and ^{15}N e.g. $^{13}C_6$-Arg, $^{13}C_6$-$^{15}N_4$-Arg (both [20] or $^{15}N_2$-Arg [21] (Fig. 2.2).

The usage of ^{15}N enriched medium has proven to be an accurate quantification method and is more frequently used than ^{13}C-modified media. In order to reduce difficulties and expenses in preparing ^{15}N enriched media for mammalian cells the usage of stable isotope labeling by amino acids in cell culture (SILAC) was introduced in 2002 for differential proteomics by Ong et al 2002. Several quantitative analyses have been carried out ranging from analysis of proteome dynamics of the nucleolus [20], proteome changes in muscle cell differentiation [19] to repertoire changes of major histocompatibility complex class I (MHC-I) peptides after *Salmonella enterica* infection [21].

A principle obstacle concerning metabolic labeling is that the number of heavy atoms respectively amino acids incorporated into one peptide is not deducible without sequence information. As mass differences between isotopic labeled peptides depend on the sequence they differ from peptide to peptide (Fig. 2.1-A

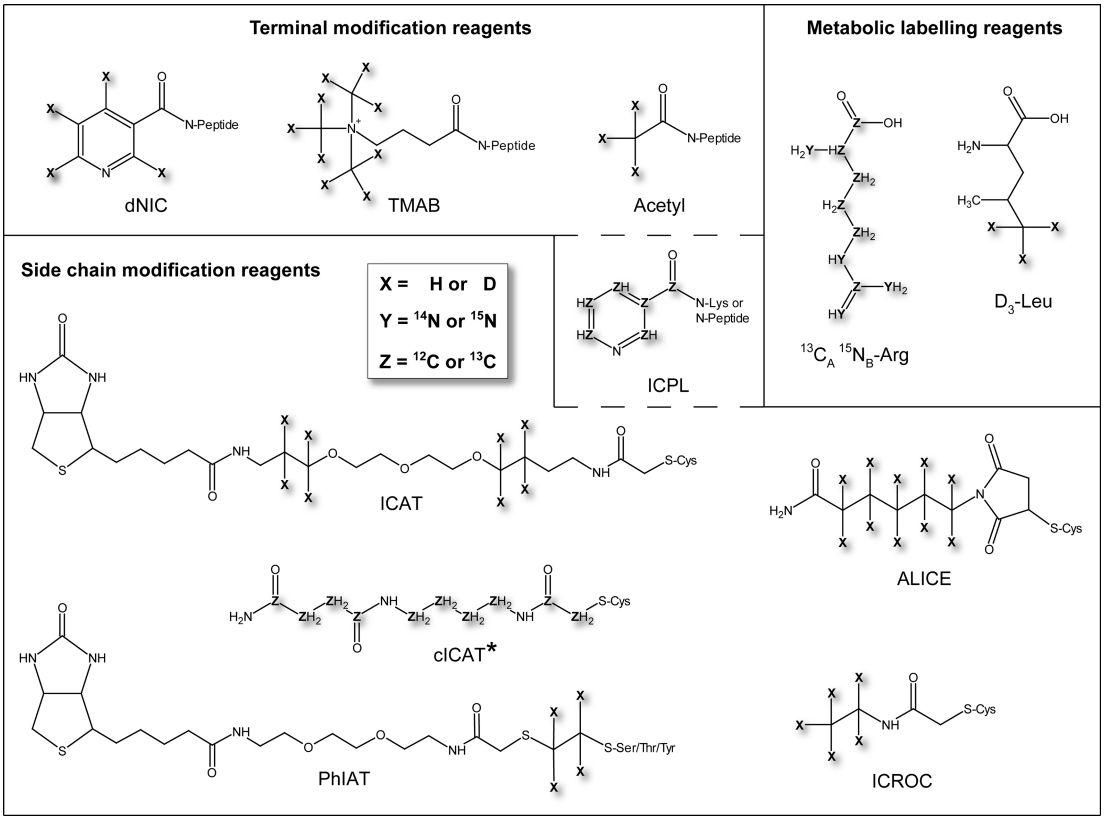

Figure 2.2.: Structure formulas of isotope bearing modification reagents described in this review. Sites bearing heavy isotopes in the respective heavy isotope-reagents are marked with X, Y and Z. For further details regarding the reagents see the respective sections of the text. * The heavy cICAT reagent contains nine ^{13}C-atoms out of the ten C-atoms in total. The exact position of these nine ^{13}C-atoms is a corporate secret of Applied Biosystems, Framingham, MA.

and D, compare peptides 1-4). An additional restraint is the fact that highly abundant peptides leading to just one MS-signal are not distinguishable from also existing unlabeled peptides (Fig. 2.1-D, peptide 4). This in turn reduces the impact of quantifiable MS-data without sequence information dramatically.

2.3.2. Chemical labeling

For proteomic samples which cannot be derived from *in vitro* cultured cells the metabolic labeling approach is not applicable. These samples can be chemically modified with isotope labels at their terminus or at side chains. This chemical modification can be divided into side chain modifications (Fig. 2.1-B and E) and terminal modifications (Fig. 2.1-C and F). The major difference between side chain modification and terminal modification is not only the modification reagent (see Fig. 2.2 dNIC and ICPL) but mainly the number of peptides available for

modification and thus for quantification. After side chain labeling, only peptides containing the targeted amino acid are modified and the number of labels per peptide is also sequence dependent (Fig. 2.1-B and E, compare peptides 1-4), whereas terminal labeling introduces exactly one label in every peptide (Fig. 2.1-C and F, compare peptides 1-4).

2.3.3. Side chain labeling

Favorite targets for specific modifications of side chains are mainly the amino group of lysine residues or the thiol group of cysteines. In principle proteins are at first isolated then modified and digested enzymatically and subsequently subjected to MS-analysis (Fig. 2.1-B).

2.3.3.1. ICAT and cICAT

Gygi et al introduced in 1999 the first isotope-coded affinity tag (ICAT). The ICAT reagent consists of a biotin tag, an isotope carrying linker and a reactive sulfhydryl group for the attachment of the ICAT reagent to reduced cysteine residues. The biotin tag allows the purification of labeled peptides, reducing the sample complexity on the one hand and omitting singular signals from unmodified peptides on the other. Although this ICAT reagent was successfully used in the analysis of differences in the *Saccaromyces cerevisiae* proteome, induced by galactose or ethanol as carbon source [2] the ICAT reagent itself has some restraints. The biotin isotope tag is relatively large (442 Da) and can result in incomplete peptide sequence information [13], the purification via the biotin-avidin interaction is susceptible to unspecific binding and incomplete elution [22] and finally the eight deuterium atoms of the heavy ICAT reagent caused a relatively strong isotope effect on the chromatography column, which could partially resolve isotopic peptide pairs [23]. These problems were largely solved with several improvements of the ICAT reagent. The cleavable ICAT (cICAT, Fig. 2.2) has two major improved features. It contains an acid-cleavable linker, which allows removal of the biotin tag from the labeled peptide. This removal of the biotin decreases the mass of the ICAT-tag for MS-analysis and results in better interpretable peptide fragmentation spectra. Additionally the isotope effect in chromatography was omitted by replacement of the deuterium with ^{13}C-isotope labels [24].

2.3.3.2. ICROC and ALICE

Several other strategies vary the basic idea of ICAT modification. For the isotope-coded reduction off of a chromatographic support (ICROC, Fig. 2.2) method [25] proteins are also reduced and cleaved with trypsin. The thus generated cysteine containing peptides are covalently coupled to pyridyl disulfide beads.

After washing the cysteine peptides are eluted using a reducing agent. Afterwards the free cysteine thiol groups are alkylated either with N-ethyl or with N-D_5-ethyl-iodacetamide and subjected to LC-MS for quantitative analysis.

The usage of acid-labile isotope-coded extractants (ALICE, Fig. 2.2) [26] also circumvents biotin/avidin based purification of cysteine-peptides. Cysteine containing peptides derived from tryptic digests are linked with their thiol group to the ALICE reagent, which contains apart from a thiol group an isotope-coded linker and an acid labile unit, which is covalently coupled to an inorganic resin. After linking the cysteine peptides to the resin, they are washed and afterwards eluted by acid treatment and analyzed consecutively.

2.3.3.3. PhIAT

The principle of ICAT was also conveyed to the quantitative analysis of phosphorylated proteins using phosphoprotein isotope-coded affinity tags (PhIAT). In 2001 Oda et al. [27] developed a labeling strategy for phosphorylated proteins with a reagent resembling ICAT. Proteins were subjected to tryptic digestion, thiol groups of cysteine residues were oxidized to cysteic acid and thus blocked from further reactions. Phosphate residues were removed by β-elimination using high pH-conditions. This elimination results in a replacement of the phosphogroup with an unsaturated residue to which ethanedithiol (EDT) is added. Finally a biotin bearing linker structurally related to the ICAT linker is introduced to the free thiol group of the EDT. The further steps of processing are equal as for ICAT labeled peptides, namely biotin based purification, subsequent elution and LC-MS analysis. This approach is easily applicable for quantification using EDT-H_4 and EDT-D_4 for induction of isotope bearing tags (Fig. 2.2) for peptide modification [28].

2.3.3.4. ICPL

The purification of cysteine peptides to reduce the complexity of the peptide mixture about 10 fold is a major advantage of the ICAT method, but can also become one of its pitfalls. The reduction of complexity leads also to a reduction of protein coverage as cysteine is one of the rarest amino acids. For example approximately 66% of all *E. coli* open reading frames contain less than 5 cysteine residues [29].

Apart from cysteine thiol groups, the much more frequent amino group in lysine side chains is as well modifiable in a selective way. More than 80% of all proteins have a high lysine content and are therefore available for isotope-coded protein labels (ICPL, Fig. 2.2) using either D_4-nicotinic acid [29] or $^{13}C_6$-nicotinic acid [30] as heavy modifying reagent. Like for cysteine based quantification approaches proteins are first modified, then mixed and digested. As lysine modification impairs protein digestion with trypsin, proteins have to be digested with other

proteases such as endoproteinase Glu-C.

Although the higher protein coverage of the ICPL-technique can be very valuable, some limitations are intrinsic to this approach. Unlabeled peptide portions are not separated from the modified peptide pool and thus singulet peptide peaks in LC-MS analysis derive from highly abundant peptides as well as from unmodified ones. Additionally a multiple labeling of one peptide is likely due to the fact that lysine is a frequent amino acid. This multiple labeling in turn complicates identification of peptide pairs (Fig. 2.1-E).

2.3.4. Terminal labeling

In principle terminal labeling has one major advantage as every peptide is labeled and thus all peptides are available for quantification. Proteins labeling after digestion bears vice versa the restraint that mixing of samples is later than in all other approaches (Fig. 2.1-C). This makes this approach per se more prone to artefacts due to separated sample preparation.

2.3.4.1. Enzymatic digest in presence of $H_2{}^{18}O$

One of the first strategies used for quantitative comparisons of protein samples was the enzymatic digestion of protein samples in ^{18}O-labeled water, marking the C-terminus of the generated peptide by a mass shift of 2 Da compared to peptides digested in $H_2{}^{16}O$ [31, 32]. Although this approach is elegant and fast there are some restraints. The small mass change of the peptide mass due to incorporation of ^{18}O results in an overlay of the isotope pattern of light and heavy tagged peptides. Furthermore the 2 Da mass difference between differentially tagged peptides can be altered due to rebinding of trypsin to peptidic arginine and lysine residues after cleavage. The rebinding leads to incorporation of a second ^{18}O into the peptide [33] complicating the quantitative analysis of the sample.

2.3.4.2. Acetylation and TMAB

Acetylation of free aminogroups in a peptide is chemically simple and quantitative and thus allows easily the introduction of an isotope-coding tag into a peptide. Several protocols have been described either for specific acetylation of the N-terminus after blocking free amino groups in lysine side chains ([34], see also dNIC section) or the unspecific acetylation of all free amino groups in a peptide [35], which is less favorable for analysis.

Irrespective the specificity of the reaction, acetylation reduces the nucleophilic character of the modified aminogroup dramatically, resulting in very ineffective ionization efficiencies of the acetylated peptides (Fig. 2.3). Therefore other N-terminal modification reagents, which preserve a positive N-terminal charge, are more favorable.

One example of such a reagent is 4-trimethyl-ammoniumbutyryl (TMAB, Fig. 2.2, [14]) which also reacts with every free amino group. In 2005 Che and coworkers used this reagent without blocking lysine side-chains successfully for quantification of neuropeptides in mice [36].

2.3.4.3. dNIC

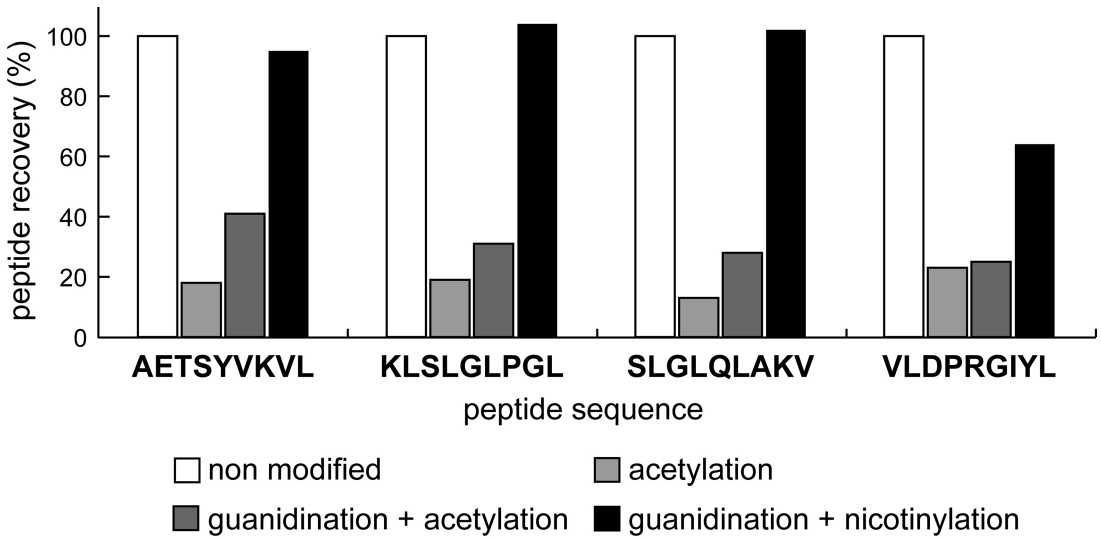

Figure 2.3.: Recoveries of different peptide derivatives. The four synthetic peptides (AETSYVKVL, KLSLGLPGL, SLGLQLAKV and VLDPRGIYL) present in equimolar concentrations were either acetylated, guanidinated and acetylated or guanidinated and nicotinylated. After complete derivatization, each peptide mixture was combined with the initial nonmodified peptide mixture (thus expecting equimolar yields of both species) and quantified by nanospray-ESI-MS analysis. This figure is adapted from Lemmel et al. 2004 [34].

Lemmel et al. 2004 introduced an approach to provide a differential N-terminal isotope coding (dNIC) selectively at the N-terminus [34]. To prevent isotope labeling of lysine side chains, these amino groups are selectively guanidinated at a pH above 10.5 using O-methylisourea [37] (for workflow see Fig. 4.1). The peptides are desalted on C_{18}-colums and afterwards the dNIC label is solely introduced at the N-terminus. This approach is advantageous in various respects. Every peptide is modified with exactly one dNIC-label, meaning, that singlet LC-MS events can not emerge from unlabeled peptides but are due to singularly existing peptides in one of the samples. The dNIC-label is very light (110 Da) and thus has not the impairments of ICAT-labels in respect of peptide sequencing. Furthermore the dNIC-label as well as the guanidination enhance MS sensitivity

[38, 39], which compensates losses during peptide labeling quantitatively (Fig. 2.3).

2.4. Conclusions

Although there is still much space for further improvements regarding e.g. quantification of very low abundant proteins or peptides and automation, quantitative comparisons of proteins using LC-MS became a well established technique during the last decade.Several different modification reagents cover a broad range of biological applications from total proteome comparisons using ICAT or ICAT-abutted techniques to reagents suited for phosphoprotein or MHC ligand quantification. Common to all these techniques is a good reproducibility and sensitivity as well as a broad dynamic range, altogether needed for accurate quantification experiments. Taken all together it can be expected that protein quantification will be used more frequently in the future, tackling multifaceted problems in various biological systems. There it will complement largely available quantitative mRNA data to a refined view of the addressed question [4].

2.5. References

[1] M. Schena, D. Shalon, R. W. Davis, and P. O. Brown. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 270(5235):467–470, 1995.

[2] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, and R. Aebersold. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.*, 17(10):994–999, 1999.

[3] M. Huber, I. Bahr, J. R. Kratzschmar, A. Becker, E. C. Muller, P. Donner, H. D. Pohlenz, M. R. Schneider, and A. Sommer. Comparison of proteomic and genomic analyses of the human breast cancer cell line T47D and the antiestrogen-resistant derivative T47D-r. *Mol. Cell Proteomics.*, 3(1):43–55, 2004.

[4] T. Ideker, V. Thorsson, J. A. Ranish, R. Christmas, J. Buhler, J. K. Eng, R. Bumgarner, D. R. Goodlett, R. Aebersold, and L. Hood. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science*, 292(5518):929–934, 2001.

[5] R. Aebersold and M. Mann. Mass spectrometry-based proteomics. *Nature*, 422(6928):198–207, 2003.

[6] J. Klose. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik.*, 26(3):231–243, 1975.

[7] P. H. O'Farrell. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.*, 250(10):4007–4021, 1975.

[8] P. James, M. Quadroni, E. Carafoli, and G. Gonnet. Protein identification by mass profile fingerprinting. *Biochem. Biophys. Res. Commun.*, 195(1): 58–64, 1993.

[9] III J. R. Yates, S. Speicher, P. R. Griffin, and T. Hunkapiller. Peptide mass maps: a highly informative approach to protein identification. *Anal. Biochem.*, 214(2):397–408, 1993.

[10] S. P. Gygi, G. L. Corthals, Y. Zhang, Y. Rochon, and R. Aebersold. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. U. S. A*, 97(17):9390–9395, 2000.

[11] M. Unlu, M. E. Morgan, and J. S. Minden. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis*, 18(11):2071–2077, 1997.

[12] Z. Yin, D. Stead, L. Selway, J. Walker, I. Riba-Garcia, T. McLnerney, S. Gaskell, S. G. Oliver, P. Cash, and A. J. Brown. Proteomic response to amino acid starvation in Candida albicans and Saccharomyces cerevisiae. *Proteomics.*, 4(8):2425–2436, 2004.

[13] J. Lill. Proteomic tools for quantitation by mass spectrometry. *Mass Spectrom. Rev.*, 22(3):182–194, 2003.

[14] R. Zhang, C. S. Sioma, R. A. Thompson, L. Xiong, and F. E. Regnier. Controlling deuterium isotope effects in comparative proteomics. *Anal. Chem.*, 74(15):3662–3669, 2002.

[15] S. A. Gerber, J. Rush, O. Stemman, M. W. Kirschner, and S. P. Gygi. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U. S. A*, 100(12):6940–6945, 2003.

[16] Y. Oda, K. Huang, F. R. Cross, D. Cowburn, and B. T. Chait. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. U. S. A*, 96(12):6591–6596, 1999.

[17] M. P. Washburn, R. Ulaszek, C. Deciu, D. M. Schieltz, and III J. R. Yates. Analysis of quantitative proteomic data generated via multidimensional protein identification technology. *Anal. Chem.*, 74(7):1650–1657, 2002.

[18] R. Stocklin, J. F. Arrighi, K. Hoang-Van, L. Vu, F. Cerini, N. Gilles, R. Genet, J. Markussen, R. E. Offord, and K. Rose. Positive and negative labeling of human proinsulin, insulin, and C-peptide with stable isotopes. New tools for in vivo pharmacokinetic and metabolic studies. *Methods Mol. Biol.*, 146:293–315, 2000.

[19] S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, and M. Mann. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell Proteomics.*, 1(5):376–386, 2002.

[20] J. S. Andersen, Y. W. Lam, A. K. Leung, S. E. Ong, C. E. Lyon, A. I. Lamond, and M. Mann. Nucleolar proteome dynamics. *Nature*, 433(7021): 77–83, 2005.

[21] J. H. Ringrose, H. D. Meiring, D. Speijer, T. E. Feltkamp, C. A. van Els, A. P. de Jong, and J. Dankert. Major histocompatibility complex class I peptide presentation after Salmonella enterica serovar typhimurium infection assessed via stable isotope tagging of the B27-presented peptide repertoire. *Infect. Immun.*, 72(9):5097–5105, 2004.

[22] M. A. Moseley. Current trends in differential expression proteomics: isotopically coded tags. *Trends Biotechnol.*, 19(10 Suppl):S10–S16, 2001.

[23] W. A. Tao and R. Aebersold. Advances in quantitative proteomics via stable isotope tagging and mass spectrometry. *Curr. Opin. Biotechnol.*, 14(1):110–118, 2003.

[24] K. C. Hansen, G. Schmitt-Ulms, R. J. Chalkley, J. Hirsch, M. A. Baldwin, and A. L. Burlingame. Mass Spectrometric Analysis of Protein Mixtures at Low Levels Using Cleavable 13C-Isotope-coded Affinity Tag and Multidimensional Chromatography. *Mol. Cell Proteomics.*, 2(5):299–314, 2003.

[25] M. Shen, L. Guo, A. Wallace, J. Fitzner, J. Eisenman, E. Jacobson, and R. S. Johnson. Isolation and Isotope Labeling of Cysteine- and Methionine-containing Tryptic Peptides: Application to the Study of Cell Surface Proteolysis. *Mol. Cell Proteomics.*, 2(5):315–324, 2003.

[26] Y. Qiu, E. A. Sousa, R. M. Hewick, and J. H. Wang. Acid-labile isotope-coded extractants: a class of reagents for quantitative mass spectrometric analysis of complex protein mixtures. *Anal. Chem.*, 74(19):4969–4979, 2002.

[27] Y. Oda, T. Nagasu, and B. T. Chait. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.*, 19(4): 379–382, 2001.

[28] M. B. Goshe, T. D. Veenstra, E. A. Panisko, T. P. Conrads, N. H. Angell, and R. D. Smith. Phosphoprotein isotope-coded affinity tags: application to the enrichment and identification of low-abundance phosphoproteins. *Anal. Chem.*, 74(3):607–616, 2002.

[29] A. Schmidt, J. Kellermann, and F. Lottspeich. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics.*, 5(1):4–15, 2005.

[30] E. O. Hochleitner, B. Kastner, T. Frohlich, A. Schmidt, R. Luhrmann, G. Arnold, and F. Lottspeich. Protein stoichiometry of a multiprotein complex, the human spliceosomal U1 small nuclear ribonucleoprotein: absolute quantification using isotope-coded tags and mass spectrometry. *J. Biol. Chem.*, 280(4):2536–2542, 2005.

[31] I. I. Stewart, T. Thomson, and D. Figeys. 18O labeling: a tool for proteomics. *Rapid Commun. Mass Spectrom.*, 15(24):2456–2465, 2001.

[32] X. Yao, A. Freas, J. Ramirez, P. A. Demirev, and C. Fenselau. Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.*, 73(13):2836–2842, 2001.

[33] K. J. Reynolds, X. Yao, and C. Fenselau. Proteolytic 18O labeling for comparative proteomics: evaluation of endoprotease Glu-C as the catalytic agent. *J. Proteome. Res.*, 1(1):27–33, 2002.

[34] C. Lemmel, S. Weik, U. Eberle, J. Dengjel, T. Kratt, H. D. Becker, H. G. Rammensee, and S. Stevanović. Differential quantitative analysis of MHC ligands by mass spectrometry using stable isotope labeling. *Nat. Biotechnol.*, 2004.

[35] F. Y. Che and L. D. Fricker. Quantitation of neuropeptides in Cpe(fat)/Cpe(fat) mice using differential isotopic tags and mass spectrometry. *Anal. Chem.*, 74(13):3190–3198, 2002.

[36] F. Y. Che, R. Biswas, and L. D. Fricker. Relative quantitation of peptides in wild-type and Cpe(fat/fat) mouse pituitary using stable isotopic tags and mass spectrometry. *J. Mass Spectrom.*, 40(2):227–237, 2005.

[37] R. L. Beardsley and J. P. Reilly. Optimization of guanidination procedures for MALDI mass mapping. *Anal. Chem.*, 74(8):1884–1890, 2002.

[38] F. L. Brancia, A. Butt, R. J. Beynon, S. J. Hubbard, S. J. Gaskell, and S. G. Oliver. A combination of chemical derivatisation and improved bioinformatic tools optimises protein identification for proteomics. *Electrophoresis*, 22(3): 552–559, 2001.

[39] T. Keough, M. P. Lacey, and R. S. Youngquist. Derivatization procedures to facilitate de novo sequencing of lysine-terminated tryptic peptides using postsource decay matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.*, 14(24):2348–2356, 2000.

3. Aims of the thesis

At the beginning of this thesis, the differential N-terminal isotope (dNIC) strategy had already been established profoundly by Claudia Lemmel (Lemmel et al., Nat Biotechnol. 2004 Apr;22(4):450-4.). However, the reliability of this strategy had not yet been tested in complex samples.

Thus, the first aim of this thesis was to test the reliability of the dNIC strategy in complex samples like MHC-peptide extracts from cell lines or primary tissue. Additionally, the mixing procedure of two samples, to be compared with each other, was to be optimized.

The second aim was the identification of patterns according to which TAP defective tumours could be made available for tumour immunologic approaches. Therefore, this optimized dNIC strategy was used for a comparison of a TAP expressing cell line with a hereof derived TAP deleted mutant cell line.

The third aim was to analyse the correlation between HLA ligand ratios and mRNA rations in order to assess the potential of gene expression based identification of tumour (associated) antigens. Thus, the differential analysis was used for a comparison of tumour and autologous normal tissue regarding HLA peptide presentation and gene expression.

Fourth, a T cell epitope, identified by the dNIC approach was to be defined in order to demonstrate the power of the dNIC approach in the identification of tumour (associated) antigens.

4. Results I: Peptide quantification using dNIC

This chapter is partly derived from a review article published in *Biotechnology and Genetic Engeneering Reviews*, 23:21-39, 2006 by:
Andreas O. Weinzierl and Stefan Stevanović.

The author of this thesis designed and performed all experiments described herein. Figure 4.4-A and -B were created by Jens Joachim.

4.1. Quantitative analysis of MHC-I ligands

MHC-I molecules are cell surface proteins presenting a non-covalently bound peptide [1]. These MHC ligands mainly derived from endogenous proteins are displayed primarily to cytotoxic T-cells for [2]. Several features of MHC ligands withdraw them from ICAT based quantification but make them perfectly suited for N-terminal modification using the dNIC-approach. The MHC-bound peptides range in size from 8-10 amino acids [3] and seldom contain cysteine residues (as can be seen from the MHC ligand database at www.syfpeithi.de). Thus all ICAT based approaches are not suited for quantitative MHC analysis. Fortunately the complexity of a MHC ligand preparation is in a magnitude which does not need reduction prior to LC-MS analysis. Ringrose et al therefore used metabolic labeling with ^{15}N-arginine to analyze differential HLA-B*2704 peptide presentation after *Salmonella enterica* infection [4]. The studied HLA-B*2704 molecule binds mostly peptides which bear in their second position an arginine as so called anchor residue. Therefore merely all isolated HLA-B*2704 ligands were differentially marked in position two.

This approach is not generally conferrable to other MHC ligand pools as most MHC motifs are not that restrictive in their anchor amino acids as HLA-B*2704. A metabolic labeling with a more abundant amino acid would impair the same general problems as described above for metabolic labeled proteins. Due to the fact that protein coverage of MHC-bound peptides per se is very low, it is in this context especially important, that virtually every peptide is quantitative analyzable. Therefore also a quantification based on lysine side chain labeling like ICPL which has no separation of labeled from unlabeled peptides would be inadequate.

All these restraints hampering metabolic and side chain labeling do not apply for the dNIC approach (Fig. 4.1) as it is independent from amino acid composition and all peptides are available for quantification (Fig. 2.1-C).

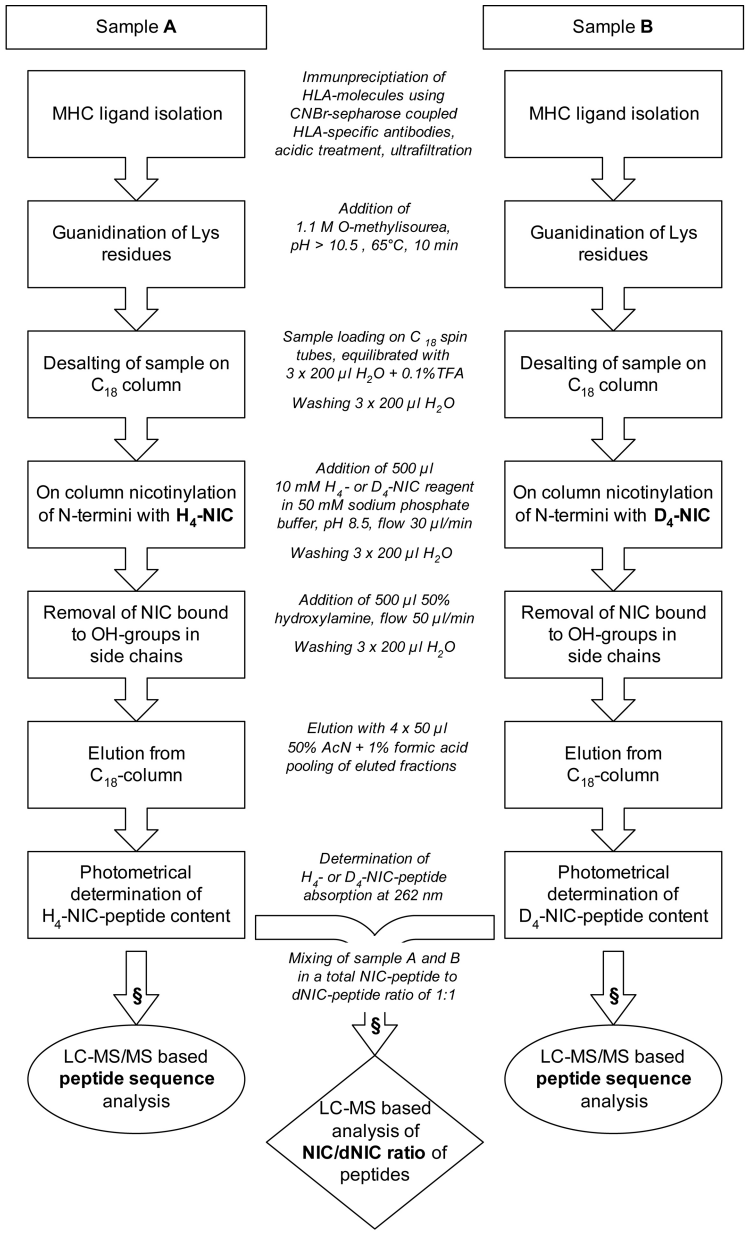

Figure 4.1.: Workflow for dNIC based quantitative comparison of MHC-I ligand levels isolated from two samples.

4.2. Kinetics of MHC-I ligand nicotinylation

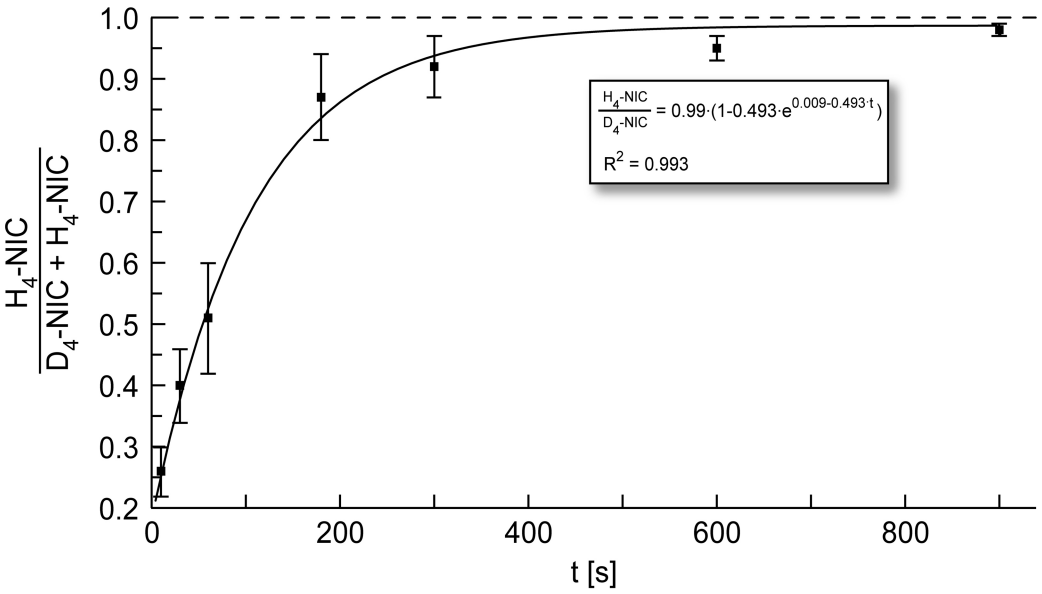

Figure 4.2.: Kinetic analysis of peptide nicotinylation. A mix of eight synthetic peptides was guanidinated and afterwards subjected to nicotinylation for 15 min in total. First peptides were pulsed with 10 mM H_4-NIC for different times indicated in the diagram. Afterwards H_4-NIC was removed and for the remaining time 10 mM D_4-NIC was added for chase. For each peptide and point in time ratios of peptides modified with light and heavy nicotinic acid were calculated. The used fitting function is indicated in the diagram.

Peptide modification at NH_2 of lysine side chains or at the N-terminus with N-nicotinoyloxy-succinimide (NIC-NHS) esters is carried out from 1 h [5] to 2.5 h [6, 7]. To optimize reaction time and NIC-NHS usage, we investigated nicotinylation kinetics in a pulse-chase experiment (Fig. 4.2). A mix of eight synthetic peptides (4 nmol each peptide) was on-column modified within 15 min in total with 10 mM light or heavy nicotinylation reagent. For pulsing the pre-guanidinated peptide mix was incubated for 10 sec to 15 min with H_4-NIC. The chase was carried out by replacing light nicotinylation solution with D_4-NIC to give a total nicotinylation time of 15 min. For kinetic analysis, the H_4-NIC to D_4-NIC ratio of modified peptides was determined by mass spectrometry. Summarizing the results, a nicotinylation efficiency of 99% was achieved for every synthetic peptide after 15 min and thus the protocol for the dNIC approach could be optimized.

4.3. Appropriate mixing of dNIC-labeled MHC-I ligands

Table 4.1.: Proof of principle for densitometry based mixing of peptides modified with the dNIC strategy. 42 nmol of a peptide mix were guanidinated and split into portions of 6, 12 and 24 nmol. These portions were separately modified with H_4-NIC (6 and 24 nmol) and D_4-NIC (12 nmol). The optical density at 262 nm was determined for each sample after modification. The 6 nmol and 12 nmol portion were mixed according to their optical density in a 1:1 ratio. The 12 and 24 nmol portions were mixed in a 0.8:1 ratio respectively. Actual peptide ratios were determined by MS.

amount of modified peptide	6nmol	12nmol		24nmol	12nmol	
	Intensity			**Intensity**		
Peptide sequence	**NIC**	**dNIC**	**ratio**	**NIC**	**dNIC**	**ratio**
PGSYTYEWNFRKDVN	386	393	0.98	612	755	0.81
DVETQFNQYK	547	506	1.08	966	1217	0.79
KLKEFIPKV	495	460	1.08	514	629	0.82
GLDEVKSSL	696	607	1.15	1021	1251	0.82
VVQLTLAFR	1683	1633	1.03	1777	2462	0.72
ILNSWNISK	799	703	1.14	1286	1444	0.89
average			**1.08±0.06**			**0.81±0.05**
expected			**1.00**			**0.80**

Adequate mixing of the samples prior to quantification analysis is one of the most crucial steps for every quantitative application. Regarding adequate mixing, metabolic labeling is generally advantageous as with no other isotope labeling method different samples can be mixed that early in the workflow (Fig. 2.1-A). For quantitative analysis total cell preparations can be normalized to equal cell number or total protein content, directly mixed and processed together. This early mixing procedure reduces any artificial protein quantification differences resulting from different handling of the samples. Principally terminal modification implies a later mixing than all other modification techniques, but this does not apply to MHC-I ligand samples to full extent. MHC-I ligands are short peptides, therefore an enzymatic digestion step is surplus. Nevertheless accurate determination of total MHC ligands content is difficult. Before acid treatment of immunoprecipitated MHC the total amount of MHC can be determined and hence the quantity of the thereof isolated MHC ligands can only be roughly deduced.

Due to the sensitivity of LC-MS based dNIC quantification experiments a more accurate determination of total peptide content of a sample prior to mixing is required. As peptide nicotinylation introduces a strongly UV-absorbing nicotinic

acid into every peptide the total amount of peptide in a sample can be determined by UV-absorption at 262 nm. To prove that this densitometric based mixing is in principal appropriate for LC-MS quantification experiments pools of synthetic peptides were differentially labeled and mixed in fixed ratios according to their OD_{262nm}. Afterwards the actual peptide ratios were determined by MS. In detail 6, 12 and 24 nmol of six synthetic peptides were modified either with light or heavy nicotinic acid. After elution from the C_{18} columns OD_{262nm} was determined for each sample, whereas an equally treated sample without peptide served as blank. According to their optical densities, the samples were mixed in a ratio of 1:1 respectively 0.8:1. The actual peptide ratios were determined by mass spectrometric analysis (Tab. 4.1). The results from this experiment demonstrate that UV-absorption based peptide mixing is accurate for nicotinylated peptides as the relative error was as low as 8% which is in a range similar to that obtained with other isotope based quantification strategies [8].

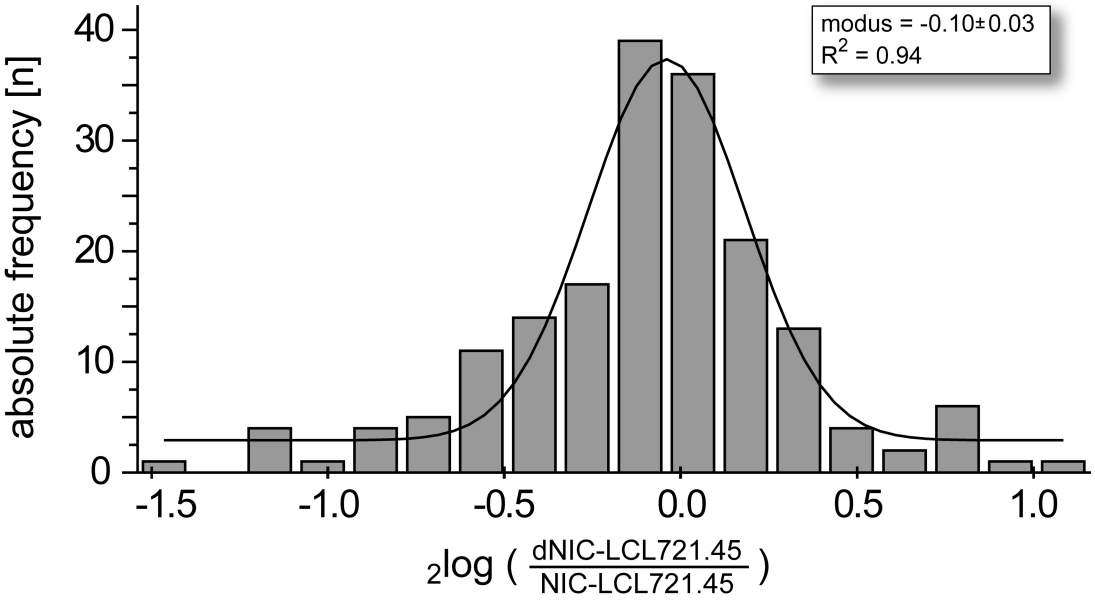

Figure 4.3.: Densitometry based mixing of two complex peptide samples after dNIC modification. MHC-I ligands were isolated from two preparations of LCL721.45, modified with D_4-NIC respectively H_4-NIC and mixed according to their OD_{262nm} in a 1:1 ratio. A frequency count (bin size 0.157) of the logarithmized ratios of dNIC to NIC modifiedligands was performed and fitted using a Gaussian curve, modus and R^2 are indicated within the figure.

This mixing method was also successfully applied to more complex samples. Total MHC ligands were isolated independently from two batches of the cell line LCL721.45 and modified with D_4-NIC respectively H_4-NIC (Fig. 4.3). Here most MHC ligands seem to be equally presented on both batches of the LCL721.45 cell line. Assuming a Gaussian distribution of the logarithmized ratio of dNIC to

NIC modified peptides (R^2=0.94), the ratios can be further normalized setting the modus of the curve at zero.

4.4. Automatization of peptide quantification using dNIC-labels

The dNIC strategy labels every peptide and introduces a mass shift from light to heavy modification reagent of 4 Da, independent from the peptide sequence. This constant mass difference and the isotopic effect of the peptides modified with the deuterium bearing nicotinic acid can be used for reliable automated analysis of peptide pair ratios. This makes the dNIC quantification approach well suited for high-throughput experiments. Figure 4.4 shows an example how an automated quantification can be approached. Individual measured values of one peptide including its isotope pattern eluting from the HPLC (Fig. 4.4-A) is integrated using a two-dimensional model consisting of an elution profile and a theroretical isotope pattern (Fig. 4.4-B). Finally the room integral of such features is calculated. Pairs of feature ((Fig. 4.4-C) are easily allocatable due to the mass difference and the isotopic shift in retetnion time. Such automated analysis have been integrated in the OpenMS software (http://open-ms.sourceforge.net).

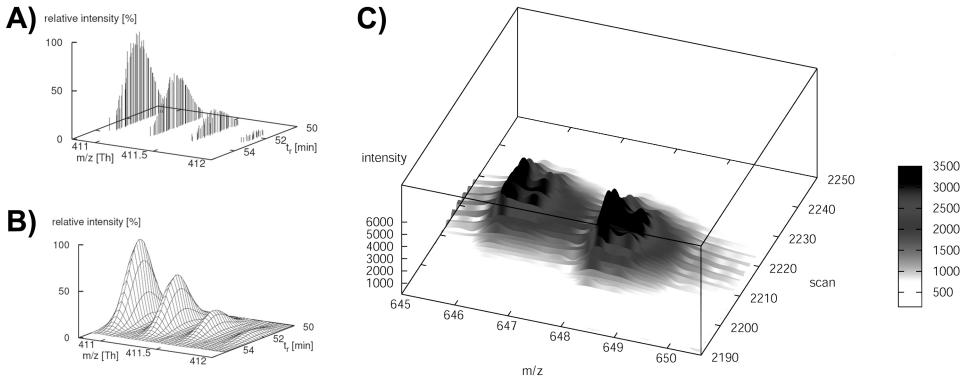

Figure 4.4.: Exemplaric approach of an automated peptide quantification analysis in an LC-MS experiment. A, The individual measured m/z values and intensities of one peptide which eluting from the HPLC are depicted. B, Two-dimensional model of an elution profile combined with a theroretical isotope pattern. The measured values of the peptide depicted in A is depicted. C, Three-dimensional image using data of a modified peptide pair with a m/z = 2 mass shift. The dNIC peptide is eluting earlier due to the isotopic effect of deuterium.

4.5. Facilitated *de novo* sequencing by dNIC-labeling

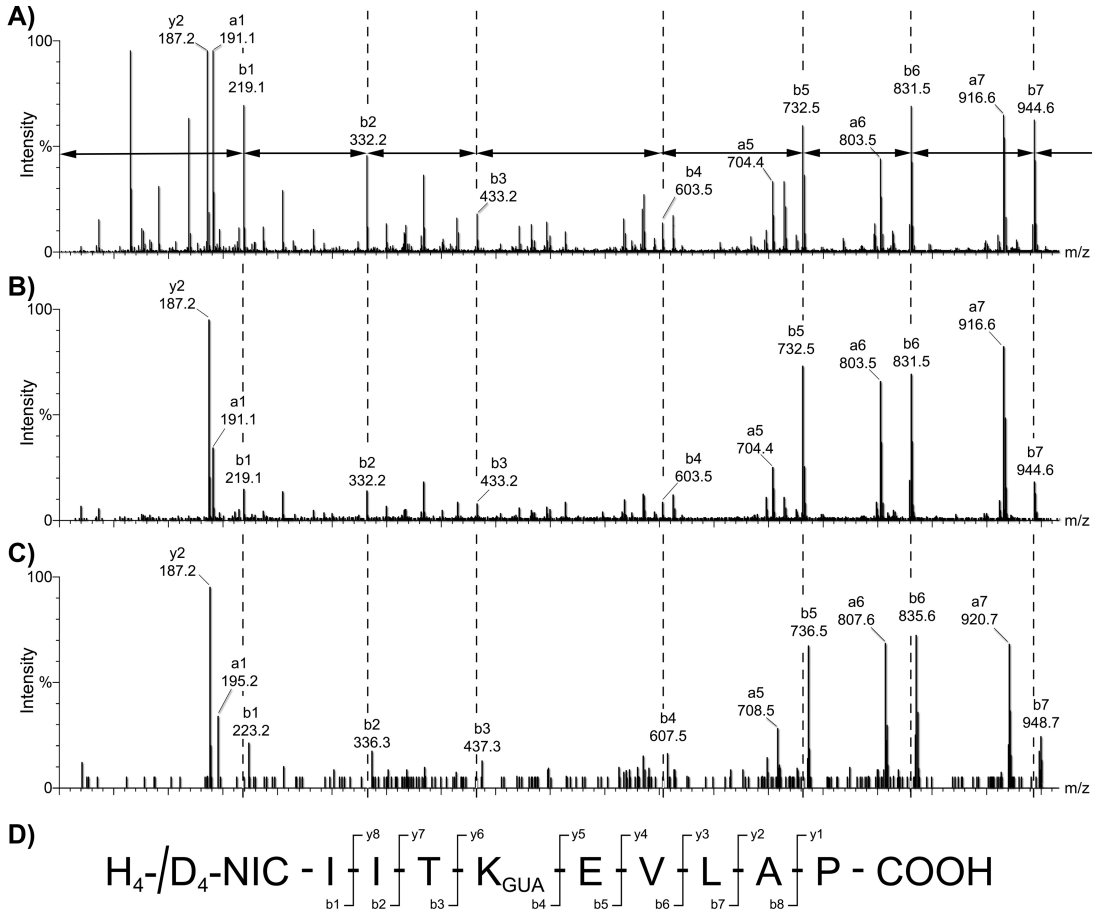

Figure 4.5.: Faciliated *de novo* sequencing of the modified peptide IITKEVLAP using the dNIC approach. A, MS/MS fragmentation spectrum of synthetic IITKEVLAP modified with D_4-NIC after guanidination. B, MS/MS fragmentation spectrum of IITKEVLAP isolated from LCL721.174 also modified with D_4-NIC after guanidination. C, MS/MS fragmentation spectrum of guanidinated and H_4-NIC modified IITKEVLAP, isolated from LCL721.174. D, Nomenclature of MS/MS fragments derived from modified IITKEVLAP.

Apart from the fact that N-terminal modification with heavy and light nicotinic acid enables quantitative comparison of two samples, it also facilitates *de novo* peptide sequencing. Identical peptides modified with light and heavy dNIC-reagent result in identical MS/MS-spectra except that every fragment containing the N-terminus differs in mass by 4 Da. Taking advantage of that fact, we could *de novo* sequence the MHC ligand IITKEVLAP, which origins only in an expressed sequence tag (EST) corresponding to a yet unknown protein (Fig. 4.5).

The modified synthetic peptide (Fig. 4.5-a) displays an identical fragmentation pattern compared to the MHC ligands isolated from the LCL721.174 cell line modified with light or heavy modification reagent (Fig. 4.5-b/c). Another feature of nicotinylation facilitating sequencing is the fact that the N-terminal b-ion (Fig. 4.5-d) becomes observable in MS/MS fragmentation spectra. Normally N-terminal b1-ions (Fig. 4.5-d) are not generated during peptide fragmentation as b-ion generation requires cyclic transition states [11]. Instead b2-ions are generally well observable. As nicotinylation introduces a new peptide bond at the N-terminus the b2-ions generated from nicotinylated peptides contain only nicotinic acid and the N-terminal amino acid and thus make the N-terminal amino acid easily identifiable.

4.6. Acknowledgements

We thank Claudia Lemmel who invented the dNIC approach for invaluable help during optimization of the method and Lynne Yakes for expert proof reading. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 685 and SFBTR 19).

4.7. References

[1] M. A. Saper, P. J. Bjorkman, and D. C. Wiley. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. *J. Mol. Biol.*, 219(2):277–319, 1991.

[2] R. M. Zinkernagel and P. C. Doherty. Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature*, 248(450):701–702, 1974.

[3] H. G. Rammensee, K. Falk, and O. Rotzschke. Peptides naturally presented by MHC class I molecules. *Annu. Rev. Immunol.*, 11:213–244, 1993.

[4] J. H. Ringrose, H. D. Meiring, D. Speijer, T. E. Feltkamp, C. A. van Els, A. P. de Jong, and J. Dankert. Major histocompatibility complex class I peptide presentation after Salmonella enterica serovar typhimurium infection assessed via stable isotope tagging of the B27-presented peptide repertoire. *Infect. Immun.*, 72(9):5097–5105, 2004.

[5] E. O. Hochleitner, B. Kastner, T. Frohlich, A. Schmidt, R. Luhrmann, G. Arnold, and F. Lottspeich. Protein stoichiometry of a multiprotein complex, the human spliceosomal U1 small nuclear ribonucleoprotein: absolute quantification using isotope-coded tags and mass spectrometry. *J. Biol. Chem.*, 280(4):2536–2542, 2005.

[6] M. Münchbach, M. Quadroni, G. Miotto, and P. James. Quantitation and facilitated de novo sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation-directing moiety. *Anal. Chem.*, 72(17): 4047–4057, 2000.

[7] A. Schmidt, J. Kellermann, and F. Lottspeich. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics.*, 5(1):4–15, 2005.

[8] S. P. Gygi, Y. Rochon, B. R. Franza, and R. Aebersold. Correlation between protein and mRNA abundance in yeast. *Mol. Cell Biol.*, 19(3):1720–1730, 1999.

[9] B. H. Koller, D. E. Geraghty, R. DeMars, L. Duvick, S. S. Rich, and H. T. Orr. Chromosomal organization of the human major histocompatibility complex class I gene family. *J. Exp. Med.*, 169(2):469–480, 1989.

[10] M. L. Wei and P. Cresswell. HLA-A2 molecules in an antigen-processing mutant cell contain signal sequence-derived peptides. *Nature*, 356(6368): 443–446, 1992.

[11] T. Yalcin, C. Khouw, I. G. Csizmadia, M. R. Peterson, and A. G. Harrison. Why Are B Ions Stable Species in Peptide Spectra? *J. Am. Soc. Mass Spectrom.*, 6(12):1165–1174, 1995.

5. Results II: Features of TAP-independent MHC class I ligands revealed by quantitative mass spectrometry

This chapter has been submitted for publication by:
Andreas O. Weinzierl, Despina Rudolf, Nina Hillen, Stefan Tenzer, Peter van Endert, Hansjörg Schild, Hans-Georg Rammensee and Stefan Stevanović.

The author of this thesis designed and performed all HLA quantification and FACS experiments as well as gene expression and bioinformatical analyses.

5.1. Summary

The TAP transporter is responsible for transferring cytosolic peptides into the ER where they can be loaded onto MHC molecules. Deletion of TAP results in a drastic reduction of MHC surface expression and alters the presented peptide pattern. This key molecule in antigen processing is tackled by several viruses and in some tumours rendering the altered cells less vulnerable to T cell based immune surveillance. Using the TAP deficient cell line LCL721.174 and its TAP expressing progenitor cell line LCL721.45, we have identified and quantified more than 160 HLA ligands, 50 of which were presented TAP-independently. Peptides which were predominantly presented on the TAP deficient LCL721.174 cell line had a decreased MHC binding affinity according to their SYFPEITHI and BIMAS score. About half of the identified TAP-independently presented peptides were not derived from signal sequences and may partly be generated by the proteasome. Furthermore, we have excluded the possibility that different HLA presentation ratios were due to varying expression of the respective protein or due to changes in the antigen loading complex. Features of peptides presented independently of TAP as well as proteasomal contribution to their generation provides an insight into basic immunological mechanisms.

5.2. Introduction

Peptides derived from endogenously expressed proteins are displayed on the cell surface by MHC-I molecules to cytotoxic T Lymphocytes (CTLs), enabling the

CTLs to distinguish normal cells from infected cells.

Proteins are degraded within cells by the proteasome – generating 3-22 aa peptides [1], which can be further trimmed in the cytosol by proteases like the tripeptidyl peptidase II (TPP2) [2]. Such generated peptides – with an optimal length of 9-16 aa [3] – are transported into the ER by the TAP complex. Before the peptides are loaded onto MHC-I molecules they can be further trimmed by the ER aminopeptidase associated with antigen processing (ERAAP) [4]. The resulting peptides with a length of 8-11 aa [5] are loaded onto MHC-I molecules by a complex consisting of tapasin, calreticulin and ERp57. Finally, the MHC:peptide complexes are exported via the Golgi network onto the cell surface.

Apart from this major pathway, it has been described [6, 7] that TAP1/2 deficient cells display signal sequence derived peptides which have been liberated in the ER by the signal peptidase (SP) complex. Both reports found 4, respectively 3, HLA-A*02-bound peptides derived mainly from the signal peptide of the lysosomal thiol reductase IFI30. Up to now, about 10 human TAP-independent MHC-I ligands are known [8], mostly restricted to HLA-A*02. This is due to the fact that the HLA-A*02 can bind signal sequence derived peptides with a higher affinity than most other HLA types because its HLA motif has a high similarity to the protein signal sequence pattern.

The use of TAP blockade or deletion as a means of evading CTL immune surveillance has been described for various viruses such as HSV or HCMV [9] or in the case of several tumours [10]. In tumours, TAP deletion correlates with poor prognosis and tumorigenesis [11].

5.3. Materials and Methods

5.3.1. Elution of HLA presented peptides

HLA presented peptides were obtained by immune precipitation of HLA molecules from LCL721.45 and LCL721.174 cell lines [13] as described elsewhere in detail [12] using the pan-HLA antibody W6/32. $1 \cdot 10^{10}$ LCL721.45 cells respectively $5 \cdot 10^{10}$ LCL721.174 [14] cells were used for the isolation of HLA peptides. Elution experiments were performed twice.

5.3.2. Peptide modification and analysis

Modification of peptides was carried out as previously described [15]. LCL721.174 and LCL721.45-derived peptides were modified with heavy (4 hydrogens replaced by deuterium), respectively light (all hydrogen), dNIC-NHS solution. Using inverted labels, the reproducibility of the quantification was assessed to be on average about 90% (data not shown). For quantification, HLA peptides isolated from LCL721.45 and LCL721.174 were mixed in a total peptide ratio of 1:1 and

recorded in a single LCMS experiment without fragmentation using a Q-TOF-II (Micromass, Manchester, UK). For sequence analysis each HLA peptide sample was analyzed separately in individual LC-MS/MS experiments. Interpretation of MS/MS fragmentation spectra was carried as described [15].

5.3.3. Gene expression analysis

RNA isolation and gene expression analysis were carried out using Affymetrix Human Genome U133 Plus 2.0 oligonucleotide microarrays (Affymetrix, Santa Clara, CA, USA). Data were analyzed with the GeneChip Operating Software (GCOS, Affymetrix, Santa Clara, CA, USA). Pairwise comparisons between tumor and autologous normal kidney were calculated using the respective normal array as baseline. Microarray data are available from the Gene Expression Omnibus repository (www.ncbi.nih.gov/geo, GSE9437).

5.3.4. Subfractionation of cellular extracts and tryptic digest

Differential detergent fractionation [16] and centrifugal fractionation [17] were performed as previously described. All fractions were precipitated using the ProteoExtract Protein Precipitation Kit (Merck, Darmstadt, Germany) and solubilized in 25 mM ammonium bicarbonate containing 0.1 % RapiGest (Waters, Eschborn, Germany). Proteins were reduced by adding 5 mM DTT and free cysteines alkylated with iodoacetamide (Sigma, Taufkirchen, Germany). 0.2 μg porcine sequencing grade trypsin (Promega, Mannheim, Germany) was added and the samples were incubated overnight at 37°C. After digestion, RapiGest was hydrolyzed by adding 10 mM HCl and the resulting precipitate was removed by centrifugation. The supernatant was analyzed by LCMS/MS.

5.3.5. LCMS/MS analysis of tryptic digests

Capillary liquid chromatography of tryptic peptides was performed with a Waters NanoAcquity UPLC system online coupled to a Waters Q-TOF Premier API system as described [18]. The continuum LCMSE data were processed and searched using the IDENTITYE-Algorithm of ProteinLynx Global Server (PLGS, v2.3). Protein identifications were assigned by searching the UniProtKB/Swiss-Prot Protein Knowledgebase Release 52.3 for human proteins (16150 entries). Peptide identifications were restricted to tryptic peptides with no more than one missed cleavage. For valid protein identification, the following criteria had to be met: at least 2 peptides detected together with at least 7 fragments. All reported peptide identifications provided by the IDENTITYE-algorithm were correct with >95% probability. The false positive rate for protein identification was set to 1% based on search of a reversed database.

5.3.6. Proteasomal processing

Purification of 20S proteasomes, *in vitro* degradation of precursor peptides, separation, and analysis of cleavage products were performed as described [19]. 10 nmol of the peptides (HPLC purity >95%) were incubated for 6 h with 2 μg immunoproteasome or for 6 h with 2 μg of constitutive proteasome, respectively, in digestion buffer (20 mM Tris-HCl; pH 7.6; 10 mM NaCl; 10 mM KCl; 2 mM $MgCl_2$; 0.5 mM DTT) and the reaction stopped by adding formic acid to a final concentration of 0.5% and freezing the reaction mixture at -80^{o}C. Fragments were analyzed as described [18].

5.3.7. Cell culture and protease inhibition assay

LCL721.45 and LCL721.174 cells were cultivated in RPMI 1640 (C.C.Pro, Neustadt, Germany) medium containing 10% FCS (Pan, Aidenbach, Germany). For the protease inhibition assay, $2{\cdot}10^5$ LCL721.174 cells were incubated at 37°C in 200 μl cell culture medium containing 20 μM Ala-Ala-Phe-chloromethylketone (AAF-CMK, Bachem, Weil am Rhein, Germany), 200 μM Butabindide (Tocris Bioscience, Bristol, UK), 50 μM Lactacystin (biomol, Hamburg, Germany), 2 μM Epoxomycin (biomol, Hamburg, Germany), 5 μg/ml Chloroquine (Sigma-Aldrich, Taufkirchen, Germany) or 7.5 mM NH_4Cl. After 6 h HLA surface expression was analyzed by flow cytometry as described below.

5.3.8. FACS analysis

After washing, cells were incubated for 30 min on ice with the primary antibody W6/32 (HLA-A, -B, C), B1.23.2 (HLA-B, C) or BB7.2 (HLA-A). Subsequently, cells were washed three times and incubated with an FITC-conjugated goat anti-mouse antibody (Jackson ImmunoResearch Laboratories, Newmarket, Suffolk, UK) for 30 min on ice. Cells were analyzed by flow cytometry on a FACSCalibur cytometer (BD Biosciences, Heidelberg, Germany). Dead cells were stained with 7-AAD (BD Biosciences, Heidelberg, Germany) and excluded from the analysis.

5.4. Results and Discussion

5.4.1. Validation of the experimental setting

In this report we quantitatively analyze the MHC:peptide repertoire of the TAP1/2-deficient cell line LCL721.174 in comparison to its TAP expressing progenitor cell line LCL721.45 (5.1-A). To exclude influences on MHC:peptide levels caused by alterations in the respective mRNA and protein levels, both cell lines were additionally compared on the protein and mRNA level.

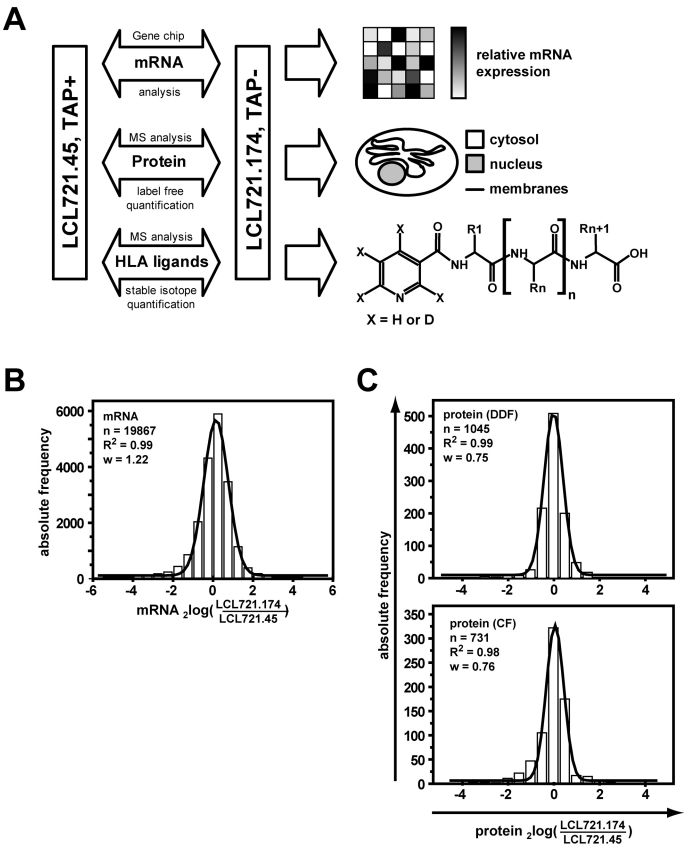

Figure 5.1.: LCL721.45 and LCL721.174 differ only slightly in protein and mRNA expression. A, Schematic overview of the quantitative comparison of LCL721.174 and LCL721.45 on the mRNA, protein and HLA ligand level. B and C, Relative gene and protein expression levels were evaluated in a histogram analysis. The full width at half maximum (w) was calculated assuming a Gaussian distribution. DDF: differential detergent fractionation, CF: centrifugal fractionation.

mRNA levels of LCL721.174 and LCL721.45 were compared in a gene chip experiment. Figure 5.1-B shows that 90% of all genes differed at the most by a factor of two in their expression. A similar analysis was performed on the protein level, whereby cells were fractionated into cytosolic, nuclear and membrane bound proteins. In order to minimize methodical influences, two independent fractionation methods were used: differential detergent fractionation (DDF) and centrifugal fractionation (CF). Tryptic digests of the different fractions were analyzed in five repetitive HPLC-MS/MS experiments. The identification of a specific protein was judged to be valid if it was identified in at least three out of five HPLC-MS/MS experiments.

Table 5.1.: A quantitative analysis of protein and mRNA levels demonstrates partial deletion of chromosome 6p21.3. The analysis comprised constitutive proteasomal subunits (PSMB1, 2 and 5), immunoproteasomal subunits (LMP7, 2 and MECL1), subunits of the immunoproteasomal regulator PA28 (PSME1-3), proteins of the HLA peptide loading complex (TAP1 and 2, tapasin, calreticulin, calnexin and PDIA3), HLA class I and II (HLA-A, -B, -C, -DR, -DP and -DQ), the endoplasmic reticulum aminopeptidase associated with antigen processing (ERAAP) and proteins of the SEC61 translocon which can be involved in peptide export from the ER. Genes encoded in the deleted region of chromosome 6p21.3 in LCL721.174 cells are highlighted in bold.

Gene Symbol	Gene Title	$_2log$ (LCL721.174 vs. LCL721.45) protein	mRNA
PSMB1	proteasome subunit beta 1	-0.29	n.d.
PSMB2	proteasome subunit beta 2	-0.54	n.d.
PSMB5	proteasome subunit beta 5	LCL721.174 only	0.2
LMP7	**proteasome subunit beta 8**	**721.45 only**	**721.45 only**
LMP2	**proteasome subunit beta 9**	**721.45 only**	**721.45 only**
MECL1	proteasome subunit beta 10	721.45 only	-1.4
PSME1	proteasome activator subunit 1	-0.82	-0.7
PSME2	proteasome activator subunit 2	-0.56	0.3
PSME3	proteasome activator subunit 3	721.45 only	0.9
CALR	calreticulin	n.a.	0.1
CANX	calnexin	n.a.	-1.0
TAP1	**transporter 1, ATP-binding cass.**	**721.45 only**	**721.45 only**
TAP2	**transporter 2, ATP-binding cass.**	**721.45 only**	**721.45 only**
TAPBP	Tapasin	721.45 only	-1.7
HLA-A	major histocomp. complex I A	-2.78	-0.8
HLA-B	major histocomp. complex I B	n.s.d.	-1.1
HLA-C	major histocomp. complex I C	-2.04	-1.0
HLA-DP	**major histocomp. complex II DP**	**721.45 only**	**721.45 only**
HLA-DQ	**major histocomp. complex II DQ**	**721.45 only**	**721.45 only**
HLA-DR	**major histocomp. complex II DR**	**721.45 only**	**721.45 only**
PDIA3	protein disulfide isomerase A3	n.a.	-0.8
SEC61A1	Sec61 alpha 1 subunit	721.45 only	-0.2
SEC61A2	Sec61 alpha 1 subunit	n.d.	-0.1
SEC61B	Sec61 alpha 1 subunit	721.45 only	-0.3
SEC61G	Sec61 alpha 1 subunit	n.d.	-0.9
ERAP1	ER aminopep. assoc. w. antigen process.	n.d.	0.7

n.d. not detected, n.s.d. no significant detection ($n < 3$) , n.a. no definite ratio available

This proteome analysis revealed an even higher similarity between LCL721.174 and LCL721.45 compared to the gene chip experiment. Figure 5.1-C shows that over 95% of all proteins for which an expression ratio was available differed between LCL721.174 and LCL721.45 at the most by a factor of two.

Table 5.1 shows an overview of mRNA and protein ratios of the key molecules involved in antigen processing, MHC-peptide loading and antigen presentation. During its generation, the LCL721.174 cell line has lost part of chromosome

6q21.3, including the genes of TAP1 and 2, LMP2 and 7 as well as MHC class II. This was detectable both on mRNA and on the protein level. Apart from these differences we could not observe any major changes in the antigen presentation machinery.

5.4.2. Analysis of HLA ligands by mass spectrometry

Due to TAP deletion, the total HLA surface expression was reduced by a factor of five in LCL721.174 (Fig. 5.2-A, left panel). HLA-A*0201 surface expression was sustained in LCL721.174 to some extent (Fig. 5.2-A, middle) due to the presence of signal peptide fragments in the ER. The expression of HLA-B*5101 expression was reduced by about forty fold (Fig. 5.2-A, right panel).

For quantification experiments, 15 nmol of HLA:peptides were isolated from LCL721.174 and LCL721.45 and each modified with a stable isotope label (Fig. 5.1-A). In total, 48 peptide sequences which were presented on the TAP1/2 deficient LCL721.174 cell line could be identified (Tab. 5.2). 25 of these peptides were also presented on LCL721.45; for 2 peptides no quantification was possible. 116 peptides specifically presented on LCL721.45 were identified; quantification was possible for 105 peptides (Supplementary Tab. A.6). The number of identified peptides, their presentation ratio and their HLA assignment is summarized in Figure 5.2-B. 91% of the peptides which were over-presented on LCL721.174 by a factor greater than ten were presented on HLA-A*0201 and were predicted to be derived from signal sequences (SignalP, http://www.cbs.dtu.dk/services/SignalP). Interestingly, five HLA-A*0201 presented peptides were not derived from signal sequences. Furthermore, both HLA-B*5101-bound peptides in this group were actually derived from signal sequences; however, the C-terminus was most probably not generated by the SP complex.

This observation that TAP deletion does not restrict the HLA-peptide repertoire solely to signal peptide derived, SP cleaved peptides is even more obvious for peptides presented both on LCL721.174 and LCL721.45 cells. About 50% of such peptides were presented on HLA-B*5101 and only one-third was derived from within in the signal sequence.

Table 5.2.: 52 peptides isolated from the TAP-deficient cell line LCL721.174 were sequenced by mass spectrometry. Signal sequence cleavage sites were predicted using SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/); peptides located within the putative signal sequence are highlighted in bold. Signal sequence-derived peptides which required additional trimming were marked with an asterisk. Predominantly TAP-independently presented peptides are found above the dashed line.

Sequence	Gene Symbol	Gene Title	position	signal seq.	HLA	HLA present. 721.174 vs. 721.45
LLSAEPVPA	**CD79B**	**CD79B antigen**	**20-28**	**28 | 29**	**A*0201**	**556.8**
LLGPRLVLA	**TMP21**	**transmembrane trafficking protein**	**22-31**	**31 | 32**	**A*0201**	**253.2**
SLWGQPAEA	**COL4A5**	**collagen, type IV, alpha 5**	**18-26**	**26 | 27**	**A*0201**	**124.4**
VLAPRVLRA	**RCN1**	**reticulocalbin 1**	**21-29**	**29 | 30**	**A*0201**	**120.3**
ALVVQVAEA	**HEXB**	**hexosaminidase B**	**34-42**	**42 | 43**	**A*0201**	**116.1**
HGVFLPLV	**KIAA0247**	**KIAA0247**	**21-28***	**39 | 40**	**B*5101**	**92.3**
LLAAWTARA	**APP**	**amyloid beta A4 precursor protein**	**9-17**	**17 | 18**	**A*0201**	**91.9**
VLLKARLVPA	**KIAA1946**	**KIAA1946**	**19-28**	**28 | 29**	**A*0201**	**47.7**
KMDASLGNLFA	FAM3C	family with seq simil. 3, member C	30-40	24 | 25	A*0201	36.4
LLFSHVDHVIA	**SLC8A1**	**solute carrier family 8, member 1**	**25-35**	**35 | 36**	**A*0201**	**28.1**
FLGPWPAAS	**LRPAP1**	**LDL rec.-rel. prot. assoc. prot. 1**	**22-30***	**32 | 33**	**A*0201**	**23.7**
SLYALHVKA	**VKORC1**	**vit. K epoxide reduc. complex**	**23-31**	**31 | 32**	**A*0201**	**23.1**
LLLSAEPVPA	**CD79B**	**CD79B antigen**	**19-28**	**28 | 29**	**A*0201**	**22.2**
AMAPPSHLLL	PELP1	Pro-Glu-Leu-rich protein 1	473-482	21 | 22	A*0201	18.3
MAPLALHLL	**IL4I1**	**interleukin 4 induced 1**	**1-9***	**21 | 22**	**B*5101**	**18.0**
FLLGPRLVLA	**TMP21**	**transmembrane trafficking prot.**	**22-31**	**31 | 32**	**A*0201**	**17.9**
LLLDVPTAAV	**IFI30**	**interferon, gamma-induc. prot. 30**	**15-24***	**26 | 27**	**A*0201**	**17.8**
LLLDVPTAAVQA	**IFI30**	**interferon, gamma-induc. prot. 30**	**15-26**	**26 | 27**	**A*0201**	**14.8**
LLLDVPTAA	**IFI30**	**interferon, gamma-induc. prot. 30**	**15-23***	**26 | 27**	**A*0201**	**14.7**
LLDVPTAAV	**IFI30**	**interferon, gamma-induc. prot. 30**	**16-24***	**26 | 27**	**A*0201**	**14.4**
VLFRGGPRGLLAVA	SSR1	signal sequence receptor, alpha	19-32	20 | 21	A*0201	12.8
ALLSSLNDF	NIF3L1	NIF3 NGG1 interacting factor 3-like 1	5-13	no	A*0201	12.8
IITKEVLAP	EST sequence	EST sequence	frame 3	n.d.	A*0201	12.8
QLQEGKNVIGL	TAGLN2	transgelin 2	166-176	no	A*0201	7.9
ILAPAGSLPKI	TERE1	transitional epithelia response protein	328-338	no	A*0201	6.2
MASRWGPLIG	**Cab45**	**calc. bind. prot. Cab45 prec.**	**8-17***	**36 | 37**	**B*5101**	**4.5**
AVLALVLAPAGA	**NRP1**	**neuropilin 1**	**10-21**	**21 | 22**	**A*0201**	**4.3**
LAPRVLRA	**RCN1**	**reticulocalbin 1**	**22-29**	**29 | 30**	**A*0201**	**4.2**
AALLDVRSVP	GDF5	growth differentiation factor 5	275-284	27 | 28	A*0201	3.8
SLPKKLALL	HSPC023	HSPC023 protein	72-80	no	A*0201	3.2
KAPVTKVAA	PDLIM1	PDZ and LIM domain 1	241-249	no	B*5101	1.7
NPLPSKETI	TMSB4X	thymosin, beta 4, X-linked	27-35	no	B*5101	1.2
MFPLVKSAL	COX7B	cytochrome c oxidase subunit VIIb	1-9	no	B*5101	0.9
MAPRTLVL	**HLA-A**	**m. histocomp. complex, I-A**	**4-11***	**24 | 25**	**B*5101**	**0.5**
TLLGHEFVL	CDC27	cell division cycle 27	605-613	no	A*0201	0.4
LPHVPLGVI	AUP1	ancient ubiquitous protein 1	374-382	37 | 38	B*5101	0.3
YLTAEILEL	HIST1H2AJ	histone 1, H2aj	58-66	no	A*0201	0.3
LPREILNLI	LOC116064	hypothetical protein LOC116064	259-267	no	B*5101	0.3
LLDRFLATV	CCNI	cyclin I	72-80	no	A*0201	0.2
GSHSMRYF	HLA-A/B/C/G	m. histocomp. complex I-A,-B,-C,-G	25-32	24 | 25	B*5101	0.2
HLINYIIFL	TMEM41B	transmembrane protein 41B	240-248	no	A*0201	0.2
YVPRAILV	TUBB3	tubulin, beta 3	59-66	no	B*5101	0.1
TLAEIAKVEL	NONO	non-POU domain cont., octamer-bind.	120-129	no	A*0201	0.1
DALDVANKIGII	RPL23A	ribosomal protein L23a	145-156	no	B*5101	0.1
RIIEETLAL	ARPC2	act. rel. prot. 2/3 complex.	9-17	no	A*0201	0.1
DGLVVLKI	EIF3S3	eukar. transl. initiat. fact. 3.	42-49	no	B*5101	0.1
MAPRTLLL	**HLA-A/B/C**	**m. histocomp. complex, I-A,-B,-C**	**4-11***	**24 | 25**	**B*5101**	**0.1**
VMAPRTLVL	**HLA-A**	**m. histocomp. complex, I-A**	**3-11***	**24 | 25**	**A*0201**	**0.1**

n.d. not determinable, n.a. ratio not available

5.4.3. Features of TAP-independent presented HLA ligands

The bulk of TAP independently presented peptides allowed an analysis of their HLA-binding features in comparison to a TAP-dependent peptide repertoire identified from the parental LCL721.45 cell line. For both groups of peptides the SYFPEITHI score (http://www.syfpeithi.de), as well as the BIMAS score (http://www-bimas.cit.nih.gov/molbio/hla_bind/), were calculated. The SYF-

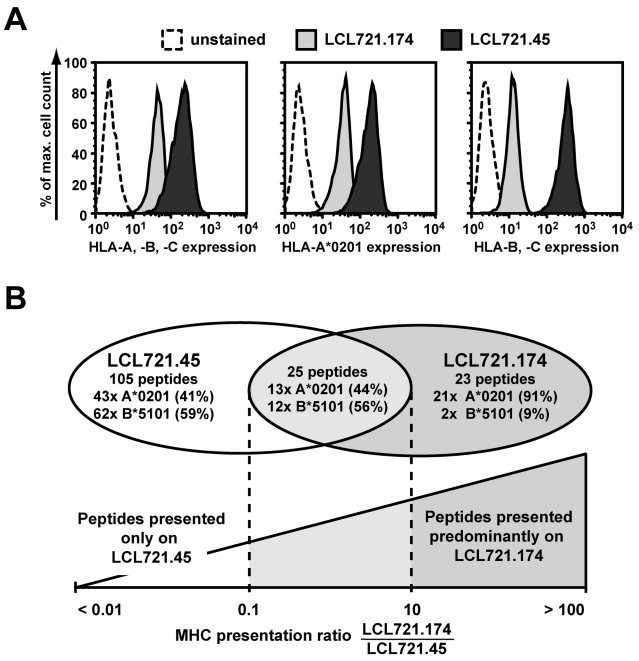

Figure 5.2.: Surface HLA and its ligands on LCL721.45 and LCL721.174. A, HLA expression of LCL721.174 and LCL721.45 was assessed by FACS analysis using the antibodies W6/32 (HLA-A, -B, -C), BB7.2 (HLA-A*0201) and B1.23.2 (HLA-B, -C). B, Overview of HLA ligands identified in LCL721.174 and LCL721.45 including their presentation ratio and HLA allele assignment.

PEITHI analysis assessed the quality with which the peptide matched its HLA motif; a SYFPEITHI score above 25 indicates high HLA affinity. With an excess of peptide in the ER (LCL721.45, TAP+), only peptides with a strong affinity managed to be exported as MHC:peptide complexes to the cell surface (Fig. 3-A, left). However, in the case of an excess of HLA molecules with limiting peptide amount (LCL721.174, TAP-), HLA molecules were less specific in their peptide selection. This resulted in a drop in the SYFPEITHI score from 27 to approximately 22. This decrease in stringency of HLA-molecules regarding their peptide choice was also reflected by the BIMAS score, which is a measure for the half-time of HLA:peptide dissociation. On average, TAP-independently presented peptides had a 15-fold decrease in their BIMAS score, also indicating a reduced binding affinity to their respective HLA molecules (Fig. 5.3-A, right).

5.4.4. Generation of non signal sequence-derived peptides

According to the SignalP prediction, 24 out of the 50 HLA ligands which were identified on LCL721.174 cells were not derived from signal sequences. To identify proteases required for generating these peptides, we analyzed the HLA-surface expression of LCL721.174 after inhibition of TPP2, the proteasome, or lysosomal

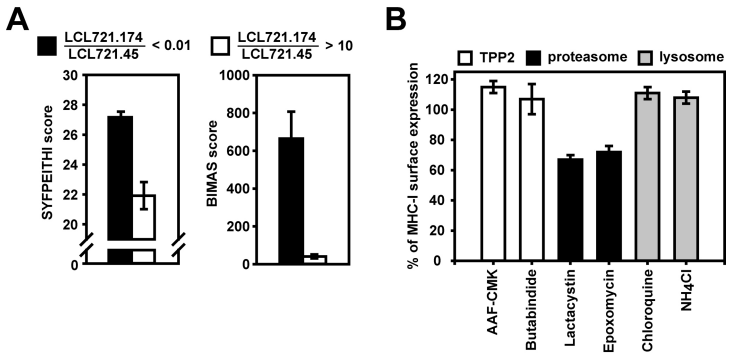

Figure 5.3.: HLA ligands of LCL721.174 bind HLA with low affinity and are dependent of proteasome activity. A, SYFPEITHI and BIMAS scores were calculated for peptides which were presented in a TAP-dependent (black) or TAP-independent (white) manner. SYFPEITHI scores above 25 indicate high HLA affinity; the BIMAS score is a measure for the HLA:peptide dissociation half-time. B, Reduction of surface HLA expression in LCL721.174 cells upon 6 h incubation with inhibitors of various proteases as revealed by FACS analysis. The color of the bars indicate the inhibited protease respectively compartment. All error bars represent the SEM.

proteases (Fig. 5.3-B). Only proteasomal inhibition resulted in a decrease of HLA surface expression, indicating that proteasome activity was necessary to process TAP-independently presented peptides which are not located in signal sequences. In order to strengthen this hypothesis, prolonged precursors of 13 non signal sequence peptides were synthesized and digested *in vitro* with purified constitutive proteasomes (expressed in LCL721.174) and immunoproteasomes (predominantly expressed in LCL721.45). The C-terminus of eight of these peptides was generatable by the proteasome; for three peptides the cleavage efficiency differed slightly between the two proteasomes (Supplementary Tab. A.7). Therefore, the proteasome could be responsible for generating the C-terminus for a part of the TAP-independently presented peptides not derived from signal sequences. Thus, precursors of HLA-ligands might be transported into the ER via the Sec translocon or a pathway that is not yet known. HLA surface expression was reduced by over 30%. Considering the fact that a residual activity remains despite proteasomal inhibition, this could correlate with the observation that approximately 50% of all identified TAP-independent peptides were not derived from signal sequences.

5.4.5. Concluding Remarks

Previous reports have identified only low monadic numbers of TAP-independently presented peptides. This low number prevented the deduction of common features possessed by such peptides except for the fact that most peptides were derived

from signal sequences and were bound to HLA-A*0201.

We could verify the finding that HLA ligands that are predominantly presented in a TAP-independent manner are mainly derived from the C-terminus of the respective signal peptide. However, this is only true for abundant peptides presented on HLA-A*02. In our analysis, more than 30% of all identified TAP-independent peptides were presented by HLA-B*5101 and about 50% were derived from mature protein chains. These peptides could not be created by the catalytic activity of the SP complex but may be generated partially by the proteasome. However, the exact mechanism by which peptides derived from within mature proteins were translocated into the ER remains unclear. Furthermore, TAP-independently presented peptides had on average a reduced HLA affinity. A detailed understanding of such features of TAP-independently presented peptides will facilitate the identification of tumour and virus-derived CTL epitopes which are specific for cells with deleted or blocked TAP transporter.

5.5. Acknowledgements

We thank Lynne Yakes for critical reading of the manuscript and Patricia Hrstić fro perfect technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft (SFB-TR19 for AOW; SFB 685 for SS), the European Union (Genomes to Vaccines, LSHB-CT-2003-503321 for NH and SS).

5.6. References

[1] A. F. Kisselev, T. N. Akopian, K. M. Woo, and A. L. Goldberg. The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J. Biol. Chem.*, 274(6):3363–3371, 1999.

[2] I. A. York, N. Bhutani, S. Zendzian, A. L. Goldberg, and K. L. Rock. Tripeptidyl peptidase II is the major peptidase needed to trim long antigenic precursors, but is not required for most MHC class I antigen presentation. *J. Immunol.*, 177(3):1434–1443, 2006.

[3] P. M. Van Endert, R. Tampe, T. H. Meyer, R. Tisch, J. F. Bach, and H. O. McDevitt. A sequential model for peptide binding and transport by the transporters associated with antigen processing. *Immunity.*, 1(6):491–500, 1994.

[4] T. Serwold, F. Gonzalez, J. Kim, R. Jacob, and N. Shastri. ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature*, 419(6906):480–483, 2002.

[5] K. Falk, O. Rötzschke, S. Stevanovic, G. Jung, and H. G. Rammensee. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*, 351(6324):290–296, 1991.

[6] R. A. Henderson, H. Michel, K. Sakaguchi, J. Shabanowitz, E. Appella, D. F. Hunt, and V. H. Engelhard. HLA-A2.1-associated peptides from a mutant cell line: a second pathway of antigen presentation. *Science*, 255(5049): 1264–1266, 1992.

[7] M. L. Wei and P. Cresswell. HLA-A2 molecules in an antigen-processing mutant cell contain signal sequence-derived peptides. *Nature*, 356(6368): 443–446, 1992.

[8] M. V. Larsen, M. Nielsen, A. Weinzierl, and O. Lund. TAP-Independent MHC Class I Presentation. *Curr. Immunol. Rev.*, 2:233–245, 2006.

[9] F. Momburg and H. Hengel. Corking the bottleneck: the transporter associated with antigen processing as a target for immune subversion by viruses. *Curr. Top. Microbiol. Immunol.*, 269:57–74, 2002.

[10] B. Seliger, A. Höhne, A. Knuth, H. Bernhard, B. Ehring, R. Tampe, and C. Huber. Reduced membrane major histocompatibility complex class I density and stability in a subset of human renal cell carcinomas with low TAP and LMP expression. *Clin. Cancer Res.*, 2(8):1427–1433, 1996.

[11] A. K. Johnsen, D. J. Templeton, M. Sy, and C. V. Harding. Deficiency of transporter for antigen presentation (TAP) in tumor cells allows evasion of immune surveillance and increases tumorigenesis. *J. Immunol.*, 163(8): 4224–4231, 1999.

[12] M. Schirle, W. Keilholz, B. Weber, C. Gouttefangeas, T. Dumrese, H. D. Becker, S. Stevanovic, and H. G. Rammensee. Identification of tumor-associated MHC class I ligands by a novel T cell-independent approach. *Eur. J. Immunol.*, 30(8):2216–2225, 2000.

[13] P. Kavathas, F. H. Bach, and R. DeMars. Gamma ray-induced loss of expression of HLA and glyoxalase I alleles in lymphoblastoid cells. *Proc. Natl. Acad. Sci. U. S. A*, 77(7):4251–4255, 1980.

[14] R. DeMars, C. C. Chang, S. Shaw, P. J. Reitnauer, and P. M. Sondel. Homozygous deletions that simultaneously eliminate expressions of class I and class II antigens of EBV-transformed B-lymphoblastoid cells. I. Reduced proliferative responses of autologous and allogeneic T cells to mutant cells that have decreased expression of class II antigens. *Hum. Immunol.*, 11(2): 77–97, 1984.

[15] A. O. Weinzierl, C. Lemmel, O. Schoor, M. Müller, T. Krüger, D. Wernet, J. Hennenlotter, A. Stenzl, K. Klingel, H. G. Rammensee, and S. Stevanovic. Distorted relation between mRNA copy number and corresponding major histocompatibility complex ligand density on the cell surface. *Mol. Cell Proteomics.*, 6(1):102–113, 2007.

[16] F. M. McCarthy, S. C. Burgess, B. H. van den Berg, M. D. Koter, and G. T. Pharr. Differential detergent fractionation for non-electrophoretic eukaryote cell proteomics. *J. Proteome. Res.*, 4(2):316–324, 2005.

[17] J. Adachi, C. Kumar, Y. Zhang, and M. Mann. In-depth analysis of the adipocyte proteome by mass spectrometry and bioinformatics. *Mol. Cell Proteomics.*, 6(7):1257–1273, 2007.

[18] E. M. Krämer-Albers, N. Bretz, S. Tenzer, C. Winterstein, W. Möbius, H. Berger, K. A. Nave, H. Schild, and J. Trotter. Oligodendrocytes secrete exosomes containing major myelin and stress-protective proteins: trophic support for axons? *Proteomics*, in press, 2007.

[19] S. Tenzer, L. Stoltze, B. Schönfisch, J. Dengjel, M. Muller, S. Stevanovic, H. G. Rammensee, and H. Schild. Quantitative analysis of prion-protein degradation by constitutive and immuno-20S proteasomes indicates differences correlated with disease susceptibility. *J. Immunol.*, 172(2):1083–1091, 2004.

6. Results III: Distorted relation between mRNA copy number and corresponding MHC ligand density on the cell surface

This chapter has been published in *Molecular and Cellular Proteomics*, 6(1):102-113, 2007 by:
Andreas O. Weinzierl, Claudia Lemmel, Oliver Schoor, Margret Müller, Tobias Krüger, Dorothee Wernet, Jörg Hennenlotter, Arnulf Stenzl, Karin Klingel, Hans-Georg Rammensee and Stefan Stevanović.

The author of this thesis designed and performed all experiments except gene expression analysis of RCC100 as well as large parts of the MS-analysis of RCC100.

6.1. Summary

The major histocompatibility complex (MHC) presents peptides derived from degraded cellular proteins to T-cells and is thus crucial for triggering specific immune responses against viral infections or cancer. Up to now, there has been no evidence for a correlation between levels of mRNA (the 'transcriptome') and the density of MHC:peptide complexes (the 'MHC ligandome') on cells. Since such dependencies are of intrinsic importance for the detailed understanding of translation efficiency and protein turnover, and thus for systems biology in general and for tumour immunotherapy in practical application, we quantitatively analysed the levels of mRNA and corresponding MHC ligand densities in samples of renal cell carcinomas and their autologous normal kidney tissues. Relative quantification was carried out by gene chip analysis and by stable isotope peptide labeling, respectively. In comparing more than 270 pairs of gene expression and corresponding peptide presentation ratios, we demonstrate that there is no clear correlation ($R = 0.32$) between mRNA levels and corresponding MHC peptide levels in RCC. A significant number of peptides presented predominantly on tumour or normal tissue showed no or only minor changes in mRNA expression levels. In several cases, peptides could even be identified in spite of the virtual absence of the respective mRNA. Thus we conclude that a majority of epitopes from tumour associated antigens will not be found in approaches based mainly on mRNA expression studies, as mRNA expression reflects a distorted picture of the situation on the cell surface as visible for T cells.

6.2. Introduction

Major histocompatibility complex (MHC) bound peptides reflect a snapshot of the dynamic protein pool [1] and can therefore provide insights into protein turnover and the transient cellular proteome [2–4]. Class I MHC molecules present peptides (8-11 AA in length) [5] from degraded proteins on the cell surface to cytotoxic T-cells in order to allow the T-cells to discriminate between self and non-self [6]. Additionally to its original purpose, this detection system can be used for the selective removal of mutated cells displaying aberrant MHC:peptide complexes on their cell surface [7]. Therefore the identification of disease-related MHC:peptide complexes is of utmost interest for immunodiagnostics and -therapy. Huge efforts are being undertaken in order to identify cancer specific MHC:peptide complexes applicable for tumour immunotherapy [8–10]. In this context, the identification of tumour associated antigens (TAAs), which are predominantly expressed in tumour tissue, is one strategy being followed. Apart from differential comparisons of tumour and normal tissue proteomes [11], mRNA based techniques are also being widely applied. DNA expression cloning, for example, has proved to be one successful method for the identification of TAAs [12] but it is being increasingly complemented by large scale gene expression analysis using DNA microarrays [13, 14].

The generation of MHC:peptide complexes starts in the cytosol, where proteins are degraded by the proteasome [15]. The source proteins for proteasomal digestion range from long-lived proteins over short-lived proteins to defective ribosomal products (DRiPs) [16] which do not reach proper folding and functional state. DRiPs are thought to be responsible for more than 30% of all MHC bound peptides [17]. Irrespective of their origin, a small part of peptides generated in the cytosol escapes further degradation and is transported via the transporter associated with antigen processing (TAP) into the endoplasmatic reticulum [18] where they can be further trimmed N-terminally by proteases such as the endoplasmic reticulum aminopeptidase associated with antigen processing (ERAAP) [19]. Finally, the peptides are loaded onto MHC molecules with the help of several proteins that includes tapasin, calnexin and calreticulin, and are exported as MHC:peptide complexes onto the cell surface. An alteration which is thought to be crucial for the repertoire of MHC ligands is the induction of immunoproteasomal subunits by interferon gamma (IFNγ), for example. Upon IFNγ stimulation cells exchange the three proteolytic subunits of the proteasome (PSMB1, 2, and 5) with subunits of different proteolytic activity (LMP2, LMP7, and MECL1), a hallmark for converting normal proteasomes into immunoproteasomes [20, 21].

The entire MHC pathway has been investigated in depth often with a focus on the therapeutic potential. MHC bound peptides can be used for stimulation and activation of cytotoxic T-cells which have, for example, been used successfully for the treatment of melanoma [22, 23]. The identification of suitable T-cell epitopes is one of the bottlenecks in the large scale use of this tumour immunological

strategy. Utilisable T-cell epitopes have to be both tumour-associated and match the patient's MHC repertoire. Several T-cell epitopes for T-cell based tumour immunotherapy have already been described but the identification is laborious and not all *in vitro* defined T-cell epitopes are potent for triggering an immune response also *in vivo*.

Gene chip analysis is frequently used for the identification of TAAs [24, 25], but it has not yet been investigated whether an over-expression of mRNA also results in an MHC over-presentation of peptides derived from the respective protein. Thus, in this study MHC peptide presentation levels are compared with their corresponding transcription levels. We analysed human clear cell renal cell carcinoma and autologous normal tissue in a combined transcriptome and human leukocyte antigen (HLA) ligandome approach. HLA:peptide levels were quantified by mass spectrometry using stable isotope labels, mRNA levels were determined by gene chip analysis using DNA microarrays.

6.3. Materials and Methods

6.3.1. Materials

The HPLC reagents, trifluoroacetic acid, acetonitrile, formic acid and HPLC water, were purchased from Merck. Peptide modification reagent O-Methyl isourea hemisulfate was purchased from Acros Organics, 1-([$^1H_4$/$^2D_4$] nicotinoyloxy)succinimide (light or heavy dNIC-NHS) was synthesized as described elsewhere [26].

6.3.2. Elution of HLA presented peptides

HLA presented peptides were obtained by immune precipitation of HLA molecules from solid tissues using a slightly modified protocol [27] that involves the HLA-A, -B and -C specific antibody W6/32 coupled to protein A-sepharose or CNBr-activated sepharose (Roche) followed by acid elution and subsequent ultrafiltration.

6.3.3. Modification of peptides

Modification of peptides was carried out as described elsewhere [26] with slight modifications. In brief, peptide solutions were adjusted to pH 11 using 10 M NaOH and guanidinated at 65°C for 10 min using 2.5 M O-methyl isourea hemisulfate solution. The reaction was halted adjusting the pH to 3 and peptides were desalted by using Peptide Cleanup C18 Spin Tubes (Agilent) as described in the manual. Peptides were loaded three times in 200 μl aliquots onto the Spin Tube columns. These were then washed three times with H_2O and the peptides were

nicotinylated on column for 15 min at room temperature. 500 μl of a 20 mM light or heavy dNIC-NHS solution in 50 mM phosphate buffer pH = 8.5 were thus passed over the C18-material with a flow rate of 33 μl/min. After washing the columns as described above, aminolysis of unwanted Tyr-dNIC esters was carried out by treatment with 500 μl 50% hydroxylamine for 10 min at room temperature with a flow rate of 50 μl/min. Subsequently, the columns were washed again and the peptides were eluted using four times 50 μl of 50% acetonitrile, 1% formic acid.

6.3.4. Mixing of peptides

In order to mix tumour and normal MHC peptides in a total peptide ratio of 1:1, absorption of isotope labeled peptide pools was determined at 260 nm [28]. Peptide content of the tumour and normal sample were calculated using the molar extinction coefficient for nicotinic acid in 50% acetonitrile (ϵ = 1430 M^{-1}) and equal amounts of peptides were mixed.

6.3.5. Microcapillary LCMS

Peptide analysis was carried out as described elsewere [26] using an Ultimate HPLC system (Dionex) with a gradient ranging from 15-55% solvent B within 170 minutes. Mixed tumour and normal samples were recorded in an LCMS experiment without fragmentation using a hybrid quadrupol orthogonal acceleration time of flight MS/MS (Q-TOF, Micromass), equipped with a micro-ESI source. For sequence analysis tumour and normal samples were analysed separately in individual LCMS/MS experiments.

6.3.6. Peptide sequence analysis and peptide quantification

Peptide sequence analysis was carried out manually. Therefore MS/MS spectra were smoothed using the MassLynx 4.0 software (Savitzky-Golay, 3 smooth windows, 2 smoothes). For manual peptide identification, sequence tag searches were done using Mascot 2.0 software (peptide and MS/MS tolerance 0.2 Da, NCBInr database monthly updated and restricted for search to human entries, 138263 entries for human proteins at the time when the searches were done) and relevant hits were assessed manually and not by evaluation of the Mascot score. Criteria for manual identification were: a reasonable interpretation of at least 95% of all of all fragment peaks, complete sequence coverage with MS/MS fragments and signal intensities of fragment ion peaks which match breakage probabilities of the respective sequence. For more than 50% of all sequenced peptides MS/MS spectra with both light and heavy dNIC isotope label were available, which facilitated manual sequencing of the peptides. Ions containing the N-terminus showed a mass shift of 4 Da due to the dNIC isotope label (see Fig. 6.2-C and D). In total

363 peptides were sequenced in the three analyzed RCC samples. All identified peptides were blasted against the NCBInr database (see above) and all peptides which could not be unambiguously allocated to only one protein sequence were excluded from further analysis unless the mRNA expression ratios for these potential source proteins were equal. For all peptides which could be linked to an mRNA expression ratio Supplementary Tab. A.2 lists protein information, HLA presentation ratio, mRNA expression ratio, HLA restriction, source, precursor mass and charge state. All relevant MS/MS spectra used for peptide identification can be found at
http://www.uni-tuebingen.de/uni/kxi/PaperSupplements/Weinzierl_MCP2006/.
For peptide quantification mixed tumour and normal tissue samples were analyzed in an MS experiment. Peptide pairs were identified manually. Afterwards mass chromatograms for the light and heavy version of the peptide were calculated and scans containing the peptide of interest were summed up. For background subtraction an equal number of scans with a similar retention time was subtracted. The relative amount of light and heavy dNIC peptide was calculated using the peak heights of the first isotope peaks (see Fig. 6.2-B).

6.3.7. Assessment of false positive rate of peptide identification

In order to estimate the false positive rate of the identified peptides, a database was designed which contained both the EBI IPI-Human database (IPI, version 3.21, containing 60822 entries) and the reversed EBI IPI-Human database (rIPI). 50 randomly chosen peptides from RCC099 were searched using the manually identified sequence tags in this database with the MASCOT software. In summary 20% of all searches returned only one peptide with a significant MASCOT score, 66 % of all searches returned only one hit (all in IPI). Only 6 % of all searches resulted in more than three peptide hits (both in IPI and rIPI, Supplementary Tab. A.3). 8 of theses peptides and 8 additional peptides – also identified in RCC099 – were chemically synthesized and modified (Supplementary Tab. A.1). Fragmentation spectra of these synthetic peptides were compared to the fragmentation spectra recorded from RCC099. All synthetic peptides showed exactly the same fragmentation pattern as the peptides identified in RCC099. Thus we conclude that the false positive rate in our sequence analysis is below 1%. Comparisons of MS/MS spectra obtained from synthetic peptides and from RCC099 can also be found at http://www.uni-tuebingen.de/uni/kxi/PaperSupplements/Weinzierl_MCP2006/.

6.3.8. Gene expression analysis by high-density oligonucleotide microarrays

RNA isolation from tumour and autologous normal kidney specimens as well as gene expression analysis by Affymetrix Human Genome U133 Plus 2.0 oligonucleotide microarrays (Affymetrix, Santa Clara, CA, USA) were performed as described previously [29]. Data were analysed with the GCOS software (Affymetrix). Pairwise comparisons between tumour and autologous normal kidney were calculated using the respective normal array as baseline.

6.4. Results

6.4.1. Principles of quantitative peptide and mRNA analysis

The general strategy used to assess the correlation between mRNA ratios and their corresponding ratios of HLA presented peptides is depicted in Fig. 6.1-A. mRNA was isolated from three different RCC tumour tissues and respective autologous normal tissues for subsequent gene chip analysis. HLA presented peptides were isolated from the same tumour and normal tissue samples [30] and quantified relatively as shown in Fig. 6.1-B using the differential N-terminal isotope coding (dNIC) strategy [26, 28]: The HLA ligandome isolated from tumour tissue was derivatized with 'heavy' nicotinic acid (deuterium-bearing, Fig. 6.1-C) whereas for the HLA ligandome of normal tissue the 'light' nicotinic acid (hydrogen-bearing, Fig. 6.1-C) was used. As a consequence, all peptides from tumour tissue possessed the very same physico-chemical features as their counterparts with identical sequence from normal tissue, but differed in mass by 4 Da. In order to achieve the highest possible accuracy and sensitivity both for quantification and sequence identification, quantification and sequence identification experiments were carried out seperately. Therefore, after derivatization one aliquot of each peptide pool was used for relative quantification. Heavy dNIC-peptides from tumour samples and light dNIC-peptides from normal tissue were mixed in a total peptide ratio of 1:1 and subjected to liquid chromatography online coupled to an electrospray ionization mass spectrometer (ESI-LCMS). Pairs of identical peptides derived from tumour and normal tissue were identified on the basis of a 4 Da mass difference introduced by the isotope label and due to their identical retention times in chromatography. For differential quantification, the ratio of signal intensities between monoisotopic peaks of the tumour and normal tissue derived peptide was calculated (Fig. 6.1-B, right panel). For peptide sequence analysis, the remaining aliquot of derivatized HLA ligand pools was used for peptide sequence analysis by online coupled electrospray ionization tandem mass spectrometry (ESI-LCMS/MS).

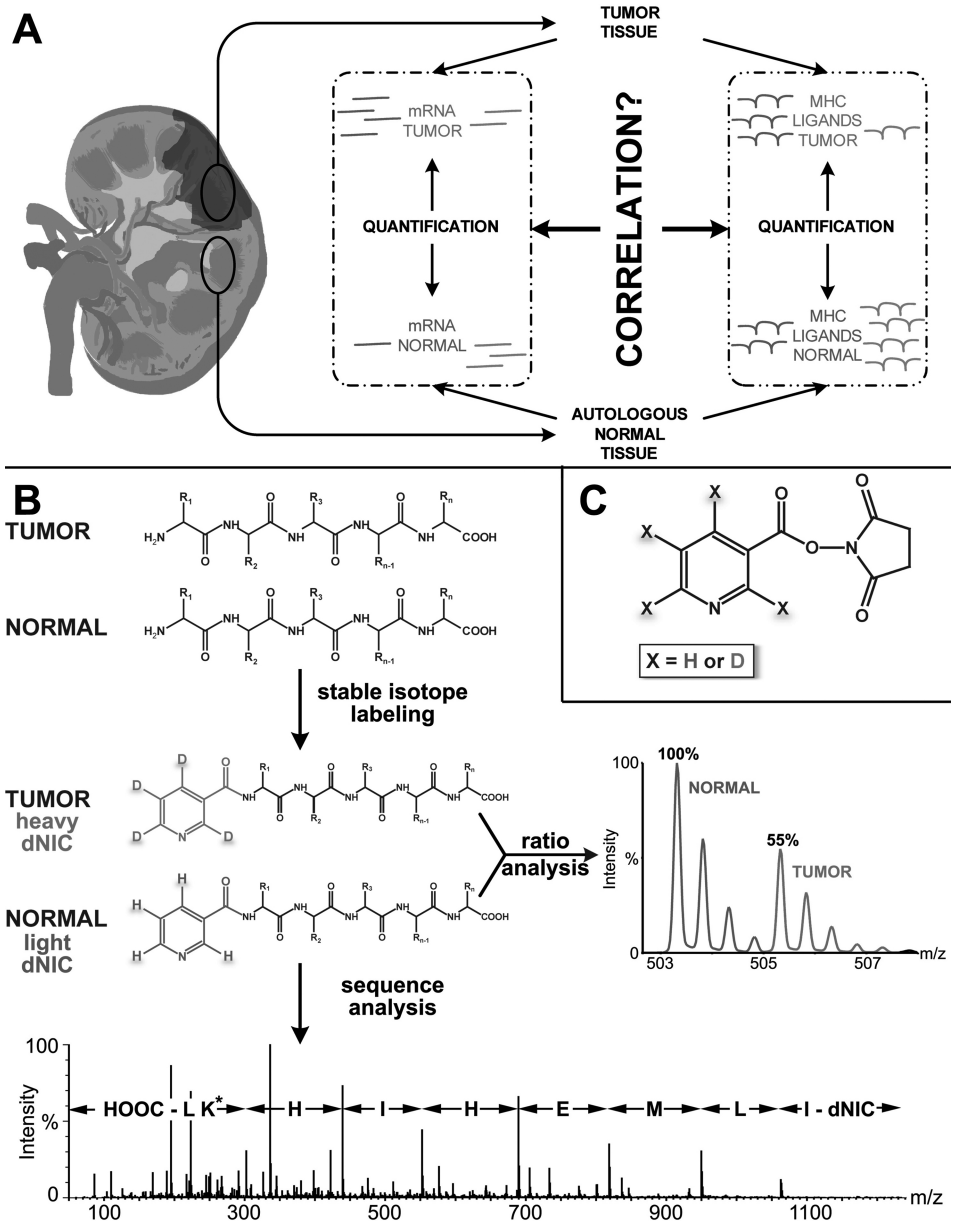

Figure 6.1.: Strategy for differential mRNA and MHC ligand analysis. A, mRNA and MHC ligands were isolated from tumour and autologous normal tissue. Relative quantification was carried out for both specimens and the correlation between the quantitative data was determined. B, Modification strategy for MHC ligands: Tumour derived ligands were modified with heavy dNIC reagent, normal tissue with the light form. Sequence analysis was carried out individually for tumour and normal samples, K* indicates a guanidinated lysine side chain; for ratio analysis both samples were mixed in a total peptide ratio of 1:1. C, dNIC modification reagent with X representing the site of stable isotope labeling.

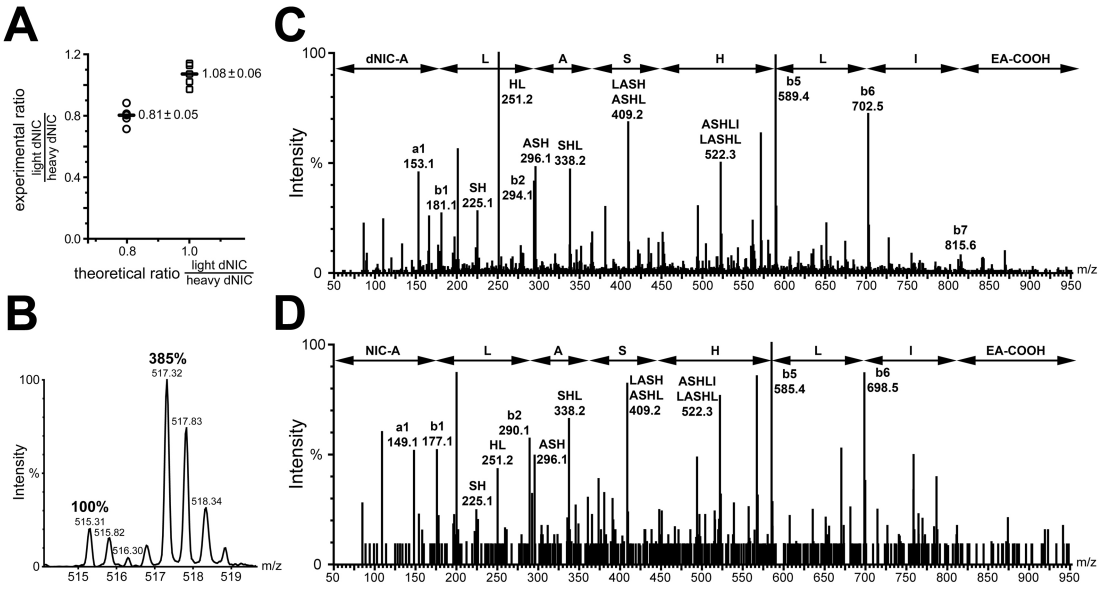

Figure 6.2.: Peptide ratio analysis and sequence identification. A, For estimation of methodical deviations a mix of six synthetic peptides was modified with the dNIC strategy using either heavy or light dNIC in two independent experiments. Heavy and light samples were mixed in a total peptide ratio of 0.8:1 and 1:1. Actual peptide ratios were determined by mass spectrometry, mean and standard deviation are stated. B-D, MHC ligands were isolated from RCC100 tumour and normal tissue and subsequently labeled differentially with heavy respectively light dNIC. Ratio determination of differentially labeled ALASHLIEA was done in MS mode (B). MS/MS fragment spectra (C and D) were used for sequence identification.

Fragmentation experiments were carried out separately for each tumor and normal tissue (Fig. 6.1-B, bottom panel). Thus an unambiguous allocation to tumor or normal tissue was possible for each peptide signal, even in the case of peptides for which the corresponding sequence could not be determined. The separate analysis of tumour and normal tissue also allowed the exclusion of singlet signals, which occurred due to incomplete modification with dNIC, from further analysis. Quantitative information was linked to sequence information by peptide individual features combining mass, charge state and retention time. In order to estimate the intrinsic methodological error, this quantification strategy was also applied to synthetic peptide pools, resulting in a calculated error for peptide pairs with high MS-signal intensity of less than 10% (Fig. 6.2-A).

6.4.2. mRNA vs HLA peptide: quantitative comparisons

In our study, this relative mRNA and HLA peptide quantification strategy was extended to three RCC samples sharing two HLA-alleles: RCC099 (HLA-A*02, -A*03, -B*27, -B*57), RCC100 (HLA-A*02, -A*03, -B*07, -B*18) and RCC110

(HLA-A*02, -A*68, -B*18, -B*27). Tab. 6.1 gives an overview of the identified mRNA and HLA peptide ratios, the number of identified sequences and the peptides for which sequence information could be linked to its corresponding presentation ratios. In total, intensity ratios between tumour and normal tissue were determined for 728 peptide pairs (Fig. 6.2-B), 363 peptides of which could be sequenced (Fig. 6.2-C and D). 273 peptide sequences could be linked to a specific mRNA allowing the evaluation of a possible correlation between differences in gene expression and HLA peptide presentation (Supplementary Tab. A.2).

Table 6.1.: Absolute numbers of evaluated mRNA ratios, HLA ligand ratios and peptide sequences identified in this study. In the last column for all RCCs the number of successful allocations of peptide sequences to the corresponding HLA ligand ratios and mRNA ratios is given.

source	mRNA ratios	HLA ligand ratios	identified peptide sequences	allocated sequence, peptide and mRNA ratios
RCC099	18976	452	213	166
RCC100	24485	142	59	43
RCC110	26571	134	91	64
total	70032	728	363	273

In order to be able to combine the HLA-mRNA comparisons of all three RCCs datasets, relative mRNA expression ratios and HLA peptide presentation levels had to be normalised individually for each analyzed tissue pair (Fig. 6.3-A). To do so, histograms of logarithmized HLA peptide presentation levels were plotted and subsequently fitted assuming a Gaussian distribution. Each fitted HLA histogram plot had its maximum near 0, representing a 1:1 ratio of presentation on tumour and normal tissue. To take experimental inaccuracies into account regarding the mixing of HLA peptides from each tumour and normal tissue pair, the HLA presentation ratios were normalised such that the maxima of their fitted histogram plots were located at 0. Normalised HLA peptide presentation data of all three RCCs were cumulated and Gaussian fitted in order to determine the presentation level representing the 5% highest over- or under-presentation ($x_{TOP5\%}$, Fig. 6.3-B). An analogue normalisation and assessment of the $x_{TOP5\%}$ was carried out for the mRNA expression data.

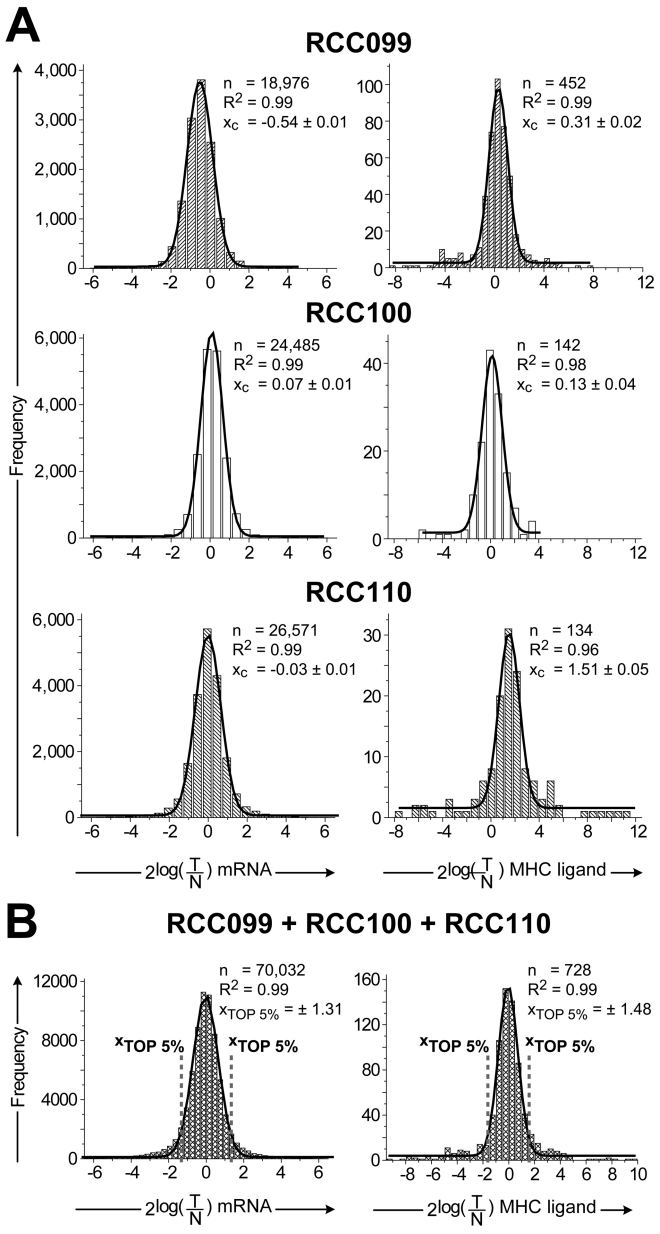

Figure 6.3.: Normalization of quantitative mRNA and MHC ligand data. mRNA ratios (left panels) and MHC ligand ratios (right panels) between tumour and normal tissue were logarithmized, frequency counts were made with bin sizes from 0.3 to 0.7. The bin sizes were chosen so that the Gaussian fit resulted in the best R^2-values. Using these Gaussian fit functions, the mean value (x_c) of each data set was calculated. Panel A, shows data sorted for RCC099, RCC100, and RCC110. All data were normalised to x_c=0. B, The margins of relative gene transcript expression representing the 5% strongest over/ under-expression ($x_{TOP5\%}$) were calculated as follows: Normalised mRNA data for all RCCs were cumulated and Gaussian fitted, the area under this curve was calculated and the x-values corresponding to 5% (for under-presentation) and 95% (for over-presentation) of the area were determined. $x_{TOP5\%}$ of the HLA ligandome ratios were calculated similarily.

Finally, normalised mRNA ratios were plotted against their corresponding normalised HLA presentation ratio Fig. 6.4) and the Spearman Rank correlation coefficient was calculated at R = 0.32, showing an even weaker correlation of transcriptome to HLA ligandome than the correlation of transcriptome to proteome (R = 0.45) [4]. For qualitative data analysis and simplified data discussion Fig. 6.4 was divided into nine areas by the $_{TOP5\%}$margins of mRNA and peptide ratios as indicated by the dashed lines Fig. 6.3-B).

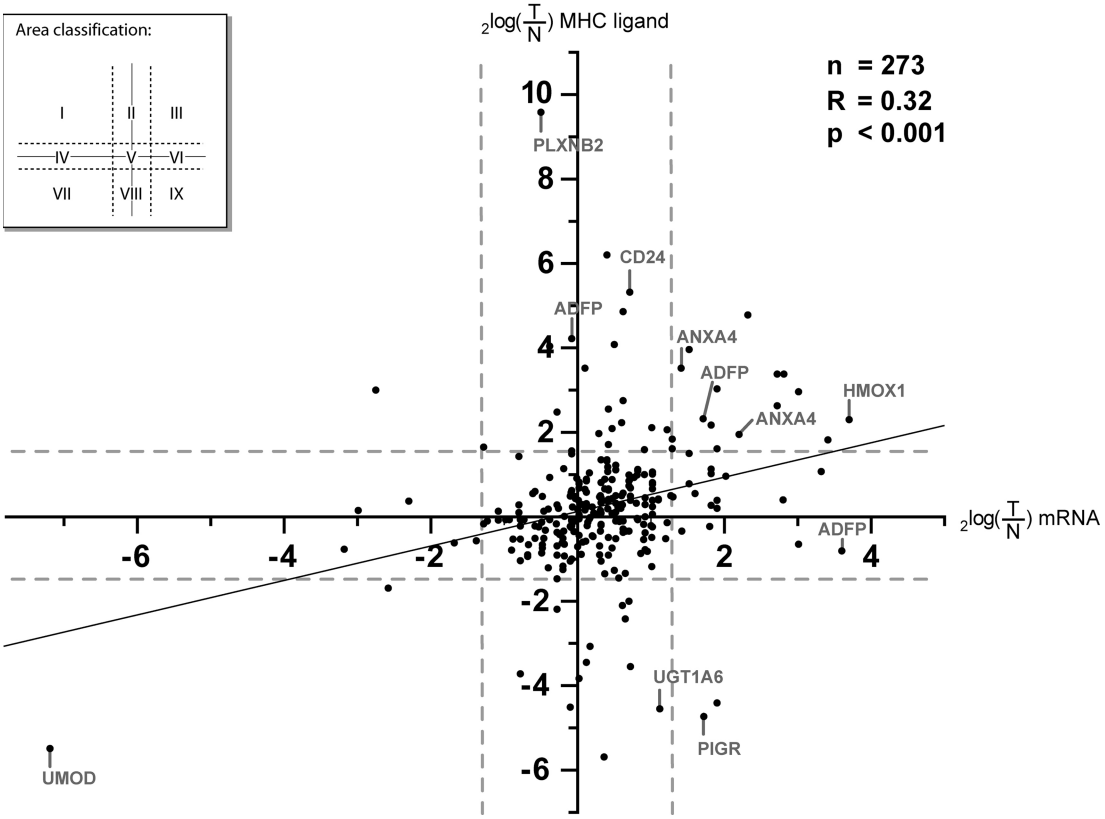

Figure 6.4.: Scatter plot representation of the correlation between relative mRNA expression and relative MHC ligand presentation. The $_2$log of the mRNA and MHC ligand ratios of tumour vs. normal tissue were calculated and plotted for 273 pairs of mRNA and MHC ligand ratios. Linear regression is shown as a solid line and the Spearman Rank correlation coefficient was calculated. Dashed lines designate the margins for the 5% of mRNA (x = ± 1.3) and peptide (y = ± 1.5) specimens with the highest under- and over-expression respectively presentation. Thus, the plot is divided into 9 areas (see insert). Selected loci are highlighted.

6.4.3. HLA ligand pools derived from tumour and normal tissue

For approximately 75% of all identified HLA ligands we could not detect significant changes neither in their HLA presentation ratio nor in their corresponding mRNA expression levels (Fig. 6.4, area V). Regarding mRNA expression this might have been expected as it indicates on a transcriptional basis that the tumour tissue still resembles its corresponding normal tissue. As to the HLA level this is more surprising as gene chip analysis shows upregulation of interferon gamma inducible proteins in the antigen processing machinery (Tab. 6.2).

Table 6.2.: Gene Chip analysis of gene transcripts involved in antigen processing. Analysis comprised constitutive proteasomal subunits (PSMB1, 2 and 5), immunoproteasomal subunits (LMP7, 2 and MECL1), subunits of the immunoproteasomal regulator PA28 (PSME1-3), proteins of the HLA peptide loading complex (TAP1 and 2, tapasin, calreticulin, calnexin and PDIA3), HLA class I and II (HLA-A, -B, -C, -DR, -DP and -DQ), the endoplasmic reticulum aminopeptidase associated with antigen processing (ERAAP) and proteins of the SEC61 translocon which can be involved in peptide export from the ER. Proteins inducible by interferon gamma are highlighted in bold.

Gene Symbol	Gene Title	RCC099 $_2log$ mRNA ratio	RCC100 $_2log$ mRNA ratio	RCC110 $_2log$ mRNA ratio
PSMB1	proteasome subunit beta 1	0.4	0.4	0.6
PSMB2	proteasome subunit beta 2	0.3	0.1	0.5
PSMB5	proteasome subunit beta 5	0.0	-0.2	0.2
LMP7	**proteasome subunit beta 8**	**2.5**	**2.1**	**2.5**
LMP2	**proteasome subunit beta 9**	**1.8**	**1.7**	**3.5**
MECL1	**proteasome subunit beta 10**	**1.3**	**0.8**	**2.4**
PSME1	**proteasome activator subunit 1**	**0.2**	**0.2**	**0.7**
PSME2	**proteasome activator subunit 2**	**0.1**	**-0.3**	**1.4**
PSME3	**proteasome activator subunit 3**	**-0.1**	**-0.9**	**-0.6**
CALR	calreticulin	0.0	0.0	0.1
CANX	calnexin	1.5	0.8	2.0
TAP1	**transp. 1, ATP-bind.cass.**	**1.5**	**1.8**	**2.6**
TAP2	**transp. 2, ATP-bind.cass.**	**0.3**	**1.1**	**2.6**
TAPBP	**Tapasin**	**1.4**	**1.1**	**1.6**
HLA-A	**m.histocomp.complex I A**	**0.8**	**0.3**	**1.1**
HLA-B	**m.histocomp.complex I B**	**1.3**	**0.8**	**2.1**
HLA-C	**m.histocomp.complex I C**	**0.9**	**0.5**	**1.6**
HLA-DP	**m.histocomp.complex II DP**	**1.4**	**0.6**	**2.5**
HLA-DQ	**m.histocomp.complex II DQ**	**2.0**	**1.3**	**3.5**

continued on next page

Table 6.2 – contiuned

Gene Symbol	Gene Title	RCC099 $_2log$ mRNA ratio	RCC100 $_2log$ mRNA ratio	RCC110 $_2log$ mRNA ratio
HLA-DR	**m.histocomp.complex II DR**	**1.1**	**0.6**	**2.4**
PDIA3	prot. disulfide isomerase A3	0.2	0.1	0.7
SEC61A1	Sec61 alpha 1 subunit	0.6	-0.5	0.5
SEC61A2	Sec61 alpha 1 subunit	0.0	-0.9	-1.2
SEC61B	Sec61 alpha 1 subunit	-0.1	-0.2	0.6
SEC61G	Sec61 alpha 1 subunit	1.9	0.9	1.2
ERAP1	ER aminopep.assoc.AG.process.	0.7	0.7	0.8

In all tumour tissues analysed, immunoproteasomal subunits were clearly overexpressed (up to 11-fold for LMP2 in RCC110). Although immunoproteasomal activity is thought to change the HLA bound peptide repertoire [31] and although it has been shown that both normal and immunoproteasomes are capable of destroying each other's epitopes [32, 33], we could not detect major differences in the HLA ligandome of tumour versus normal tissue.

6.4.4. Multiple HLA ligands from one source protein

Tab. 6.3 lists mRNA and HLA presentation ratios of proteins which are highlighted in Fig. 6.4 and discussed below. Analysis of individual peptide pairs indicated tumour associated changes in HLA presentation ratios (Tab. 6.4).

Table 6.3.: Diversity of peptide and RNA ratios highlighted by 11 examples. Genes were sorted according to the area classification introduced in Fig. 6.4. HLA ligand ratios and mRNA ratios were calculated between tumour and autologous normal tissue. Gene Symbol and Title can be found at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene.

Area	Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	Source
II	PLXNB2	plexin B2	TYTDRVFFL	9.58	−0.50	RCC110
	ADFP	adip.diff.-related prot.	VRLGSLSTK	4.22	−0.08	RCC099
	CD24	CD24 antigen	RAMVARLGL	5.32	0.71	RCC100
III	ANXA4	annexin A4	DEVKFLTV	3.52	1.41	RCC100
	ADFP	adip.diff.-related prot.	SLLTSSKGQLQK	2.32	1.71	RCC100
	ANXA4	annexin A4	DEVKFLTV	1.95	2.20	RCC110
	HMOX1	heme oxygenase 1	KIAQKALDL	2.30	3.70	RCC110
VI	ADFP	adip.diff.-related prot.	VRLGSLSTK	−0.81	3.60	RCC110
VII	UMOD	uromodulin	RAFSSLGLLK	−5.49	−7.19	RCC100
VIII	UGT1A6	UDP glucuron.transfer.1-A6	ALGKIPQTV	−4.55	1.12	RCC099
IX	PIGR	polymeric Ig receptor	FSVVINQLR	−4.73	1.72	RCC099

Table 6.4.: Multiple peptides from one source protein identified in one tumour and normal tissue pair. HLA presentation ratios were calculated between a corresponding tumour and normal tissue pair. Peptides which were identified also in N-terminally trimmed forms are marked in bold. Gene Symbol and Title can be found at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene.

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source
ALDOA	aldolase A	ALSDHHIYL	3.38	2.72	A02	RCC099
ALDOA	aldolase A	RTVPPAVTGITF	2.63	2.72	B57	RCC099
CCNI	cyclin I	LLDRFLATV	−0.93	−0.18	A02	RCC099
CCNI	cyclin I	SLLDRFLATV	0.01	−0.18	A02	RCC099
EEF2	euk.transl.elongat.fact. 2	ILTDITKGV	−0.91	0.42	A02	RCC099
EEF2	euk.transl.elongat.fact. 2	RRWLPAGDAL	0.37	0.42	B27	RCC099
FLNA	filamin A	GTHKVTVLF	0.06	1.02	B57	RCC099
FLNA	filamin A	GTHTVSVKY	1.01	1.02	B57	RCC099
FLNA	filamin A	GVHTVHVTF	−0.24	1.02	B57	RCC099
GNB2L1	guan.nucleot.bind.prot. b2-L 1	KTIKLWNTL	1.06	1.02	A02	RCC099
GNB2L1	guan.nucleot.bind.prot. b2-L 1	YTDNLVRVW	0.24	1.02	B57	RCC099
HLA-A/G	m.histocomp.complex I A/G	VMAPRTLLL	−0.23	1.60	A02	RCC110
HLA-A	m.histocomp.complex I A	VMAPRTLVL	0.40	1.60	A02	RCC110
HLA-B	m.histocomp.complex I B	AAQITQRKW	1.84	1.29	B57	RCC099
HLA-B	m.histocomp.complex I B	TAAQITQRKW	1.61	1.29	B57	RCC099
HNRPM	heterogen.nucl.rib.prot. M	KSRGIGTVTF	−0.74	0.47	B57	RCC099
HNRPM	heterogen.nucl.rib.prot. M	LLFDRPMHV	0.43	0.47	A02	RCC099
LMNA	lamin A/C	KAGQVVTIW	0.29	0.07	B57	RCC099
LMNA	lamin A/C	MRARMQQQL	−0.60	0.07	B27	RCC099
PPP1CA/B/C	prot. phosphatase 1 alpha/b/c	KYPENFFLL	6.20	0.35	C	RCC110
PPP1CA	prot. phosphatase 1 alpha	SIIGRLLEV	0.47	0.35	A02	RCC110
RBBP4/7	retinoblastoma bind.prot. 4/7	HTAKISDFSW	−0.13	0.14	B57	RCC099
RBBP4/7	retinoblastoma bind.prot. 4/7	TAKISDFSW	0.43	0.14	B57	RCC099
RPS16	ribosomal prot. S16	ISKALVAYY	1.12	0.82	B57	RCC099
RPS16	ribosomal prot. S16	KLLEPVLLL	1.09	0.82	A02	RCC099
SCD	stearoyl-CoA desaturase	ARLPLRLFL	3.96	1.52	B27	RCC099
SCD	stearoyl-CoA desaturase	ITAGAHRLW	1.50	1.52	B57	RCC099
SSR1	signal sequence receptor	VLFRGGPRGLLAVA	−2.00	0.70	A02	RCC110
SSR1	signal sequence receptor	VLFRGGPRGSLAVA	−2.42	0.70	A02	RCC110
SSR1	signal sequence receptor	VLFRGGPRGLLAVA	0.89	0.37	A02	RCC099
SSR1	signal sequence receptor	VLFRGGPRGSLAVA	−1.35	0.37	A02	RCC099
STAT3	s.transd.activ.transcript.3	EERIVELF	0.10	−0.80	B18	RCC110
STAT3	s.transd.activ.transcript.3	EELQQKVSY	0.38	−0.80	B18	RCC110
TMED10	transmemb.traff.prot. 10	FLLGPRLVLA	−0.91	−0.08	A02	RCC099
TMED10	transmemb.traff.prot. 10	LLGPRLVLA	1.49	−0.08	A02	RCC099
TMEM66	transmembrane prot. 66	KGWDGYDVQW	0.39	0.72	B57	RCC099
TMEM66	transmembrane prot. 66	RRLDPIPQL	−3.55	0.72	B27	RCC099
TNS1	tensin 1	HAKVLEFGW	0.81	0.32	B57	RCC099
TNS1	tensin 1	FLIETGPRGV	0.45	0.32	A02	RCC099
VIM	vimentin	ARLDLERKV	1.13	1.82	B27	RCC099
VIM	vimentin	NLAEDIMRL	0.27	1.82	A02	RCC099
VIM	vimentin	NYIDKVRFL	1.02	1.82	C	RCC099

For 20% of all proteins with several identified HLA ligands, HLA presentation ratios differed between the individual peptides more than 4-fold. For four peptides, N-terminally trimmed forms were identified (CCNI, HLA-B, RBBP4 and TMED10). The presentation ratios of these peptide variants changed up to 5-fold.

This might indicate tumour associated changes in antigen processing [34] and especially alterations in trimming by the endoplasmic reticulum aminopeptidase associated with antigen processing (ERAAP) [35].

6.4.5. HLA ligands identified only in one tissue specimen

Several HLA ligands were found to be presented exclusively on one tissue specimen (Tab. 6.5), although mRNA ratios changed only marginally in most cases. Peptides derived from UDP glucuronosyltransferase 1 family, polypeptide A6 (UGT1A6), uromodulin (UMOD) and from the polymeric Ig receptor (PIGR) were presented uniquely on normal tissue (Fig. 6.4, areas VII-IX); peptides derived from annexin A4 (ANXA4), plexin B2 (PLXNB2), decycling heme oxygenase 1 (HMOX1) and ADFP, as well as from the small cell lung carcinoma cluster 4 antigen CD24, were presented solely on tumour tissue (Fig. 6.4, areas II and III). Peptides presented specifically on one tissue specimen might reflect both differences in antigen processing between one tumour/ normal tissue pair and patient specific variations: identical peptides from ANXA4 and ADFP were identified in different tumours but without being presented uniquely.

Table 6.5.: HLA ligands exclusively identified in either tumour or normal tissue. Values were calculated against background and sorted according to their HLA presentation ratio between tumour and autologous normal tissue. Gene Symbol and Title can be found at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene.

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source
PLXNB2	plexin B2	TYTDRVFFL	9.58	0.00	C	RCC110
PPP1CA/B/C	prot. phosphatase 1 alpha/b/c	KYPENFFLL	6.20	0.35	C	RCC110
CD24	CD24 antigen	RAMVARLGL	5.32	0.71	n/a	RCC100
SLC17A3	solute carrier family 17, member 3	ARYGIALVL	4.86	0.62	B27	RCC099
ADFP	adip.diff.-related prot.	VRLGSLSTK	4.22	−0.08	B27	RCC099
MAT1A/2A	met.adenosyltransf. Ia/Iia	RRVLVQVSY	4.08	0.50	B27	RCC110
SCD	stearoyl-CoA desaturase	ARLPLRLFL	3.96	1.52	B27	RCC099
ANXA4	annexin A4	DEVKFLTV	3.52	1.41	B18	RCC100
POLR2C	polymerase II pol.pep.C,33kDa	KLSDLQTQL	3.52	0.10	A02	RCC110
HMOX1	heme oxygenase 1	KIAQKALDL	2.30	3.70	A02	RCC110
CXCL14	chemokine ligand 14	RLLAAALLL	2.09	0.47	A02	RCC099
OGG1	8-oxoguan.DNA glycosyl.	VLADQVWTL	−3.07	0.17	A02	RCC099
MYL6	myosin, light pol.pep. 6	FVRHILSG	−3.45	0.12	n/a	RCC099
GNB5	guan.nucleot.bind.prot. b5	ILFGHENRV	−3.72	−0.78	A02	RCC099
HLA-B/C	m.histocomp.complex I B/C	DTAAQITQR	−4.41	1.58	A68	RCC110
RPL15	ribosomal prot. L15	EVILIDPFHK	−4.51	−0.75	A68	RCC110
UGT1A6	UDP glucuron.transfer.1-A6	ALGKIPQTV	−4.55	1.12	A02	RCC099
PIGR	polymeric Ig receptor	FSVVINQLR	−4.73	1.72	A03	RCC099
UMOD	uromodulin	RAFSSLGLLK	−5.49	−7.19	A03	RCC100
MYO1C	myosin IC	FLDHVRTSF	−5.69	0.36	A03	RCC100

For 60% of all source proteins in the group of the top 5% over-presented peptides (Fig. 6.4, area I-III), tumour association has been reported and this number actually increases to nearly 80% for peptides presented exclusively on tumour tissue (Supplementary Tab. A.5). As far as tumour immunotherapy is concerned, these uniquely presented peptides are excellent targets but they also underline the importance of patient specific adaptations in such an approach [9].

6.4.6. HLA ligands without corresponding mRNA

The fact that there are differences between the transcriptome and HLA ligandome becomes even more obvious when one considers that several peptides could be identified for which no mRNA was detectable in the respective tissues (Tab. 6.6). In RCC100, for example, the mRNA for the EH-domain containing protein 2 (EHD2) was detectable neither in tumour nor in normal tissue. Nevertheless, the EHD2-derived peptide ALASHLIEA was identified on tumour as well as on normal tissue (Fig. 6.2-B to D). These HLA ligands are presumably derived from long-lived proteins from which the mRNA has already been degraded [36], or represent mutated peptides for which the correct allocation to the correct mRNA was not possible any more.

Table 6.6.: HLA ligands identified from RCC099, RCC100 and RCC110 with no detectable mRNA in tumour and/ or normal tissue.Values were calculated against background; mRNA detection categorized by GCOS software in present (P), marginal present (M) or absent (A). Gene Symbol and Title can be found at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene.

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	Detect. mRNA Tumor vs. Normal	Source
CLIC5	Cl- intracell.chan. 5	NLLPKLHVV	3.00	-2.75	**A** / P	RCC099
ACVRL1	activin A receptor type II-L 1	SPRKGLLML	−0.65	3.01	P / **A**	RCC100
EHD2	EH-domain containing 2	ALASHLIEA	1.07	3.32	P / **A**	RCC099
EHD2	EH-domain containing 2	ALASHLIEA	1.82	3.41	P / **A**	RCC100
LSP1	lymphocyte-specific prot. 1	KLIDRTESL	0.39	1.90	P / **A**	RCC110
PLVAP	plasmalem.vesic.assocc.prot.	KVKTLEVEI	0.50	0.52	P / **A**	RCC099
SRXN1	sulfiredoxin 1 homolog	TLSDLRVYL	0.43	0.32	P / **M**	RCC099
WDR78	WD repeat domain 78	TSVVYDVAW	0.32	0.52	P / **A**	RCC099
ATOX1	ATX1 antiox.prot. 1 homol.	RVLNKLGGVK	−1.01	n/a	**A** / **A**	RCC100
CCDC21	coiled-coil dom.cont. 21	RLQMEQMQL	−1.14	n/a	**A** / **A**	RCC100
CYHR1	cysteine/histidine-rich 1	HLGPEGRSV	−1.21	n/a	**M** / **A**	RCC099
DHX38	DEAH box pol.pep. 38	VLFGLLREV	0.81	n/a	**A** / **A**	RCC099
FLJ32206	hypothetical protein FLJ32206	GSHFISHLS	−0.89	n/a	**A** / **A**	RCC099
GBP4	guanylate binding prot. 4	KRLGTLVVTY	−0.60	n/a	**A** / **A**	RCC099
KIAA1305	KIAA1305	TLADIIARL	−0.10	n/a	**A** / **A**	RCC099
RASL11A	RAS-L, family 11, member A	YLLPKDIKL	−0.37	n/a	**A** / **A**	RCC099

6.4.7. HLA ligands identified in several patient samples

Peptides identified in more than one patient allowed the evaluation of patient specific variations in mRNA expression ratios and HLA presentation ratios (Supplementary Tab. A.4). More than 12% of the HLA presented peptides exhibit a change in the presentation level in different patients that is greater than 4-fold and includes many of the significantly over- or under-presented peptides. The peptide VRLGSLSTK derived from the tumour associated antigen adipophilin (ADFP) [37] is under-presented on RCC110 but mRNA levels are clearly up-regulated in the tumour tissue (Fig. 6.4, area VI). Yet the same peptide is presented exclusively on RCC099 although the mRNA levels of ADFP remain unchanged between tumour and normal tissue (Fig. 6.4, area II). Another different ADFP-derived peptide (SLLTSSKGQLQK) is presented on RCC100 by HLA-A*03 with an mRNA ratio that matches its HLA presentation ratio (Fig. 6.4, area III). HLA peptide ratios exhibit stronger patient individual changes than mRNA ratios might indicate. This is due to the fact that changes in the HLA ligandome reflect not only mRNA ratio alterations, but also changes in the degradome, which to some extent is conserved in the HLA ligandome [38].

6.5. Discussion

6.5.1. Stable isotope labeling of *ex vivo* isolated MHC molecules

Stable isotope labeling is the method of choice for protein quantification and is widely used in proteomics. For *ex vivo* prepared MHC peptide samples the standard isotope labeling methods are either not feasible or are not applicable to the majority of MHC bound peptides due to their amino acid composition. To overcome these impediments we have established the differential N-terminal isotope coding (dNIC) strategy; to our knowledge this is the first study using an isotope-based quantitative MS-method for a large scale *ex vivo* comparison of the HLA ligandome of two tissue specimen. Patient samples were chosen for analysis according to two criteria: Categorization of the RCCs into the clear cell type by histology and the broadest possible overlap of the MHC alleles between the different patients. Tumour samples were taken from its central area, control samples from renal cortex assessed as 'normal' and as distant as possible from the tumour.

6.5.2. Individual variances in antigen processing

In the three RCC tissue pairs analysed in this study, each sample displays to some extent individual features regarding antigen processing. Apart from HLA mismatches, the activity of enzymes involved in antigen processing also varied.

We found, for example, varying mRNA levels of immunoproteasome subunits in the different samples (Tab. 6.2). All individual variances in antigen processing summarize at the level of the MHC ligandome. For example, we detected ADFP-derived peptides ranging in their presentation from exclusive presentation on tumour tissue to equal presentation on tumour and normal tissue. These individual variances in MHC:peptide levels are one of the reasons why for example tumour vaccination studies can vary extremely in their outcome.

6.5.3. Four major hypotheses

The HLA ligandome is more complex in its variations than mRNA but it is of far greater immunological relevance; it reflects a particularly complex aspect of systems biology that is exemplified here by a naturally occurring change in tissue (tumour genesis). The long and complex path from mRNA transcription to HLA presentation is reflected in our observation of a generally weak mRNA-HLA-ligand correlation. Four hypotheses can thus be deduced from Fig. 6.4:

First, HLA ligand generation implies that one mRNA species is able to account for several different peptides which differ in their presentation ratios and which are potentially presented by different HLA allotypes (Tab. 6.4).

Second, for more than 75% of all identified HLA ligands, we could detect no major differences in the HLA peptide repertoire [39] (Fig. 6.4, areas IV-VI), although immunoproteasomes were up-regulated on a transcriptional level. Assuming transcription of immunoproteasomal mRNA and activity of the immunoproteasomal subunits, this would limit the implication of the proteasome composition on the HLA ligandome in our *ex vivo* setting [36].

Third, several HLA ligands could be identified although no corresponding mRNA was detectable. In these cases HLA ligands might be derived either from long-lived proteins or from mutated proteins for which the correct peptide-mRNA allocation was not possible.

Fourth, strong differences in HLA presentation without mRNA changes (Fig. 6.4, area II and VIII) are due to altered peptide generation which can be triggered by several scenarios [40]: Higher translation efficiency leads to a larger number of defective ribosomal products (DRiPs) which are known to be a major source for HLA ligands [41]. The amount of DRiPs from certain source mRNAs might also be elevated due to mutations in some of the mRNAs, distorting transcription and thus increasing DRiP formation [17]. On the other hand, enhanced protein turnover increases the total amount of substrates for proteases such as the proteasome. Moreover, proteolytic activity in the cytosol or ER can vary for some peptides when tumour and normal tissue are compared.

6.5.4. Context and limitations of the chosen experimental setting

The exact molecular mechanism responsible for alterations in the level of our *ex vivo* isolated MHC:peptide complex levels could not be deduced. This is due to the fact that these alterations are most probably not caused by one single change in the peptide generation pathway or in the antigen processing machinery but are rather complex, multifactorial changes established during tumour genesis. Defining molecular mechanisms in a complex interaction network such as the antigen processing machinery is in many cases not possible. In a recent study by Milner et al. [38], the turnover kinetics of MHC pepides was determined in a cell culture setting. Not even in this plain setting could the 'simple' discrimination between short-lived proteins and DRiPs clearly be made. In these experiments the only strong indication for MHC:peptide complexes derived from DRiPs was a biphasic turnover kinetics in pulse chase experiments using incorporation of isotope labeled amino acids.

In 1999 and 2003 the first quantitative experiments were carried out to compare the yeast transcriptome with its proteome [2, 4]. Here, the authors showed only a weak correlation between transcriptome and proteome (R = 0.45). In our study, the correlation between transcriptome and HLA ligandome was even weaker (R = 0.32) which might reflect that the generation of HLA ligands is a pathway that begins only after mRNA transcription. At the present time, large-scale data regarding the extent of correlation between proteome and HLA ligandome are not available. Taking into consideration that a major portion of the HLA ligandome is supposed to be derived from the transient proteome – that is short-lived proteins and DRiPs – one could expect an even weaker correlation between the HLA ligandome and the well quantifiable permanent proteome.

6.6. Summary and outlook

mRNA-based identification of tumour associated antigens is well established and widely used. As we could show in this study that mRNA levels alone do not adequately reflect cellular reality at the HLA level, our results have a direct impact on T-cell based immunotherapy. We conclude that for the rational selection of appropriate tumour specific T-cell targets, quantitative HLA ligand analysis contributes greatly to the benefits of quantitative transcriptome analysis; a combination of both strategies enables the identification of new target candidates and helps to avoid false positive targets. In summary, comparing changes in the transcriptome to those in the HLA ligandome of renal tumour versus normal tissue without further reflection is like comparing apples to oranges. Conveying this finding to RNA based vaccination, which was shown to be a potent vaccine also against cancer [42], one cannot take for granted that every mRNA leads to actual

peptide presentation on HLA molecules. Thus for each mRNA that is designed for vaccination, intensive individual *in vitro* tests are advisable. Our findings provide insight into tumour associated changes in translation efficiency and protein turnover and provide immunologically relevant tumour information to a depth that could not be achieved before with hitherto existing standard techniques.

6.7. Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (SFB-TR19, SFB 685 and Graduiertenkolleg 794), the Nationales Genomforschungsnetz, and the Jürgen Manchot Stiftung. We thank Lynne Yakes for critical reading of the manuscript.

6.8. References

[1] A. Admon, E. Barnea, and T. Ziv. Tumor antigens and proteomics from the point of view of the major histocompatibility complex peptides. *Mol. Cell Proteomics.*, 2(6):388–398, 2003.

[2] S. P. Gygi, Y. Rochon, B. R. Franza, and R. Aebersold. Correlation between protein and mRNA abundance in yeast. *Mol. Cell Biol.*, 19(3):1720–1730, 1999.

[3] M. F. Princiotta, D. Finzi, S. B. Qian, J. Gibbs, S. Schuchmann, F. Buttgereit, J. R. Bennink, and J. W. Yewdell. Quantitating protein synthesis, degradation, and endogenous antigen processing. *Immunity.*, 18 (3):343–354, 2003.

[4] M. P. Washburn, A. Koller, G. Oshiro, R. R. Ulaszek, D. Plouffe, C. Deciu, E. Winzeler, and III J. R. Yates. Protein pathway and complex clustering of correlated mRNA and protein expression analyses in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A*, 100(6):3107–3112, 2003.

[5] K. Falk, O. Rotzschke, S. Stevanovic, G. Jung, and H. G. Rammensee. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*, 351(6324):290–296, 1991.

[6] R. M. Zinkernagel and P. C. Doherty. Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature*, 248(450):701–702, 1974.

[7] T. Wolfel, A. Van Pel, V. Brichard, J. Schneider, B. Seliger, K. H. Meyer zum Buschenfelde, and T. Boon. Two tyrosinase nonapeptides recognized

on HLA-A2 melanomas by autologous cytolytic T lymphocytes. *Eur. J. Immunol.*, 24(3):759–764, 1994.

[8] H. G. Rammensee, T. Weinschenk, C. Gouttefangeas, and S. Stevanovic. Towards patient-specific tumor antigen selection for vaccination. *Immunol. Rev.*, 188:164–176, 2002.

[9] T. Weinschenk, C. Gouttefangeas, M. Schirle, F. Obermayr, S. Walter, O. Schoor, R. Kurek, W. Loeser, K. H. Bichler, D. Wernet, S. Stevanovic, and H. G. Rammensee. Integrated functional genomics approach for the design of patient-individual antitumor vaccines. *Cancer Res.*, 62(20):5818–5827, 2002.

[10] N. Yajima, R. Yamanaka, T. Mine, N. Tsuchiya, J. Homma, M. Sano, T. Kuramoto, Y. Obata, N. Komatsu, Y. Arima, A. Yamada, M. Shigemori, K. Itoh, and R. Tanaka. Immunologic evaluation of personalized peptide vaccination for patients with advanced malignant glioma. *Clin. Cancer Res.*, 11 (16):5900–5911, 2005.

[11] S. Hanash, F. Brichory, and D. Beer. A proteomic approach to the identification of lung cancer markers. *Dis. Markers*, 17(4):295–300, 2001.

[12] van der Bruggen P., C. Traversari, P. Chomez, C. Lurquin, E. De Plaen, Eynde B. Van den, A. Knuth, and T. Boon. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science*, 254 (5038):1643–1647, 1991.

[13] M. Takahashi, D. R. Rhodes, K. A. Furge, H. Kanayama, S. Kagawa, B. B. Haab, and B. T. Teh. Gene expression profiling of clear cell renal cell carcinoma: gene identification and prognostic classification. *Proc. Natl. Acad. Sci. U. S. A*, 98(17):9754–9759, 2001.

[14] S. Mathiassen, S. L. Lauemoller, M. Ruhwald, M. H. Claesson, and S. Buus. Tumor-associated antigens identified by mRNA expression profiling induce protective anti-tumor immunity. *Eur. J. Immunol.*, 31(4):1239–1246, 2001.

[15] W. Baumeister, J. Walz, F. Zuhl, and E. Seemuller. The proteasome: paradigm of a self-compartmentalizing protease. *Cell*, 92(3):367–380, 1998.

[16] J. W. Yewdell, L. C. Anton, and J. R. Bennink. Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? *J. Immunol.*, 157(5):1823–1826, 1996.

[17] U. Schubert, L. C. Anton, J. Gibbs, C. C. Norbury, J. W. Yewdell, and J. R. Bennink. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature*, 404(6779):770–774, 2000.

[18] J. C. Shepherd, T. N. Schumacher, P. G. Ashton-Rickardt, S. Imaeda, H. L. Ploegh, Jr. C. A. Janeway, and S. Tonegawa. TAP1-dependent peptide translocation in vitro is ATP dependent and peptide selective. *Cell*, 74(3): 577–584, 1993.

[19] T. Serwold, F. Gonzalez, J. Kim, R. Jacob, and N. Shastri. ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature*, 419(6906):480–483, 2002.

[20] M. Gaczynska, K. L. Rock, and A. L. Goldberg. Gamma-interferon and expression of MHC genes regulate peptide hydrolysis by proteasomes. *Nature*, 365(6443):264–267, 1993.

[21] H. Hisamatsu, N. Shimbara, Y. Saito, P. Kristensen, K. B. Hendil, T. Fujiwara, E. Takahashi, N. Tanahashi, T. Tamura, A. Ichihara, and K. Tanaka. Newly identified pair of proteasomal subunits regulated reciprocally by interferon gamma. *J. Exp. Med.*, 183(4):1807–1816, 1996.

[22] J. Banchereau, H. Ueno, M. Dhodapkar, J. Connolly, J. P. Finholt, E. Klechevsky, J. P. Blanck, D. A. Johnston, A. K. Palucka, and J. Fay. Immune and clinical outcomes in patients with stage IV melanoma vaccinated with peptide-pulsed dendritic cells derived from CD34+ progenitors and activated with type I interferon. *J. Immunother.*, 28(5):505–516, 2005.

[23] Jr. C. L. Slingluff, G. R. Petroni, G. V. Yamshchikov, S. Hibbitts, W. W. Grosh, K. A. Chianese-Bullock, E. A. Bissonette, D. L. Barnd, D. H. Deacon, J. W. Patterson, J. Parekh, P. Y. Neese, E. M. Woodson, C. J. Wiernasz, and P. Merrill. Immunologic and clinical outcomes of vaccination with a multiepitope melanoma peptide vaccine plus low-dose interleukin-2 administered either concurrently or on a delayed schedule. *J. Clin. Oncol.*, 22(22): 4474–4485, 2004.

[24] J. M. Boer, W. K. Huber, H. Sultmann, F. Wilmer, A. von Heydebreck, S. Haas, B. Korn, B. Gunawan, A. Vente, L. Fuzesi, M. Vingron, and A. Poustka. Identification and classification of differentially expressed genes in renal cell carcinoma by expression profiling on a global human 31,500-element cDNA array. *Genome Res.*, 11(11):1861–1870, 2001.

[25] J. A. Lovisolo, B. Casati, L. Clerici, E. Marafante, A. V. Bono, N. Celato, and M. Salvadore. Gene expression profiling of renal cell carcinoma: a DNA macroarray analysis. *BJU. Int.*, 98(1):205–216, 2006.

[26] C. Lemmel, S. Weik, U. Eberle, J. Dengjel, T. Kratt, H. D. Becker, H. G. Rammensee, and S. Stevanovic. Differential quantitative analysis of MHC ligands by mass spectrometry using stable isotope labeling. *Nat. Biotechnol.*, 2004.

[27] M. Schirle, W. Keilholz, B. Weber, C. Gouttefangeas, T. Dumrese, H. D. Becker, S. Stevanovic, and H. G. Rammensee. Identification of tumor-associated MHC class I ligands by a novel T cell-independent approach. *Eur. J. Immunol.*, 30(8):2216–2225, 2000.

[28] A. O. Weinzierl and S. Stevanovic. LC-MS Based Protein and Peptide Quantification Using Stable Isotope Labels: From ICAT in General to differential N-terminal Coding (dNIC) in Special. *Biotechnol. Genet. Eng. Rev.*, 23: 21–39, 2006.

[29] T. Krüger, O. Schoor, C. Lemmel, B. Kraemer, C. Reichle, J. Dengjel, T. Weinschenk, M. Muller, J. Hennenlotter, A. Stenzl, H. G. Rammensee, and S. Stevanovic. Lessons to be learned from primary renal cell carcinomas: novel tumor antigens and HLA ligands for immunotherapy. *Cancer Immunol. Immunother.*, 54(9):826–836, 2005.

[30] G. Gastl, T. Ebert, C. L. Finstad, J. Sheinfeld, A. Gomahr, W. Aulitzky, and N. H. Bander. Major histocompatibility complex class I and class II expression in renal cell carcinoma and modulation by interferon gamma. *J. Urol.*, 155(1):361–367, 1996.

[31] R. E. Toes, A. K. Nussbaum, S. Degermann, M. Schirle, N. P. Emmerich, M. Kraft, C. Laplace, A. Zwinderman, T. P. Dick, J. Muller, B. Schonfisch, C. Schmid, H. J. Fehling, S. Stevanovic, H. G. Rammensee, and H. Schild. Discrete cleavage motifs of constitutive and immunoproteasomes revealed by quantitative analysis of cleavage products. *J. Exp. Med.*, 194(1):1–12, 2001.

[32] M. Basler, N. Youhnovski, Broek M. Van Den, M. Przybylski, and M. Groettrup. Immunoproteasomes down-regulate presentation of a subdominant T cell epitope from lymphocytic choriomeningitis virus. *J. Immunol.*, 173(6): 3925–3934, 2004.

[33] J. Chapiro, S. Claverol, F. Piette, W. Ma, V. Stroobant, B. Guillaume, J. E. Gairin, S. Morel, O. Burlet-Schiltz, B. Monsarrat, T. Boon, and B. J. Van den Eynde. Destructive cleavage of antigenic peptides either by the immunoproteasome or by the standard proteasome results in differential antigen presentation. *J. Immunol.*, 176(2):1053–1061, 2006.

[34] P. Van der Bruggen and B. J. Van den Eynde. Processing and presentation of tumor antigens and vaccination strategies. *Curr. Opin. Immunol.*, 18(1): 98–104, 2006.

[35] G. E. Hammer, F. Gonzalez, M. Champsaur, D. Cado, and N. Shastri. The aminopeptidase ERAAP shapes the peptide repertoire displayed by major histocompatibility complex class I molecules. *Nat. Immunol.*, 7(1):103–112, 2006.

[36] J. W. Yewdell. The seven dirty little secrets of major histocompatibility complex class I antigen processing. *Immunol. Rev.*, 207:8–18, 2005.

[37] S. M. Schmidt, K. Schag, M. R. Muller, T. Weinschenk, S. Appel, O. Schoor, M. M. Weck, F. Grunebach, L. Kanz, S. Stevanovic, H. G. Rammensee, and P. Brossart. Induction of adipophilin-specific cytotoxic T lymphocytes using a novel HLA-A2-binding peptide that mediates tumor cell lysis. *Cancer Res.*, 64(3):1164–1170, 2004.

[38] E. Milner, E. Barnea, I. Beer, and A. Admon. The turnover kinetics of major histocompatibility complex peptides of human cancer cells. *Mol. Cell Proteomics.*, 5(2):357–365, 2006.

[39] T. van Hall, A. Sijts, M. Camps, R. Offringa, C. Melief, P. M. Kloetzel, and F. Ossendorp. Differential influence on cytotoxic T lymphocyte epitope presentation by controlled expression of either proteasome immunosubunits or PA28. *J. Exp. Med.*, 192(4):483–494, 2000.

[40] D. Atkins, S. Ferrone, G. E. Schmahl, S. Storkel, and B. Seliger. Down-regulation of HLA class I antigen processing molecules: an immune escape mechanism of renal cell carcinoma? *J. Urol.*, 171(2 Pt 1):885–889, 2004.

[41] J. W. Yewdell, U. Schubert, and J. R. Bennink. At the crossroads of cell biology and immunology: DRiPs and other sources of peptide ligands for MHC class I molecules. *J. Cell Sci.*, 114(Pt 5):845–851, 2001.

[42] B. Scheel, S. Aulwurm, J. Probst, L. Stitz, I. Hoerr, H. G. Rammensee, M. Weller, and S. Pascolo. Therapeutic anti-tumor immunity triggered by injections of immunostimulating single-stranded RNA. *Eur. J. Immunol.*, 36(10):2807–2816, 2006.

7. Results IV: A cryptic VEGF T cell epitope: Identification and characterization by mass spectrometry and T cell assays

This chapter has been submitted for publication by:
Andreas O. Weinzierl*, Dominik Maurer*, Florian Altenberend, Nicole Schneiderhan-Marra, Karin Klingel, Oliver Schoor, Dorothee Wernet, Thomas Joos, Hans-Georg Rammensee and Stefan Stevanović.

*Andreas O. Weinzierl and Dominik Maurer contributed equally to this work.

The author of this thesis designed and performed all experiments except transfection and qRT-PCR experiments of X63-B*2705 cells as well as generation and analysis of VEGF-specific T cells.

7.1. Summary

The vascular endothelial growth factor A (VEGF) is involved in various physiological processes, such as angiogenesis or wound healing, but is also crucial in pathological events such as tumor growth. Thus, clinical anti-VEGF treatments have been developed which could already demonstrate beneficial effects for cancer patients. In this article we describe the first VEGF-derived $CD8^+$ T cell epitope. The natural HLA ligand SRFGGAVVR was identified by differential mass spectrometry in two primary renal cell carcinomas (RCC) and was significantly over-presented on both tumor tissues. SRFGGAVVR is derived from a cryptic translated region of VEGF presumably by initiation of translation at the non-classical start codon CUG^{499}. SRFGGAVVR specific T cells were generated *in vitro* using peptide loaded dendritic cells or artificial antigen presenting cells. SRFGGAVVR specific $CD8^+$ T cells – identified by HLA tetramer analysis after *in vitro* stimulation – were fully functional T-effector cells, which were able to secrete IFN-gamma upon stimulation and killed tumor cells *in vitro*. Additionally, we have quantitatively analyzed VEGF mRNA and protein levels in RCC tumor and normal tissue samples by gene chip analysis, qRT-PCR, *in situ* hybridization, and bead based immunoassay. In the future, T cells directed against VEGF as a tumor associated antigen may represent a possible way of combining peptide-based anti-VEGF immunotherapy with already existent anti-VEGF cancer therapies.

7.2. Introduction

The vascular endothelial growth factor A (VEGF), also known as vascular permeability factor, is a mitogen with high specificity for vascular endothelial cells. Its biological activities range from angiogenesis [1, 2], endothelial cell growth [3], and expression of anti-apoptotic proteins like Bcl-2 in endothelial cells [4] to induction of vascular leakage [5] by activation of MMP activity [6] . Thus, VEGF is crucial in various physiological and pathological states such as the formation of new blood vessels, wound healing, and tumor growth (for reviews see Refs. [7] and [8]).

The human *vegf* gene is organized in eight exons which can be spliced alternatively into four major isoforms ($VEGF_{121}$, $VEGF_{165}$, $VEGF_{189}$ and $VEGF_{206}$) [9, 10] and into two minor isoforms ($VEGF_{145}$ and $VEGF_{183}$) [7]. With increasing length of the VEGF isoform, the ability to bind to heparin also increases. $VEGF_{189}$ and $VEGF_{206}$ are highly basic and almost completely bound to the extracellular matrix. $VEGF_{121}$ is acidic, does not bind heparin and thus is freely diffusible [11]. $VEGF_{165}$ has intermediate properties regarding its secretion. As the mitogenic potency of VEGF decreases upon loss of the heparin binding region, $VEGF_{165}$ is twice as active as $VEGF_{121}$ [12] and thus has optimal bioavailability and biological activity [13–15]. It has been demonstrated that translation of the VEGF mRNA may not only be initiated at its presumed start codon at nucleotide 1039 (AUG^{1039}) but also at an in–frame CUG codon at nucleotide 499 (CUG^{499}). Utilization of CUG^{499} has been described for $VEGF_{121}$, $VEGF_{165}$, and $VEGF_{189}$. Although the exact extent of CUG^{499} usage differs between reports [16, 17], one can assume that transcription of $VEGF_{121}$ and $VEGF_{165}$ starts at CUG^{499} in more than 50% of the cases.

VEGF is in the spotlight of cancer therapy as its presence is crucial for neoangiogenesis. A tumor growing beyond a volume of 1-2 mm^3 requires its own blood vessels for a sufficient supply of nutrition and oxygen [18]. Therefore several strategies have been developed to hamper tumor vascularization by tackling VEGF signaling or levels of secreted VEGF. Several small molecules that interfere with VEGF signaling are being incorporated at present in clinical phase II trials. VEGF binds to two receptor tyrosine kinases (VEGFR-1 and VEGFR-2) present at the surface of most blood endothelial cells. VEGF binding leads to receptor dimerization, subsequent autophosphorylation, and activation of intracellular signaling cascades [19]. VEGFR-1 expressing cells can be selectively killed by a recombinant fusion construct of $VEGF_{121}$ with the toxin gelonin [20]. VEGFR-2, the major mediator of the mitogenic, angiogenic, and permeability enhancing effects of VEGF [7], can be blocked by several kinase inhibitors such as sunitinib, sorafenib, [21, 22] KRN951 [23] and AMG706 [24]. In order to prevent VEGF from binding to its receptors, the monoclonal antibody bevacizumab has been developed. Bevacizumab is a humanized neutralizing VEGF antibody [25] that binds all biologically active forms of VEGF. In 2004 it was approved

by the FDA e.g. for treatment of metastatic colorectal cancer. To summarize, various anti-VEGF based cancer therapies have been developed, several of which have a substantial clinical impact.

7.3. Materials and Methods

7.3.1. Materials

The HPLC reagents, trifluoroacetic acid, acetonitrile, formic acid, and HPLC water were purchased from Merck (Darmstadt, Germany). Peptide modification reagent O-Methyl isourea hemisulfate was purchased from Acros Organics (Geel, Belgium), 1-([1H4/2D4] nicotinoyloxy)succinimide (light or heavy dNIC-NHS) was synthesized as described elsewhere [26].

7.3.2. Elution of HLA presented peptides

HLA presented peptides were obtained by immune precipitation of HLA molecules from solid tissues using an adapted protocol [27]. One volume of 2x lysis buffer containing PBS, 0.6% 3-[(3-Cholamidopropyl)-dimethylammonio]-2-hydroxy-propanesulfonate (CHAPS) and complete protease inhibitor (Roche, Mannheim, Germany) was added to shock frozen tissue samples. Subsequently, the samples were homogenized using a blender and afterwards by a potter. One volume of 1x lysis buffer was added and the lysate was stirred for 1 h at 4℃ÀAfter 4x 30 seconds of sonication, the sample was centrifuged at 3,000 g, 4℃ for 20 min and afterwards at 150,000 g, 4℃ for 1 h to remove cell debris. Finally, the sample was passed through a 0.2 μm filter (Sartorius, Göttingen, Germany). For immune precipitation this lysate was applied for at least 12 h to a CNBr-activated sepharose 4B column (GE Healthcare, Freiburg, Germany; 40 mg sepharose / mg tissue) to which the HLA-A, -B and -C specific antibody W6/32 had been coupled (1 mg antibody / mg tissue) as described by the manufacturer. After binding, the column was rinsed with 250 ml PBS and subsequently with 500 ml ddH_2O. For elution and dissociation of bound HLA-peptide complexes, the column was shaken at least four times in one bed volume of 0.1% TFA for 20 min at room temperature. The four TFA-eluates were subsequently collected and combined. Finally, the HLA presented peptides were isolated by ultrafiltration through a centricon 10 kD cut-off membrane (Millipore, Schwalbach, Germany). For LC-MS analysis samples were freeze-dried and resuspended in 0.1% formic acid or 0.1% TFA for modification.

7.3.3. Peptide modification and analysis

Modification of peptides was carried out as follows. Isolated and freeze-dried HLA ligands were resuspended in 500 μl 0.1% TFA. Lysine side chains of peptides were guanidinated at pH 11 at 65°C for 10 min using 90 μl 2.5 M O-methyl isourea hemisulfate. The reaction was halted adjusting the pH to 3 with formic acid and peptides were desalted using Peptide Cleanup C18 Spin Tubes (Agilent, Palo Alto, CA, USA) as described in the manual. Peptides were nicotinylated on column for 15 min at room temperature using a 20 mM light or heavy dNIC-NHS solution in 50 mM phosphate buffer pH = 8.5. Aminolysis of undesired Tyrosine-dNIC esters was carried out by treatment with 500 μl 50% hydroxylamine for 10 min at room temperature. Subsequently, peptides were eluted using four times 50 μl of 50% acetonitrile, 1% formic acid. Peptide analysis was carried out as described elsewhere [26] using an Ultimate HPLC system (Dionex, Sunnyvale, CA, USA) with a gradient ranging from 15-55% solvent B within 170 minutes. For quantification, tumor and normal samples were mixed in a total peptide ratio of 1:1 and recorded in a single LC-MS experiment without fragmentation using a hybrid quadrupol orthogonal acceleration time of flight MS/MS (Q-TOF, Micromass, Manchester, UK), equipped with a micro-ESI source. For sequence analysis tumor and normal samples were analysed separately in individual LC-MS/MS experiments. Interpretation of MS/MS fragmentation spectra was done as described elsewhere in detail [28].

7.3.4. Database searches

BLAST searches for SRFGGAVVR peptide sequence were performed using the PAM30 matrix and an expect value of less than 1000 (http://www.ncbi.nlm.nih.gov/BLAST/). Protein versus protein searches (pblast) were done in 245,584 sequences deposited in the swissprot database; all species were included in the search. Translated protein versus nucleotide searches (tblastn) were done for all species using the NCBI non-redundant database (4,934,654 sequences), EST sequences were excluded.

7.3.5. Histology and ISH

Cells expressing VEGF mRNA were detected by radioactive in situ hybridization using single-stranded ^{35}S-labeled RNA probes which were synthesized from human $VEGF_{121}$ cloned into pBluescript II SK (Stratage, La Jolla, CA, USA). This pBSK-hVEGF121 plasmid was generously provided by Georg Breier (Institute of Pathology, Dresden, Germany). Pretreatment, hybridization, and washing procedures of 5 μm paraffin tissue sections were performed as described previously [29]. Slide preparations were subjected to autoradiography, exposed for 3 weeks at 4°C and counterstained with hematoxylin-eosin.

7.3.6. RT-PCR and qRT-PCR

RNA from cells was isolated using TRIzol reagent (Invitrogen, Karlsruhe, Germany) according to the manufacturer's recommendations. cDNA was synthesized from 1 μg total RNA as described elsewhere [30]. Real-time quantitative PCR (qRT-PCR) was performed using the ABI PRISM 7000 Sequence Detection System (Applied Biosystems, Darmstadt, Germany). SYBR Green PCR Master Mix (Applied Biosystems, Darmstadt, Germany) was used for PCR amplification and real-time detection of PCR products. Primers (biomers.net, Ulm, Germany) specific for $VEGF_{121}$ and $VEGF_{165}$ were designed to have a melting temperature of 60°C (forward primer $VEGF_{121}$: 5'-AACATCACCATGCAGATTA TGC-3', reverse primer $VEGF_{121}$: 5'-GGCTTGTCACATTTTTCTTGTC-3', forward primer $VEGF_{165}$: 5'-GTGAATGCAGACCAAAGAAAG-3', reverse primer $VEGF_{165}$: 5'-TTTTTGCAGGAACATTTACACG-3'). 18S RNA was chosen as a reference for normalizations (forward primer: 5'-CGGCTACCACATCCAA GGAA-3', reverse primer: 5'-GCTGGAATTACCGCGGCT-3'). PCR reactions were carried out in 20 μl with 400 nM of each primer. All samples were amplified in duplicate. Formation of undesired side products during PCR that contribute to fluorescence was excluded by melting curve analysis after PCR. Expression differences between tumor and normal tissue samples for different genes were calculated from PCR amplification curves by relative quantification using the comparative threshold cycle (C_T) method (http://docs.appliedbiosystems.com/pebiodocs/ 04303859.pdf). For absolute quantification standard curves of VEGF121 and VEGF165 were prepared using plasmids coding for the respective VEGF. Linear ranges of standard curves ranged from 10 fg to 10 ng of VEGF coding DNA. In order to calculate absolute VEGF copy numbers per cell, C_T values of the VEGF standard curve were allocated to C_T values of 18S RNA isolated from 2000 JY-BLCL cells.

7.3.7. Gene expression analysis by high-density oligonucleotide microarrays

RNA isolation from tumor and autologous normal kidney specimens as well as gene expression analysis by Affymetrix Human Genome U133 Plus 2.0 oligonucleotide microarrays (Affymetrix, Santa Clara, CA, USA) were performed as described previously [31]. Data were analysed with the GeneChip Operating Software (GCOS, Affymetrix, Santa Clara, CA, USA). Pairwise comparisons between tumor and autologous normal kidney were calculated using the respective normal array as baseline. Microarray data are available from the Gene Expression Omnibus repository (www.ncbi.nih.gov/geo) with the accession no. GSE8050.

7.3.8. VEGF protein quantification

Prior to the VEGF protein quantification, anti-VEGF beads were generated by covalent coupling of the anti-VEGF capture antibody to xMAP COOH-Microspheres (Luminex, Austin, TX, USA). To do so, we used a standard carbodiimide/ succinimide-based activation procedure where the bead carboxyls formed an acylamino ester that reacted with the primary amines of the antibody, yielding a stable amide bond. VEGF protein quantification was done as described elsewhere [32]. In brief 40 to 80 mg of tissue or $1 \cdot 10^7$ X63-B*2705 cells were pulverized in a Mikro-Dismembrator U (Braun/Satorius, Melsungen, Germany) at full speed for 1 min Proteins were extracted with a 10-fold excess of solubilization buffer: 50 mM Tris pH 7.5 / 400 mM NaCl / 1 mM $CaCl_2$ / 1 mM $MgCl_2$ / 1% Triton X-100 / 1x complete protease inhibitor (Roche Diagnostics, Mannheim, Germany). After 1 h incubation on ice, cellular debris was removed by centrifugation at 16,000 g for 5 min The total amount of solubilized protein was determined by Bradford assay. 10 μg of total solubilized protein were diluted in 30 μl assay buffer (ELISA blocking reagent, Roche Diagnostics, Mannheim, Germany) and mixed with 30 μl bead suspension in a 96-well filter plate. After an overnight incubation at 4°C in a plate shaker, unbound proteins were washed out and beads were resuspended in 30 μl of biotinylated detection antibody (R&D Systems, Wiesbaden, Germany). Excessive antibodies were removed after 2 h incubation at room temperature and beads were washed twice. Bound biotinylated antibodies were detected by R-phycoerythrin labeled streptavidin (S-PE). Before the beads were analyzed in a Luminex100 IS (Luminex, Austin, TX, USA), they were washed twice and finally resuspended in 100 μl assay buffer. At least 100 beads were analyzed and the median fluorescence intensities were taken as the read out signals. As a standard, a dilution series of recombinant human VEGF165 (R&D Systems, Wiesbaden, Germany) from 0.3 to 300 pg/ml was used.

7.3.9. Transfection of eukaryotic cells

The mouse myeloma cell line X63-B2705 (X63-Ag8.653 transfected with human HLA-B*2705 and human beta-2-microglobulin), generously provided by Elisabeth Weiss [33], was maintained in RPMI 1640 (C.C.Pro, Neustadt, Germany) medium containing 10% FCS (Pan, Aidenbach, Germany) and supplemented with 1 mg/ml G418S (Biochrom AG, Berlin, Germany). In the case of VEGF transfected cell clones 0.4 mg/ml hygromycin (Roche, Mannheim, Germany) was added. Stable VEGF transfectants were generated by electroporation of X63-B2705 cells (260 V, 925 μF; Gene Pulser II, Bio-Rad, München, Germany) followed by cloning using the limiting dilution method. For transfection a VEGF construct containing the complete mRNA of VEGF165 including the 'untranslated' region was used (VEGF-UTR165). Furthermore a VEGF165 construct containing only the protein coding nucleotides from ATG^{1039} was used (VEGF-

165). All VEGF constructs used for transfection experiments were generously provided by Ben-Zion Levi (Technion Haifa, Israel).

7.3.10. Peptides, recombinant HLA molecules and fluorescent tetramers

Peptides derived from VEGF (SRFGGAVVR) and HCMV pp65 (NLVPMVATV) were synthesized by standard Fmoc chemistry using an Economy Peptide Synthesizer EPS 221 (ABIMED, Langen, Germany). Peptides were checked by MALDI and HPLC for identity and purity. Afterwards they were dissolved at 10 mg/ml in DMSO (Merck, Darmstadt, Germany), diluted 1:10 in ddH_2O, and aliquots stored at –20°CḂiotinylated recombinant HLA class I molecules and fluorescent HLA tetramers for $CD8^+$ T cell analysis were produced as described earlier [34]. Briefly, fluorescent tetramers were generated by co-incubating biotinylated HLA monomers with streptavidin-PE or streptavidin-APC (Molecular Probes, Leiden, the Netherlands) at a 4:1 molar ratio. The HLA-B*2705 tetramers used had a cysteine-to-serine mutation in position 67.

7.3.11. *In vitro* stimulation of human $CD8^+$T cells using dendritic cells

PBMCs from healthy blood donors were isolated by standard Ficoll (PAA Laboratories, Cölbe, Germany) centrifugation. PBMCs ($12x10^7$ cells/10 ml per flask) were incubated in tissue culture flasks (CELLSTAR, Greiner Bio-One GmbH, Frickenhausen, Germany) with serum-free X-VIVO 20 medium (Lonza, Verviers, Belgium). After 2 hours of incubation at 37°C non-adherent cells were removed and adherent cells (12%-19% of the incubated cells) were cultured in T cell culture medium supplemented with 20 ng/ml IL-4 (R&D Systems, Wiesbaden, Germany), 100 ng/ml human recombinant GM-CSF (Leukine, Berlex Laboratories, Richmond, USA). T cell culture medium consisted of RPMI 1640, 25 mM HEPES, 2 mM L-glutamine (all from Invitrogen, Karlsruhe, Germany), 10% heat-inactivated human AB serum (CC pro, Neustadt/W., Germany), 50 U/ml penicillin, 50 μg/ml streptomycin and 20 μg/ml gentamycin (all from Lonza, Verviers, Belgium). Cultures were fed with fresh medium and cytokines every 2 to 3 days. After 7 days, DCs were activated with 10 ng/ml human TNF-α (R&D Systems, Wiesbaden, Germany) and 1 μg/ml human PGE-2 (Sigma-Aldrich, Taufkirchen, Germany). Mature DCs were predominantly $CD14^{high}$, $CD40^{high}$, $CD80^{high}$, $CD83^{high}$, $CD86^{high}$ and HLA-DR^{high} (data not shown). DCs were pulsed for 2 hours with 25 μg/ml of SRFGGAVVR peptide in T cell culture medium and washed 3 times with PBS prior to subsequent usage. *In vitro* stimulations were initiated in 24-well plates with $25x10^5$ responder cells (autologous PBMCs, *in vitro* cultured for one week) plus $5x10^5$ DCs per well in 1.5 ml

of T cell culture medium. Cryopreserved, peptide–loaded irradiated autologous PBMC were used for restimulation after priming with DCs.

7.3.12. *In vitro* stimulation of human CD8$^+$ T cells using artificial antigen presenting cells (aAPCs)

Untouched CD8$^+$ T cells were MACS enriched by negative depletion of PBMCs (Miltenyi Biotec, Bergisch Gladbach, Germany). Stimulations were initiated in 96-well plates with 1x10^6 responder cells in 250 μl of T cell culture medium complemented with 5 ng/ml human IL-12 (PromoKine, Heidelberg, Germany) plus 2x10^5 microbeads coated with HLA-B*2705-SRFGGAVVR-complexes, HLA-A*0201-NLVPMVATV-complexes plus anti-CD28 and anti-4-1BB antibodies as costimulatory signals [35]. After 3 to 4 days of incubation at 37°C fresh medium with 80 U/ml human IL-2 (Chiron, Emeryville, CA) was added and cells were incubated for another 3 to 4 days [36]. This stimulation cycle was performed two or three times.

7.3.13. Tetramer staining

Tetramer analyses were performed with SRFGGAVVR-tetramer-APC, NLVPMVATV-tetramer-PE, and antibody CD8-PerCP clone SK1 (BD Biosciences, Heidelberg, Germany). Cells were incubated with the antibody at 4°C for 20 min in the dark, followed by 30 min incubation with fluorescent HLA tetramers in the dark at room temperature. After resuspending with PBS containing 1% paraformaldehyde, 0.5% BSA, 2mM EDTA and 0.02% NaN_3, cells were analyzed by flow cytometry on a four-color FACSCalibur cytometer (BD Biosciences, Heidelberg, Germany).

7.3.14. Intracellular IFN-γ staining

After two and three rounds of stimulation, autologous PBMCs were thawed, washed twice in X-VIVO 20 medium (Lonza, Verviers, Belgium), resuspended at 1x10^7 cells/ml in T cell culture medium, and cultured overnight. On the next day, PBMCs pulsed with 10 μg/ml peptide were incubated with effector cells in a ratio of 1:1 for 6 hours. Golgi-Stop (BD Biosciences, Heidelberg, Germany) was added for the final 5 hours of incubation. Cells were analyzed using a Cytofix/Cytoperm Plus kit (BD Biosciences, Heidelberg, Germany) and CD8-PerCP clone SK1 (BD Biosciences, Heidelberg, Germany), IFN-γ-PE (BD Biosciences, Heidelberg, Germany). For negative controls, cells were incubated either with irrelevant peptide or without peptide, respectively. Stimulation with phorbol-12-myristate-13-acetate/Ionomycin was used as positive control. Cells were analyzed on a four-color FACSCalibur (BD Biosciences, Heidelberg, Germany).

7.3.15. ^{51}Cr-release assay

For the cytolytic assays, $CD8^+$ T cells were stained with a SRFGGAVVR-PE tetramer. Tetramer-positive cells were incubated with anti-PE MACS beads (Miltenyi Biotec, Bergisch Gladbach, Germany) and then sorted using magnetic cell separation according to the manufacturer's recommendations. Cells were subsequently cultured in the presence of $5x10^5$ cells/ml irradiated fresh allogenic PBMCs, $5x10^4$ cells/ml irradiated LG2-EBV cells, 150 U/ml IL-2, and 0.5 μg/ml PHA-L (Roche Diagnostics, Mannheim, Germany). Cells were further expanded in T cell culture medium containing 150 U/ml IL-2. Target cells were labeled with 100 μCi Na_2 $^{51}CrO_4$ (Hartmann Analytic, Braunschweig, Germany). Per well, 10000 target cells were used at different effector-to-target cell ratios in a standard 6 h ^{51}Chromium release assay. Spontaneous (target cells in medium alone) and maximal (target cells in 1% Triton X-100) ^{51}Cr release was determined and % specific lysis was calculated as follows: (cpm experimental release – cpm spontaneous release)/(cpm total release – cpm spontaneous release) x 100. Varying between different cell types, spontaneous releases ranged from 800 to 1.400 cpm, maximal releases were in the range between 8.000 and 15.000 cpm.

7.4. Results and Discussion

7.4.1. A cryptic VEGF derived HLA-ligand identified by mass spectrometry

Renal cell carcinoma of the clear cell type (RCC) is a frequent genitourinary tumor which is very resistant to standard therapy. In order to discover potential targets for T cell based tumor immunotherapy, we have analyzed the HLA ligandome of several RCCs and their autologous normal kidney tissue in detail [28, 31]. Therefore, we isolated HLA-presented peptides from tumor and normal tissue, labeled the peptides with a stable isotope tag [26], and subjected them to liquid chromatography coupled to an electrospray ionization mass spectrometer (ESI-LC-MS): All normal tissue derived peptides were labeled N-terminally with nicotinic acid (H_4-NIC); all peptides derived from tumor tissue were labeled in exactly the same manner but with a nicotinic acid bearing four deuterium atoms (D_4-NIC). The introduced isotope labels allowed – due to the isotopic mass shift – an accurate allocation of peptide signals to tumor or normal tissue after mixing of the two samples in a total peptide ratio of 1:1. Comparing the signal intensity of a heavy D_4-NIC peptide with its light H_4-NIC counterpart allowed peptide quantification between tumor and autologous normal tissue in one ESI-LC-MS experiment [28]. In such an analysis of RCC099 a very prominent tumor derived peptide with a mass to charge ratio (m/z) of 529.3 attracted attention (Fig. 7.1-A).

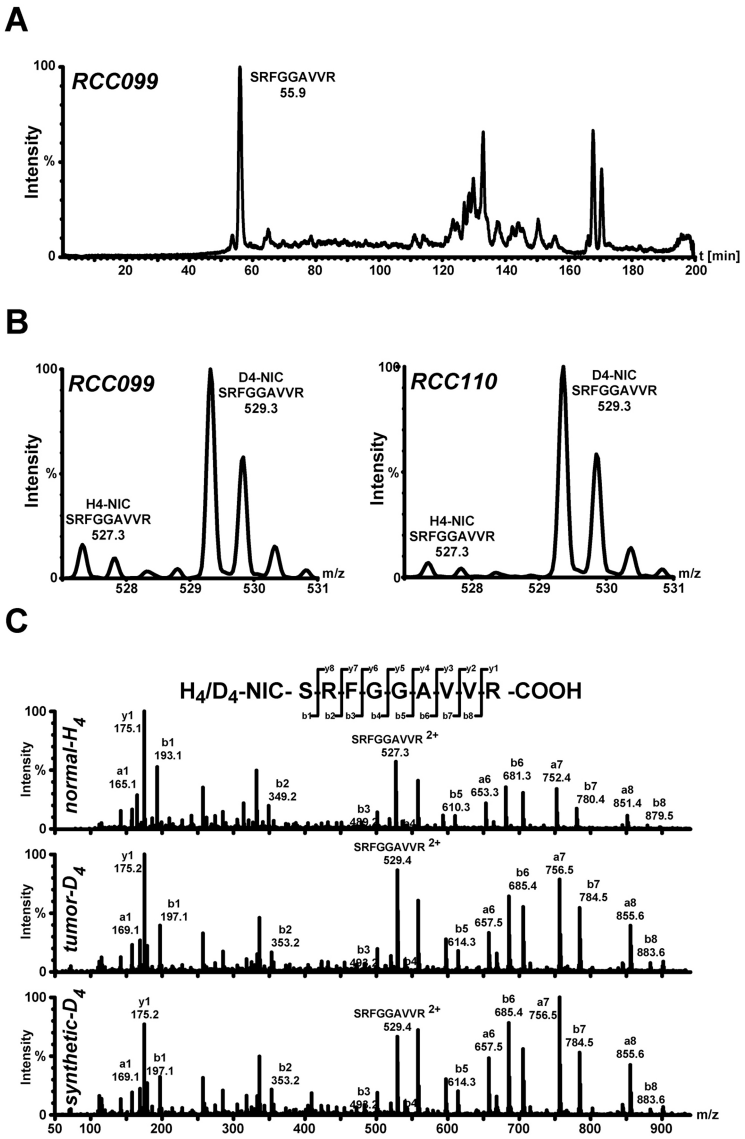

Figure 7.1.: The SRFGGAVVR peptide is over-presented on HLA molecules of RCC099 and RCC110. HLA-presented peptides isolated from RCC099 and RCC110 were analyzed by mass spectrometry. For analysis samples were differentially labeled with an isotope tag. Tumor samples were modified with deuterated nicotinic acid (D_4-NIC), normal tissue samples with normal nicotinic acid (H_4-NIC). A, Mass chromatogram of m/z 529.3 corresponding to D_4-NIC-SRFGGAVVR. B, Ratio analysis of SRFGGAVVR presented on tumor (D_4-NIC-SRFGGAVVR, m/z = 529.3) and normal tissue (H_4-NIC-SRFGGAVVR, m/z = 527.3). After normalization to the average presentation ratio, SRFGGAVVR appeared over-presented by 6.5-fold on RCC099 and 5.4-fold on RCC110. C, Fragment spectra of differentially nicotinylated SRFGGAVVR derived from RCC099 tumor and normal tissue, respectively synthetic D_4-NIC-SRFGGAVVR.

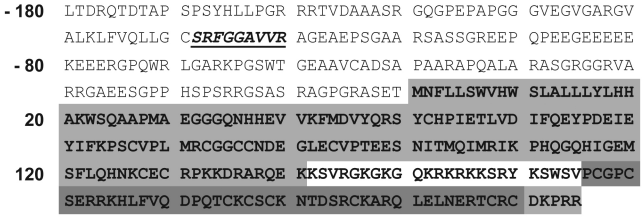

Figure 7.2.: SRFGGAVVR is located in the 'untranslated' region of VEGF. The translated protein sequence of VEGF mRNA (NCBI RefSeq Database ID NM_003376) starting at CUG^{499} is shown. VEGF protein sequence as deposited in swissprot (primary accession number P15692) is highlighted in bold. $VEGF_{121}$ is composed of amino acids shaded in light grey. The protein sequence of $VEGF_{165}$ contains additionally the amino acids shaded in dark grey. In frame SRFGGAVVR is highlighted in bold italics and is underlined.

After normalization to the average peptide presentation ratio [28], the quantitative analysis of this peptide showed a more than sixfold over-presentation of the peptide in tumor (D_4-NIC) compared to normal tissue (H_4-NIC; Fig. 7.1-b, left panel). In RCC110 the same peptide appeared more than fivefold over-presented on tumor tissue (Fig. 7.1-B, right panel). The corresponding peptide sequence (SRFGGAVVR) was identified by *de novo* sequencing using tandem mass spectrometry (MS/MS; Fig. 7.1-c). Furthermore, SRFGGAVVR fragmentation spectra from RCC samples were identical with fragmentation spectra of synthetic SRFGGAVVR (Fig. 7.1-C, lower panel). The peptide matched the HLA motif of HLA-B*2705 (www.syfpeithi.de, score = 30) for which both RCC099 and RCC110 were positive.

In order to identify the source protein from which the peptide SRFGGAVVR was derived, BLAST searches were performed against protein and translated nucleotide databases. Using the swissprot protein database, no protein in any species was found containing the peptide SRFGGAVVR. However, searching in a translated nucleotide database, the SRFGGAVVR peptide was found in the 'untranslated' region of VEGF in humans, chimpanzee, and rhesus monkey. SRFGGAVVR is in frame with the actual VEGF protein sequence and can be transcribed using the non-classical CUG^{499} [16, 17] as start codon, thus being cryptic (Fig. 7.2). Furthermore, the C-terminus of the peptide is predicted to be generated very efficiently by the proteasome (netchop 3.0 software http://www.cbs.dtu.dk/services/NetChop, score = 0.97). Integrated bioinformatical approaches, assessing the whole antigen processing pathway, rank SRFGGAVVR at least in the top 2% of all HLA-B*27 presented peptides of VEGF (WAPP score MHC 1.05, TAP -48.58, proteasome 3.31, http://www-bs.informatik.uni-tuebingen.de/WAPP ; NetCTL score 1.68, http://www.cbs.dtu.dk/services/NetCTL/; EpiJen score 1.498, http://www.jenner.ac.uk/EpiJen/).

7.4.2. Analysis of VEGF expression by in situ hybridization

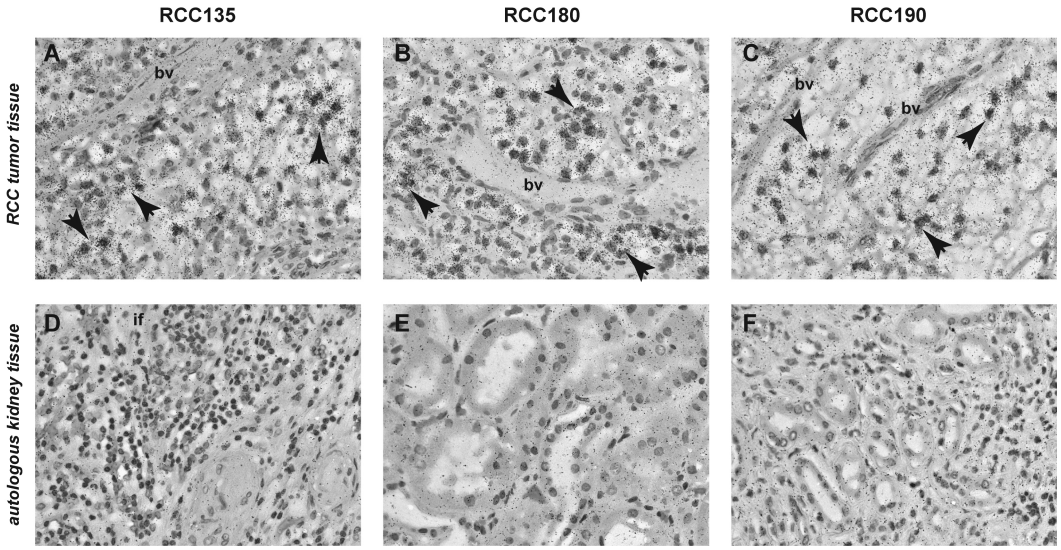

Figure 7.3.: VEGF mRNA is expressed in tumor but not in autologous normal tissue as revealed by radioactive *in situ* hybridization. RCC135, RCC180 and RCC190 tumor (A, B, C) and autologous normal tissue (D, E, F) sections were stained with hematoxylin-eosin. VEGF mRNA expressing cells were detected by *in situ* hybridized, radioactively labeled $VEGF_{121}$ RNA. VEGF positive cells are exemplarily marked with arrows. No VEGF mRNA expression is detectable in normal tissue and renal tubuli of healthy kidneys. bv, blood vessels; if, infiltrating cells.

In order to further characterize the cellular origin of VEGF leading to presentation of the SRFGGAVVR peptide in RCC tissue, *in situ* hybridization (ISH) experiments with radioactively labeled VEGF RNA probes were performed in RCC tumors and autologous normal tissues. Lacking appropriate paraffin embedded sample material for RCC099 and RCC110, this analysis was performed using three other clear cell RCCs (RCC135, RCC180 and RCC190) and their autologous normal tissues. In all analyzed samples VEGF transcripts were only detected in tumor cells (Fig. 7.3 A, B, C), thus being responsible for VEGF protein expression and subsequent presentation of the VEGF-derived peptide SRFGGAVVR. All autologous normal tissues (Fig. 7.3 D, E, F), inflammatory cells (Fig. 7.3 D, if) as well as endothelial cells of blood vessels within the tumor tissues (Fig. 7.3 A, B, C, bv) were negative for VEGF mRNA.

7.4.3. Expression profile of VEGF variants in RCCs

Table 7.1.: VEGF ratios of tumor vs. normal tissue are increased on various cellular levels. Protein, gene chip, and qRT-PCR refer to an average of $VEGF_{121}$ and $VEGF_{165}$ levels with potential minor contributions of $VEGF_{145}$ and $VEGF_{189}$;HLA-peptide ratio refers to presentation differences of the cryptic SRFGGAVVR peptide on HLA-B*2705.

	VEGF ratios tumor vs. Normal tissue			
sample	HLA peptide	protein	gene chip	qRT PCR
RCC099	6.5	4.5	2.0	1.3
RCC110	5.4	42.1	4.7	12.9

Gene chip analysis of RCC099 and RCC110 showed a 2-4-fold over-expression of VEGF in tumor tissue (Tab. 7.1). However, the gene chip probe set detecting VEGF did not allow determination of the variant of VEGF was expressed in tumor, respectively normal tissue. Similarily, the ISH probe allowed no specific detection of any VEGF variant because it had a minimum of 90% sequence overlap with each VEGF variant. A RT-PCR was therefore set up using primers covering the VEGF sequence from the classical ATG^{1039} start codon to the TGA^{1686} stop codon. $VEGF_{121}$ (459 nt) and $VEGF_{165}$ (591 nt) could subsequently be identified as the major transcript variants of VEGF in all analyzed samples (Fig. 7.4-A). Moreover, minor amounts of $VEGF_{145}$ (531 nt) and $VEGF_{189}$ (663 nt) could also be detected. In the same samples, VEGF amounts were analyzed by qRT-PCR in an absolute manner using defined amounts of reference plasmids coding for $VEGF_{121}$ and $VEGF_{165}$. Thus, $VEGF_{121}$ specific qRT-PCR primers were designed. In the case of $VEGF_{165}$, no specific qRT-PCR primer pair could be defined which did not also recognize $VEGF_{189}$. Due to the very low amount of $VEGF_{189}$ expression (Fig. 7.4-A), its contribution to $VEGF_{165}$ derived signals in qRT-PCR was put aside. $VEGF_{165}$ was expressed in RCC099 threefold higher amounts than in $VEGF_{121}$ and in RCC110 sevenfold (Fig. 7.4-B), respectively. Comparing tumor versus normal tissue, RCC099 showed only minor differences of mRNA expression levels for both VEGF variants. In contrast, RCC110 showed a minimum of 12-fold over-expression of VEGF in tumor tissue.

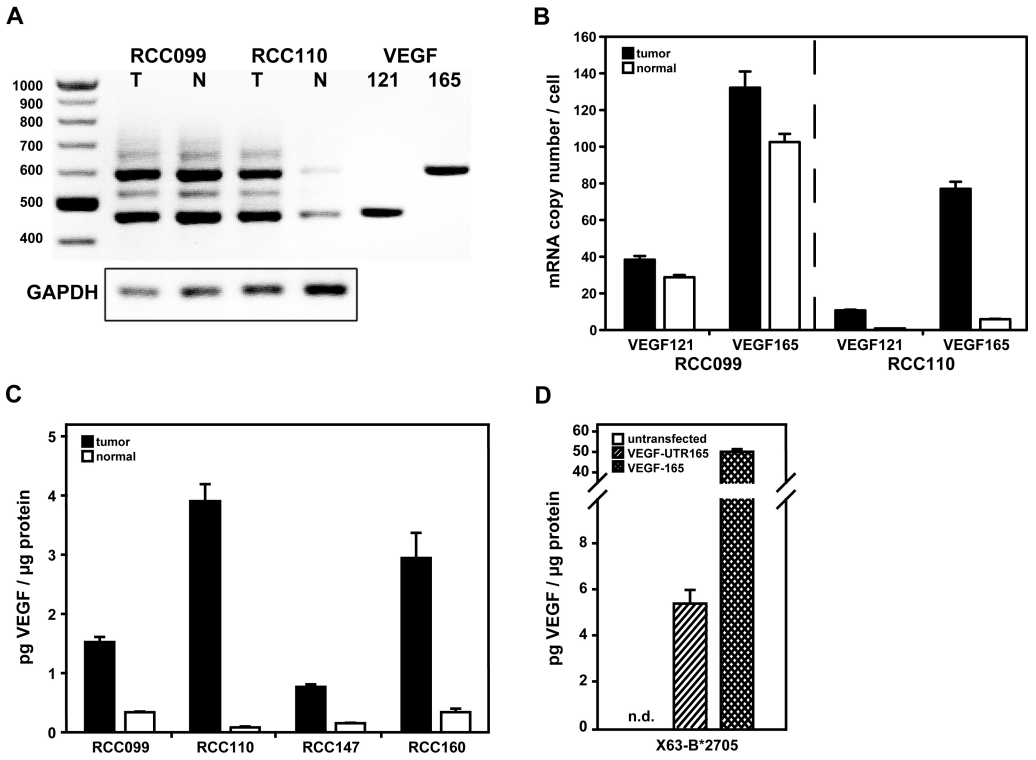

Figure 7.4.: VEGF is over-expressed in RCC099 and RCC110 both on mRNA and protein level. A, VEGF variants expressed in tumor and normal tissue were analyzed by RT-PCR by amplification of VEGF mRNA from nt 1021 to nt 1689. $VEGF_{121}$ and $VEGF_{165}$ coding plasmids were used as reference samples. $VEGF_{145}$ and $VEGF_{189}$ are observable as faint bands in RCC099 and RCC110 samples. B, Absolute quantification of $VEGF_{121}$ and $VEGF_{165}$ by qRT-PCR. Plasmids coding for $VEGF_{121}$ and $VEGF_{165}$ were used as standards. Samples were normalized to 18S RNA isolated from 2000 JY-BLCL cells. C and D, Protein quantification of VEGF by bead based sandwich immuno assay. C, Four tumor and autologous tissues were analyzed. D, Transfection efficiency of different VEGF constructs was determined in X63-B*2705 cells. In untransfected X63-B*2705 cells VEGF was not detectable. All error bars represent standard errors of the mean.

7.4.4. Quantitative analysis of VEGF protein levels in RCCs

Correlation of mRNA with protein levels [37] and HLA ligand levels [28] has been shown to be weak. Thus, protein levels of VEGF were assessed by a bead-based sandwich immunoassay [32] in RCC099 and RCC110 (Fig. 7.4-C, left panel). Additionally two other RCCs (RCC147 and RCC160), for which protein samples were available, were tested for VEGF protein expression. In all samples analyzed, VEGF protein expression was more than fourfold higher in tumor compared to normal tissue. Protein ratios in both RCC099 and RCC110 were about four times higher as their corresponding mRNA ratios. This is in agreement with reports showing enhanced VEGF translation under hypoxic conditions [38].

7.4.5. Assessment of VEGF on various cellular levels

We analyzed VEGF on mRNA, protein, and HLA ligand levels (Tab. 7.1). Expression of VEGF mRNA in RCC tumor tissue has been described in several reports [39, 40]. It is induced under hypoxic conditions mediated by the hypoxia inducible factor-1 (HIF-1). In RCCs in particular, it has been demonstrated that inactivation of the von Hippel-Lindau tumor suppressor gene leads to excessive VEGF production [41] due to constitutive HIF-1 [42] and HIF-2 [43] activation. In 16 gene chip experiments (data not shown) we compared the expression of VEGF in primary RCC tissue and autologous normal kidney tissue and found VEGF over-expression (more than 3-fold) in about 70% of all cases. In 25% of all cases, over-expression was greater than 9-fold. For RCC099 and RCC110, over-expression in tumor tissue was 2.0-fold respectively 4.7-fold. These gene chip data were confirmed by qRT-PCR. $VEGF_{121}$ and $VEGF_{165}$ were the predominant VEGF variants expressed in RCC, both being significantly over-expressed in tumor tissue. On the protein level, VEGF over-expression in tumor tissue was even more prominent. Here, differences ranged from 4- to 40-fold. Furthermore, the VEGF-derived peptide SRFGGAVVR was over-presented on HLA-molecules in both tumor tissues. The extent of HLA-peptide over-presentation did not match its corresponding mRNA-levels; this observation has been described previously for other gene products [28].

7.4.6. Generation of VEGF expressing cell lines

For subsequent T cell experiments, the mouse myeloma cell line X63-B*2705 – X63-Ag8 cells stably transfected with human HLA-B*2705 and β2-microglobulin – was used in order to circumvent contamination due to endogenous expression of human VEGF, which we could observe in human cell lines (see below). X63-B*2705 were stably transfected with two different VEGF constructs: VEGF-UTR165 was transfected with $VEGF_{165}$ containing the 'untranslated' region and thus also containing the SRFGGAVVR peptide. VEGF-165 was transfected with

$VEGF_{165}$ lacking the 'untranslated' region and thus the SRFGGAVVR peptide. Surface expression of HLA-B*2705 was confirmed by FACS analysis (data not shown), VEGF protein expression was tested as described above (Fig. 7.4-D).

7.4.7. Generation of VEGF specific T cells

The VEGF derived peptide SRFGGAVVR is presented by HLA-B*2705. In order to investigate whether the HLA-bound peptide can be recognized by functional cytotoxic T cells, $CD8^+$ T cells from healthy blood donors were stimulated *in vitro* for four weeks with dendritic cells (DCs) or artificial antigen presenting cells (aAPCs). All blood donors were HLA-A*0201 and -B*2705 positive as well as seropositive for human cytomegalovirus (HCMV). For stimulation experiments, DCs were loaded with SRFGGAVVR peptide; aAPCs were loaded with costimulatory antibodies, HLA-B*2705-SRFGGAVVR-complexes and HLA-A*0201-NLVPMVATV-complexes. The immunodominant HCMV derived peptide NLVPMVATV served as positive control. After four weeks of stimulation, cells were stained with fluorescence labeled B*2705-SRFGGAVVR- or A*0201-NLVPMVATV-tetramers and assayed in a FACS experiment. Both the HCMV derived peptide and the VEGF derived SRFGGAVVR were able to induce a specific T cell population (Fig. 7.5-A).

7.4.8. Functional analysis of VEGF-specific T cells

In order to assess the functionality of SRFGGAVVR-specific T cells, their ability to secrete cytokines and their cytolytic potential were analyzed. For these functional experiments only DC stimulations were used because it has been reported that strong and rapid T cell expansion using aAPCs bears the risk of inducing dysfunctional T cells [44]. Upon co-culture with autologous PBMCs loaded with SRFGGAVVR-peptide, $CD8^+$ T cells were able to secrete IFN-γ. This confirmed that the effector T cells were functional and specific for VEGF (Fig. 7.5-B). Furthermore, the cytolytic potential of these cells was analyzed in a chromium release assay. Therefore, *in vitro* stimulated SRFGGAVVR specific tetramer positive T cells were MACS-sorted and expanded. Autologous PBMCs loaded with an irrelevant peptide remained unlysed. Human C1R-B*2705 transfected with VEGF-UTR165 were lysed efficiently; however, the parental cell line C1R-B*2705 was lysed due to endogenous VEGF expression as well (data not shown). Mouse X63-B*2705 cells were recognized only after transfection with VEGF-UTR165 containing the cryptic SRFGGAVVR peptide sequence, whereas VEGF-165 transfected X63-B*2705 cells were not lysed – although VEGF-165 protein levels were 10-fold higher than in VEGF-UTR165 cells (Fig. 7.4-D). Finally, a cell line generated from RCC110 tumor tissue was specifically recognized by these T cells.

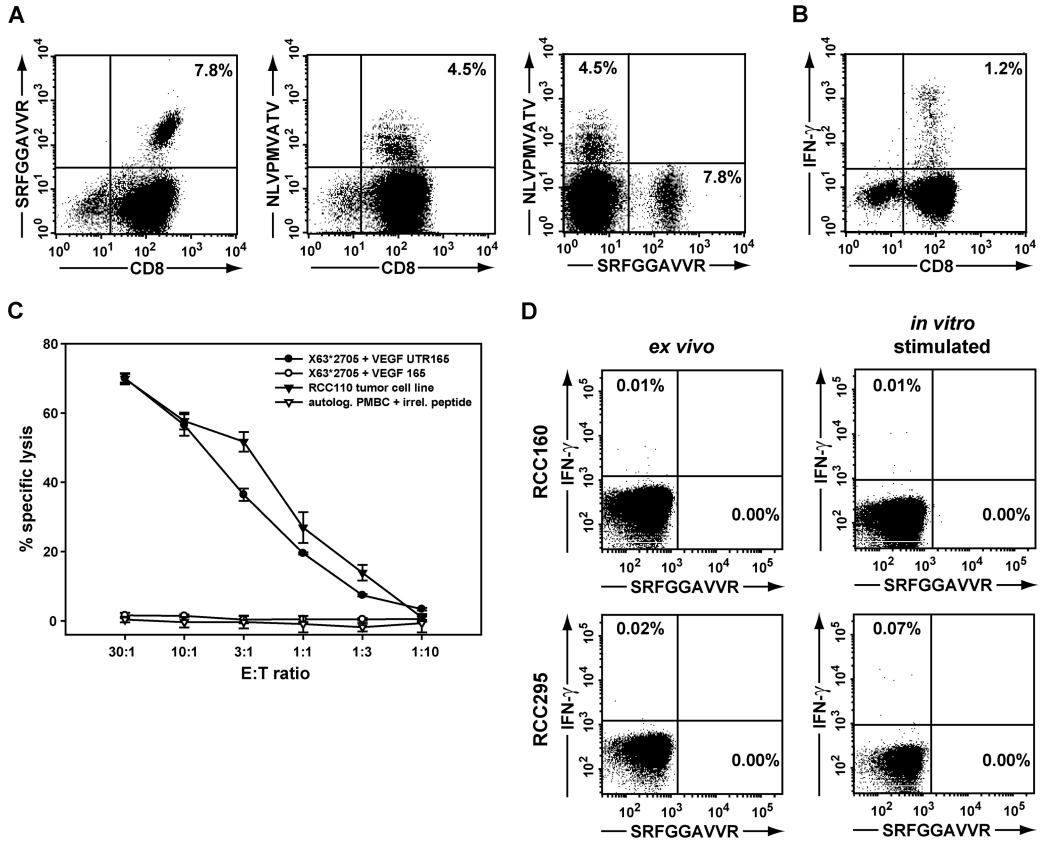

Figure 7.5.: SRFGGAVVR specific T cells are functional cytotoxic effector T cells. A, Tetramer staining of $CD8^+$ T cells that were stimulated with aAPCs loaded with HLA-A*0201-NLVPMVATV (HCMV), HLA-B*2705-SRFGGAVVR (VEGF) and with co-stimulating antibodies. The respective tetramer used for staining is indicated in each plot. Percentages are calculated from total $CD8^+$ cells. B and C, $CD8^+$ T cells, stimulated *in vitro* using dendritic cells, were used for functional assays. Stimulated T cells were tested for IFN-γ secretion upon co-culture with autologous PBMCs loaded with SRFGGAVVR. C, ^{51}Cr-cytotoxicity assay of SRFGGAVVR stimulated $CD8^+$ T cells. Two X63-B*2705 transfectants, a RCC110 tumor cell line as well as peptide loaded autologous PBMCs were used as target cells. Error bars represent standard deviations. D, Combined tetramer-IFN-γ analysis of PBMCs isolated from RCC160 and RCC294. Cells were stained *ex vivo* respectively after one week of *in vitro* stimulation with SRFGGAVVR peptide. Irrelevant peptide stimulations resulted in identical percentages regarding the IFN-γ staining. Stimulation with PMA/Ionomycin resulted in IFN-γ secretion of 70-90% of all $CD8^+$ T cells. Data are representative for PBMC samples from 10 other RCC patients.

7.4.9. Absence of VEGF specific T cells in blood of RCC patients

Using PBMCs from healthy blood donors, SRFGGAVVR specific T cells were inducible *in vitro*. In order to asses whether such SRFGGAVVR specific T cells were induced *in vivo* in RCC patients, we analyzed PBMC samples of 12 HLA-B*27 positive RCC patients by tetramer analysis and by intracellular IFN-γ staining. In none of the samples – including RCC110 – were tetramer positive or IFN-γ producing cells detectable, neither in an *ex vivo* analysis nor upon 7 days of *in vitro* peptide stimulation (Fig. 7.5-D). This suggests that SRFGGAVVR may not be spontaneously recognized by T cells from RCC patients and this may indicate a possibility for a therapeutic utilization of VEGF specific T cells.

7.5. Summary

Several anti-VEGF based tumor therapies have already been shown to have a substantial clinical impact. We have identified a cryptic HLA ligand derived from the 'untranslated' region of VEGF by mass spectrometry. The peptide is derived most probably from the VEGF variants $VEGF_{121}$ or $VEGF_{165}$, which are highly over-expressed on mRNA and on the protein level in RCC tumors (summarized in Tab. 7.1). Furthermore, the HLA-bound peptide SRFGGAVVR is highly over-presented on RCC tumor tissue and can be directly recognized by cytotoxic effector T cells. Therefore, our findings define VEGF as a T cell targetable tumor associated antigen. A peptide-based immunotherapeutic approach directed against HLA-presented VEGF epitopes may amend other anti-VEGF based therapies in the future.

7.6. Acknowledgements

We thank Lynne Yakes for critical reading of the manuscript, Tobias Krüger for establishing a cell line from RCC110 tumor tissue and Jörg Hennenlotter and Arnulf Stenzl for providing RCC sample material.

7.7. References

[1] D. W. Leung, G. Cachianes, W. J. Kuang, D. V. Goeddel, and N. Ferrara. Vascular endothelial growth factor is a secreted angiogenic mitogen. *Science*, 246(4935):1306–1309, 1989.

[2] J. Plouet, J. Schilling, and D. Gospodarowicz. Isolation and characterization of a newly identified endothelial cell mitogen produced by AtT-20 cells. *EMBO J.*, 8(12):3801–3806, 1989.

[3] N. Ferrara and T. Davis-Smyth. The biology of vascular endothelial growth factor. *Endocr. Rev.*, 18(1):4–25, 1997.

[4] H. P. Gerber, V. Dixit, and N. Ferrara. Vascular endothelial growth factor induces expression of the antiapoptotic proteins Bcl-2 and A1 in vascular endothelial cells. *J. Biol. Chem.*, 273(21):13313–13316, 1998.

[5] D. R. Senger, S. J. Galli, A. M. Dvorak, C. A. Perruzzi, V. S. Harvey, and H. F. Dvorak. Tumor cells secrete a vascular permeability factor that promotes accumulation of ascites fluid. *Science*, 219(4587):983–985, 1983.

[6] F. Q. Wang, J. So, S. Reierstad, and D. A. Fishman. Vascular endothelial growth factor-regulated ovarian cancer invasion and migration involves expression and activation of matrix metalloproteinases. *Int. J. Cancer*, 118(4): 879–888, 2006.

[7] N. Ferrara, H. P. Gerber, and J. LeCouter. The biology of VEGF and its receptors. *Nat. Med.*, 9(6):669–676, 2003.

[8] A. D. Thornton, P. Ravn, M. Winslet, and K. Chester. Angiogenesis inhibition with bevacizumab and the surgical management of colorectal cancer. *Br. J. Surg.*, 93(12):1456–1463, 2006.

[9] K. A. Houck, N. Ferrara, J. Winer, G. Cachianes, B. Li, and D. W. Leung. The vascular endothelial growth factor family: identification of a fourth molecular species and characterization of alternative splicing of RNA. *Mol. Endocrinol.*, 5(12):1806–1814, 1991.

[10] E. Tischer, R. Mitchell, T. Hartman, M. Silva, D. Gospodarowicz, J. C. Fiddes, and J. A. Abraham. The human gene for vascular endothelial growth factor. Multiple protein forms are encoded through alternative exon splicing. *J. Biol. Chem.*, 266(18):11947–11954, 1991.

[11] K. A. Houck, D. W. Leung, A. M. Rowland, J. Winer, and N. Ferrara. Dual regulation of vascular endothelial growth factor bioavailability by genetic and proteolytic mechanisms. *J. Biol. Chem.*, 267(36):26031–26037, 1992.

[12] B. A. Keyt, L. T. Berleau, H. V. Nguyen, H. Chen, H. Heinsohn, R. Vandlen, and N. Ferrara. The carboxyl-terminal domain (111-165) of vascular endothelial growth factor is critical for its mitogenic potency. *J. Biol. Chem.*, 271(13):7788–7795, 1996.

[13] P. Carmeliet, Y. S. Ng, D. Nuyens, G. Theilmeier, K. Brusselmans, I. Cornelissen, E. Ehler, V. V. Kakkar, I. Stalmans, V. Mattot, J. C. Perriard, M. Dewerchin, W. Flameng, A. Nagy, F. Lupu, L. Moons, D. Collen, P. A. D'Amore, and D. T. Shima. Impaired myocardial angiogenesis and ischemic

cardiomyopathy in mice lacking the vascular endothelial growth factor isoforms VEGF164 and VEGF188. *Nat. Med.*, 5(5):495–502, 1999.

[14] C. Ruhrberg, H. Gerhardt, M. Golding, R. Watson, S. Ioannidou, H. Fujisawa, C. Betsholtz, and D. T. Shima. Spatially restricted patterning cues provided by heparin-binding VEGF-A control blood vessel branching morphogenesis. *Genes Dev.*, 16(20):2684–2698, 2002.

[15] I. Stalmans, Y. S. Ng, R. Rohan, M. Fruttiger, A. Bouche, A. Yuce, H. Fujisawa, B. Hermans, M. Shani, S. Jansen, D. Hicklin, D. J. Anderson, T. Gardiner, H. P. Hammes, L. Moons, M. Dewerchin, D. Collen, P. Carmeliet, and P. A. D'Amore. Arteriolar and venular patterning in retinas of mice selectively expressing VEGF isoforms. *J. Clin. Invest*, 109(3):327–336, 2002.

[16] S. Bornes, M. Boulard, C. Hieblot, C. Zanibellato, J. S. Iacovoni, H. Prats, and C. Touriol. Control of the vascular endothelial growth factor internal ribosome entry site (IRES) activity and translation initiation by alternatively spliced coding sequences. *J. Biol. Chem.*, 279(18):18717–18726, 2004.

[17] M. K. Tee and R. B. Jaffe. A precursor form of vascular endothelial growth factor arises by initiation from an upstream in-frame CUG codon. *Biochem. J.*, 359(Pt 1):219–226, 2001.

[18] J. Folkman. Tumor angiogenesis: therapeutic implications. *N. Engl. J. Med.*, 285(21):1182–1186, 1971.

[19] T. Matsumoto and H. Mugishima. Signal transduction via vascular endothelial growth factor (VEGF) receptors and their roles in atherogenesis. *J. Atheroscler. Thromb.*, 13(3):130–135, 2006.

[20] K. A. Mohamedali, A. T. Poblenz, C. R. Sikes, N. M. Navone, P. E. Thorpe, B. G. Darnay, and M. G. Rosenblum. Inhibition of prostate tumor growth and bone remodeling by the vascular targeting agent VEGF121/rGel. *Cancer Res.*, 66(22):10919–10928, 2006.

[21] P. Nathan, D. Chao, C. Brock, P. Savage, M. Harries, M. Gore, and T. Eisen. The place of VEGF inhibition in the current management of renal cell carcinoma. *Br. J. Cancer*, 94(9):1217–1220, 2006.

[22] A. J. Schrader, Z. Varga, S. Pfoertner, U. Goelden, J. Buer, and R. Hofmann. Treatment targeted at vascular endothelial growth factor: a promising approach to managing metastatic kidney cancer. *BJU. Int.*, 97(3):461–465, 2006.

[23] K. Nakamura, E. Taguchi, T. Miura, A. Yamamoto, K. Takahashi, F. Bichat, N. Guilbaud, K. Hasegawa, K. Kubo, Y. Fujiwara, R. Suzuki, K. Kubo,

M. Shibuya, and T. Isae. KRN951, a highly potent inhibitor of vascular endothelial growth factor receptor tyrosine kinases, has antitumor activities and affects functional vascular properties. *Cancer Res.*, 66(18):9134–9142, 2006.

[24] A. Polverino, A. Coxon, C. Starnes, Z. Diaz, T. Demelfi, L. Wang, J. Bready, J. Estrada, R. Cattley, S. Kaufman, D. Chen, Y. Gan, G. Kumar, J. Meyer, S. Neervannan, G. Alva, J. Talvenheimo, S. Montestruque, A. Tasker, V. Patel, R. Radinsky, and R. Kendall. AMG 706, an Oral, Multikinase Inhibitor that Selectively Targets Vascular Endothelial Growth Factor, Platelet-Derived Growth Factor, and Kit Receptors, Potently Inhibits Angiogenesis and Induces Regression in Tumor Xenografts. *Cancer Res.*, 66 (17):8715–8721, 2006.

[25] L. G. Presta, H. Chen, S. J. O'Connor, V. Chisholm, Y. G. Meng, L. Krummen, M. Winkler, and N. Ferrara. Humanization of an anti-vascular endothelial growth factor monoclonal antibody for the therapy of solid tumors and other disorders. *Cancer Res.*, 57(20):4593–4599, 1997.

[26] C. Lemmel, S. Weik, U. Eberle, J. Dengjel, T. Kratt, H. D. Becker, H. G. Rammensee, and S. Stevanovic. Differential quantitative analysis of MHC ligands by mass spectrometry using stable isotope labeling. *Nat. Biotechnol.*, 2004.

[27] K. Falk, O. Rotzschke, S. Stevanovic, G. Jung, and H. G. Rammensee. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*, 351(6324):290–296, 1991.

[28] A. O. Weinzierl, C. Lemmel, O. Schoor, M. Müller, T. Krüger, D. Wernet, J. Hennenlotter, A. Stenzl, K. Klingel, H. G. Rammensee, and S. Stevanovic. Distorted relation between mRNA copy number and corresponding major histocompatibility complex ligand density on the cell surface. *Mol. Cell Proteomics.*, 6(1):102–113, 2007.

[29] K. Klingel, C. Hohenadl, A. Canu, M. Albrecht, M. Seemann, G. Mall, and R. Kandolf. Ongoing enterovirus-induced myocarditis is associated with persistent heart muscle infection: quantitative analysis of virus replication, tissue damage, and inflammation. *Proc. Natl. Acad. Sci. U. S. A*, 89(1): 314–318, 1992.

[30] O. Schoor, T. Weinschenk, J. Hennenlotter, S. Corvin, A. Stenzl, H. G. Rammensee, and S. Stevanovic. Moderate degradation does not preclude microarray analysis of small amounts of RNA. *Biotechniques*, 35(6):1192–1201, 2003.

[31] T. Krüger, O. Schoor, C. Lemmel, B. Kraemer, C. Reichle, J. Dengjel, T. Weinschenk, M. Muller, J. Hennenlotter, A. Stenzl, H. G. Rammensee, and S. Stevanovic. Lessons to be learned from primary renal cell carcinomas: novel tumor antigens and HLA ligands for immunotherapy. *Cancer Immunol. Immunother.*, 54(9):826–836, 2005.

[32] N. Schneiderhan-Marra, A. Kirn, A. Döttinger, M. Templin, G. Sauer, H. Deissler, and T. O. Joos. Protein Microarrays - A Promising Tool for Cancer Diagnosis. *Canc. Genom. Proteom.*, 2(1):37–42, 2005.

[33] M. Ulbrecht, T. Honka, S. Person, J. P. Johnson, and E. H. Weiss. The HLA-E gene encodes two differentially regulated transcripts and a cell surface protein. *J. Immunol.*, 149(9):2945–2953, 1992.

[34] J. D. Altman, P. A. Moss, P. J. Goulder, D. H. Barouch, M. G. McHeyzer-Williams, J. I. Bell, A. J. McMichael, and M. M. Davis. Phenotypic analysis of antigen-specific T lymphocytes. *Science*, 274(5284):94–96, 1996.

[35] D. Rudolf, T. Silberzahn, S. Walter, D. Maurer, J. Engelhard, D. Wernet, H. J. Buhring, G. Jung, B. S. Kwon, H. G. Rammensee, and S. Stevanovic. Potent costimulation of human CD8 T cells by anti-4-1BB and anti-CD28 on synthetic artificial antigen presenting cells. *Cancer Immunol. Immunother.*, 2007.

[36] S. Walter, L. Herrgen, O. Schoor, G. Jung, D. Wernet, H. J. Bühring, H. G. Rammensee, and S. Stevanovic. Cutting edge: predetermined avidity of human CD8 T cells expanded on calibrated MHC/anti-CD28-coated microspheres. *J. Immunol.*, 171(10):4974–4978, 2003.

[37] S. P. Gygi, Y. Rochon, B. R. Franza, and R. Aebersold. Correlation between protein and mRNA abundance in yeast. *Mol. Cell Biol.*, 19(3):1720–1730, 1999.

[38] I. Stein, A. Itin, P. Einat, R. Skaliter, Z. Grossman, and E. Keshet. Translation of vascular endothelial growth factor mRNA by internal ribosome entry: implications for translation under hypoxia. *Mol. Cell Biol.*, 18(6):3112–3119, 1998.

[39] J. Jacobsen, K. Grankvist, T. Rasmuson, A. Bergh, G. Landberg, and B. Ljungberg. Expression of vascular endothelial growth factor protein in human renal cell carcinoma. *BJU. Int.*, 93(3):297–302, 2004.

[40] D. Nicol, S. I. Hii, M. Walsh, B. Teh, L. Thompson, C. Kennett, and D. Gotley. Vascular endothelial growth factor expression is increased in renal cell carcinoma. *J. Urol.*, 157(4):1482–1486, 1997.

[41] G. Siemeister, K. Weindel, K. Mohrs, B. Barleon, G. Martiny-Baron, and D. Marme. Reversion of deregulated expression of vascular endothelial growth factor in human renal carcinoma cells by von Hippel-Lindau tumor suppressor protein. *Cancer Res.*, 56(10):2299–2301, 1996.

[42] P. H. Maxwell, M. S. Wiesener, G. W. Chang, S. C. Clifford, E. C. Vaux, M. E. Cockman, C. C. Wykoff, C. W. Pugh, E. R. Maher, and P. J. Ratcliffe. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature*, 399(6733):271–275, 1999.

[43] V. A. Carroll and M. Ashcroft. Role of hypoxia-inducible factor (HIF)-1alpha versus HIF-2alpha in the regulation of HIF target genes in response to hypoxia, insulin-like growth factor-I, or loss of von Hippel-Lindau function: implications for targeting the HIF pathway. *Cancer Res.*, 66(12):6264–6270, 2006.

[44] M. Montes, N. Rufer, V. Appay, S. Reynard, M. J. Pittet, D. E. Speiser, P. Guillaume, J. C. Cerottini, P. Romero, and S. Leyvraz. Optimum in vitro expansion of human antigen-specific CD8 T cells for adoptive transfer therapy. *Clin. Exp. Immunol.*, 142(2):292–302, 2005.

8. Summary

Molecularly defined approaches in tumour immunotherapy require the identification of appropriate tumour (associated) antigens. Therefore, a quantitative mass spectrometric approach based on the differential N-terminal isotope coding (dNIC) strategy was chosen for analysis of various HLA-bound peptide pools. In a first step, we could show that the systemic variations in the dNIC approach are low enough that biological differences in two samples can be detected reliably.

In a TAP-deficient cell line we have identified several over-presented peptides, which share some common features: Most of these peptides had a reduced binding affinity to MHC-I molecules and were derived from signal sequences. However, several of the TAP-independently presented peptides were presumably generated by the proteasome and transported into the ER *via* alternative pathways.

Using primary tumour tissue and autologous normal tissue, we were able to show that the correlation of over-presented peptides to their corresponding mRNA levels was very weak. This result emphasized the importance of quantitative MHC ligand analysis for a fast and reliable identification of tumour (associated) antigens.

Finally, we took an HLA ligand which was identified as being majorly over-presented in our tumour vs. normal tissue comparison and demonstrated that this VEGF-derived ligand can be used as a CTL epitope derived from the tumour associated antigen VEGF.

Concluding, the dNIC strategy is a powerful MS-based approach for the identification of differences in the HLA ligandome of two samples. This can be utilised for a broad spectrum of applications ranging from quantitative HLA ligandome comparisons to the identification of tumour associated CTL epitopes.

9. Zusammenfassung

Molekular definierte Ansätze in der Tumorimmuntherapie setzen die Identifikation geeigneter Tumor (assoziierter) Antigene voraus. Deshalb wurde ein massenspektrometrischer Ansatz, basierend auf der differentiellen N-terminalen Isotopencodierungs (dNIC) Strategie, gewählt um diverse Vergleiche von HLA-gebundenen Peptidpools durchzuführen. In einem ersten Schritt konnten wir zeigen, dass die systemischen Variationen der dNIC-Methode klein genug sind, um biologische Unterschiede zwischen zwei Proben verlässlich zu identifizieren.

In einer TAP-defizienten Zelllinie konnten wir mehrere überpräsentierte Peptide identifizieren, die einige Gemeinsamkeiten teilten: Die meisten dieser Peptide hatten eine verminderte Bindungsaffinität zu MHC-I Molekülen und kamen aus Signalsequenzen. Einige der TAP-unabhängig präsentierten Peptide wurden jedoch wahrscheinlich durch das Proteasom erzeugt und gelangten durch einen alternativen Weg in das ER.

Bei einem Vergleich von Tumor- und autologem Normalgewebe konnten wir zeigen, dass die Korrelation von überpräsentierten Peptiden mit ihren korrespondierenden mRNA-Verhältnissen nur sehr gering war. Dieses Ergebnis betont die Bedeutung der quantitativen MHC-Ligandenanalyse für eine schnelle und verlässliche Identifizierung von Tumor (assoziierten) Antigenen.

Zuletzt verwendeten wir einen HLA-Liganden, der in unserem Tumor versus Normalgewebsvergleich als besonders überpräsentiert identifiziert wurde. Für diesen Liganden aus VEGF konnten wir zeigen, dass er als CTL-Epitop des Tumor-assoziierten Antigens VEGF verwendet werden kann.

Schlussfolgernd betrachtet ist die dNIC-Strategie ein MS-basierter Ansatz mit großem Potential für die Identifizierung von Unterschieden im HLA-Ligandom von zwei Proben. Dies kann für ein breites Anwendungsspektrum eingesetzt werden, welches von den quantitativen HLA-Ligandomanalysen bis hin zur Identifizierung von Tumor-assoziierten Antigenen reicht.

10. Abbreviations

2D-GE	two-dimensional gel electrophoresis
AAF-CMK	Ala-Ala-Phe-chloromethylketone
ADFP	adipophilin
ALICE	acid-labile isotope-coded extractants
ANXA4	annexin A4
AQUA	absolute quantification
CCNI	cyclin I
CD24	small cell lung carcinoma cluster 4 antigen
CDK4	cyclin-dependent kinase 4
CEA	carcinoembryonic antigen
cICAT	cleavable isotope coded affinity tag
CML	chronic myeloid leukemia
CNBr	cyanogen bromide activated
CTL	cytotoxic T lymphocyte
DCs	dendritic cells
DDF	differential detergent fractionation
DIGE	difference gel electrophoresis
dNIC	differential N-terminal isotope coding
DRiPs	defective ribosomal products
EDT	ethanedithiol
EHD2	EH-domain containing protein 2
ERAAP	endoplasmic reticulum aminopeptidase associated with antigen processing
EST	expressed sequence tag
FARSLA	homologue to the phenylalanine-tRNA synthetase alpha unit
HHV	human herpes virus
HIF	hypoxia inducible factor
HLA	human leukocyte antigen
HLA-B*2705	human leukocyte antigen B*2705
HMOX1	decycling heme oxygenase 1
ICAT	isotope coded affinity tag
ICPL	isotope coded protein labels
ICROC	isotope-coded reduction off of a chromatographic support
IDO	indoleamine 2,3-dioxygenase
IFNγ	interferon gamma
IL	interleukin

LC	liquid chromatography
LCL721.174	lymphoblastoid B-cell line 721.174
LCL721.45	lymphoblastoid B-cell line 721.45
LC-MS	liquid chromatography - mass spectrometry
LMP2/7	proteasome subunit beta 9/8
MAGE	melanoma antigen
MAGE1	melanoma antigen 1
MALDI-TOF	matrix assisted laser desorption ionization - time of flight
MHC	major histocompatibility complex
MHC-I or II	major histocompatibility complex class I or II
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MUC1	mucin 1
NHS	nicotinoyloxysuccinimide
NIC	nicotinic acid
NIC-NHS	nicotinoyloxy-succinimide
OD	optical density
PhIAT	phosphoprotein isotope-coded affinity tags
PIGR	polymeric Ig receptor
PLXNB2	plexin B2
PSMA	prostate-specific membrane antigen
pVHL	von Hippel-Lindau gene product
qRT-PCR	quantitative real time PCR
RBBP4	retinoblastoma binding protein 4
RCC	renal cell carcinoma
SILAC	stabile isotope labeling by amino acids in cell culture
SPP	signal peptide peptidase
TAA	tumour associated antigen
TAP	transporter associated with antigen processing
TMAB	4-trimethyl-ammoniumbutyryl
TMED10	transmembrane trafficking protein 10
TPP2	tripeptidyl peptidase II
UGT1A6	UDP glucuronosyltransferase 1 family, polypeptide A6
UMOD	uromodulin
VEGF	vascular endothelial growth factor
ZF	zonal zentrifugation

11. Acknowledgements

Prof. Dr. Stefan Stevanović gilt ein besonderes Dankeschön dafür, dass er es geschafft hat, mir alle Freiräume zu gewähren, ohne dass ich mich verloren gefühlt hätte – und dafür, dass ich einen Teil meiner Zeit 'Artefakten' wie Mausmodellen widmen durfte.

Prof. Dr. Hans-Georg Rammensee danke ich für sein Interesse an meiner Arbeit und für die Art, wie er die Atmosphäre in der Abteilung prägt.

Claudia Lemmel, die Herrin... ohne Claudia gäbe es die HLA-Quantifizierung nicht und so hervorragend in die Auswertung von Spektren, die aus einer Ansammlung von 'Strichen' bestehen, hätte mich sonst wohl auch niemand einführen können. Vielen lieben Dank!

Vielen Dank an die (jetzt auch schon) alte MS-Crew Margret Müller, Nina Hillen, Verena Meyer und Florian Altenberend für psychische und physische Hilfe nicht nur bei Problemen mit der Kuh und der LC.

Desi Rudolf, Dominik Maurer und Stefan Löb ein riesiges Dankeschön für Monate der Zellkultur, für versierte Hilfe bei Problemen mit Zellen und für Spaß in allen Situationen.

Patricia Hrstić, Claudia Falkenburger, Beate Pömmerl und Franziska Löwenstein ein herzliches Dankeschön unter anderem für Unmengen an Peptiden, grammweise Antikörper, riesiger Hilfe beim Kampf mit DNA sowie für Millionen sauberer Pipettenspitzen.

Lynne Yakes und Gerhard Hörr danke ich für ihre Hilfe und die zuverlässige Organisation der Abteilung.

Allen restlichen Mitgliedern der Abteilung Immunologie – von den Steinles über die Deckers und Jungs bis zu Céciles – vielen Dank für die Atmosphäre, Rat und Tat bei vielen meiner Fragen und natürlich für das ein oder andere Reagenz, das dringend gebraucht wurde.

Herzlichen Dank auch an zahlreiche Kooperationspartner erfolgreiche Zusammenarbeit bei unterschiedlichsten Projekten, v.a. an Prof. Dr. Karin Klingel, Dr. Nicole Schneiderhan-Marra, Prof. Dr. Dorothee Wernet, Dr. Steve Pascolo, Prof. Dr. Oliver Planz, Prof. Dr. Hans-Jörg Schild und Prof. Dr. Arnulf Stenzl.

Besonders möchte ich mich bei meinen Eltern und meinem Bruder bedanken für die Unterstützung und den Rückhalt, die mir für alles das nötige Vertrauen gegeben haben.

Ohne Zsuzsa wäre alles viel schwerer gewesen. Es hätte sehr viel gefehlt im Alltag so wie an besonderen Momenten.

12. Publications

1 **Weinzierl,A.O.**, Rudolf,D., Hillen,N., Tenzer,S., van Endert,P., Schild,H., Rammensee,H.-G. and Stevanović,S. Features of TAP-independent MHC class I ligands revealed by quantitative mass spectrometry. *submitted*.

2 **Weinzierl,A.O.***, Maurer,D.*, Altenberend,F., Schneiderhan-Marra,N., Klingel,K., Schoor,O., Wernet,D., Joos,T., Rammensee,H.-G. and Stefan Stevanović. A cryptic VEGF T cell epitope: Identification and characterization by mass spectrometry and T cell assays. *Cancer Res.*, accepted.

3 Bourteele,S., Oesterle,K., **Weinzierl,A.**, Riemann,M., Schmid,R.M. and Planz,O. Prion protein induced inhibition of NF-kB activity leads to mitochondrial apoptosis in the brain. *Cell. Microbiol.*, 9(9):2202-17 (2007).

4 **Weinzierl,A.O.**, Lemmel,C., Schoor,O., Müller,M., Krüger,T., Wernet,D., Hennenlotter,J., Stenzl,A., Klingel,K., Rammensee,H.-G. and Stevanović,S. Distorted relation between mRNA copy number and corresponding HLA ligand density on cell surface. *Mol. Cell. Proteomics*, 6(1), 102-113 (2007).

5 Szalay,G., Sauter,M., Hald,J., **Weinzierl,A.**, Kandolf,R., and Klingel,K. Sustained Nitric Oxide Synthesis Contributes to Immunopathology in Ongoing Myocarditis Attributable to Interleukin-10 Disorders. *Am. J. Pathol.*, 196 (6):2085:2093 (2006)

6 **Weinzierl,A.O.** and Stevanović,S. LC-MS-based protein and peptide quantification using stable isotope labels: from ICAT in general to differential N-terminal coding (dNIC) in particular. *BGER*, 23, 21-39 (2006).

7 Larsen,M.V., Nielsen,M., **Weinzierl,A.** and Lund,O. TAP-Independent MHC Class I Presentation. *Curr. Immunol. Rev.*, 2:3, 233-245 (2006).

8 Knights,A.J., **Weinzierl,A.O.**, Flad,T., Guinn,B., Deeg,M., Mueller,L., Mufti,G., Stevanović,S. and Pawelec,G. A novel MHC-associated Proteinase 3 peptide isolated from primary CML cells further supports the significance of this antigen for immunotherapy of myeloid leukaemias. *Leukemia*, 20(6):1067-72 (2006).

9 Geiger,C., Regn,S., **Weinzierl,A.**, Noessner,E. and Schendel,D.J. A generic RNA-pulsed dendritic cell vaccine strategy for renal cell carcinoma. *J. Transl. Med.*, 3:29 (2005).

10 Lipps,G., **Weinzierl,A.O.**, von Scheven,G., Buchen,C. and Cramer,P. Structure of a bifunctional DNA primase-polymerase. *Nat. Struct. Mol Biol.*, 11, 157-162 (2004).

*authors contributed equally

13. Academic Teachers

Prof. Albert, Prof. Bardele, Dr. Bayer, Prof. Bisswanger, Prof. Bock, Prof. Bohley, Prof. Breyer-Pfaff, Dr. Buchmann, Prof. Buckl, Prof. Cramer, Prof. Dringen, Prof. Duszenko, Prof. Dwek, Prof. Eisele, Prof. Fischer, Prof. Fröhlich, Prof. Gauglitz, Dr. Grossmann, Dr. Günzl, Prof. Häfelinger, Prof. Hamprecht, Prof. Hanack, Prof. Hofschneider, Dr. Iglauer, Prof. Jäger, Prof. Gundram Jung, Prof. Günther Jung, Dr. Kalbacher, Prof. Kandolf, Prof. Kapurniotu, PD Dr. Klein, Prof. Klingel, Prof. Krammer, Prof. Lindner, Prof. Madeo, Prof. Maier, PD Dr. Maier, Prof. Mayer, Prof. Mecke, Prof. Müller, PD Dr. Münzel, Prof. Neubert, Prof. Neupert, Prof. Ninnemann, PD Dr. Nössner, Prof. Oberhammer, PD Dr. Pawelec, Prof. Pommer, Prof. Probst, Prof. Rammensee, Prof. Reutter, Prof. Rieß, Dr. Sarrazin, Prof. Schendel, Prof. Schild, Prof. Schwarz, Prof. Selzer, PD Dr. Steinle, Prof. Stevanović, PD Dr. Stoeva, Prof. Strähle, PD Dr. Verleysdonk, Prof. Voelter, Prof. Wagner, Prof. Weber, Prof. Werringloer, Prof. Weser, Prof. Wohlleben, Prof. Zimmermann, Prof. Zeller

14. Curriculum Vitae

Name:	Andreas Oliver Weinzierl
Date of birth:	07/06/1979
Place of birth:	Wangen im Allgäu
10/2004 - 01/2008	PhD thesis at the Institute for Cell Biology, Department of Immunology, University of Tübingen. Supervision by Prof. Dr. Stefan Stevanović Title: New approaches in tumour immunology: Implications from mass spectrometric analyses of HLA-ligands.
10/2004	Diploma in Biochemistry
03/2004 - 09/2004	Diploma thesis at the Institute for Cell Biology, Department of Immunology, University of Tübingen. Supervision by Prof. Dr. Stefan Stevanović Title: Quantitative comparison of transcriptome and HLA ligandome in tumours.
10/2002 - 09/2003	Facultative studies in Biochemistry, Ludwig Maximilians University München
04/1999 - 10/2004	Studies in Biochemistry, University of Tübingen
07/1998 - 05/1999	Military service, Stetten am kalten Markt
06/1998	Abitur
07/1989 - 06/1998	Rupert-Ness-Gymnasium, Wangen im Allgäu (grammar school)
07/1986 - 06/1989	Berger-Höhe Grundschule, Wangen im Allgäu (elementary school)

15. Lebenslauf

Name:	Andreas Oliver Weinzierl
Geburtsdatum:	07.06.1979
Geburtsort:	Wangen im Allgäu
10/2004 - 01/2008	Doktorarbeit am Interfakultären Institut für Zellbiologie, Abteilung für Immunologie, Universität Tübingen. Betreuung durch Prof. Dr. Stefan Stevanović Titel: Neue Ansätze in der Tumorimmunologie: Implikationen aus der massenspektrometrischen Analyse von HLA-Liganden.
10/2004	Diplom in Biochemie
03/2004 - 09/2004	Diplomarbeit am Interfakultären Institut für Zellbiologie, Abteilung für Immunologie, Universität Tübingen. Betreuung durch Prof. Dr. Stefan Stevanović Titel: Quantitativer Vergleich von Transkriptom und HLA-Ligandom in Tumoren.
10/2002 - 09/2003	Fakultatives Studienjahr, Ludwig Maximilians Universität München
04/1999 - 10/2004	Biochemie Studium, Universität Tübingen
07/1998 - 05/1999	Wehrdienst, Stetten am kalten Markt
06/1998	Abitur
07/1989 - 06/1998	Rupert-Ness-Gymnasium, Wangen im Allgäu
07/1986 - 06/1989	Berger-Höhe Grundschule, Wangen im Allgäu

A. Supplementary Tables

Table A.1.: Peptides identified in RCC099 for which MS/MS spectra were compared to MS/MS spectra of synthetic peptides For 16 randomly chosen peptides from RCC099 synthetic peptides were produced. Synthetic peptides were modified using the dNIC strategy and peptide fragmentation spectra were recorded. Fragmentation spectra of synthetic peptides were compared to fragmentation spectra of the naturally occurring peptide. Peptides which were also used in MASCOT searches of Supplementary Tab. A.3 are marked in bold. A comparison of the fragmentation spectra for each peptide can be found at http://www.uni-tuebingen.de/uni/kxi/PaperSupplements/Weinzierl_MCP2006/.

Sequence	m/z	charge state	MS/MS spectra natural and synthetic peptide match
ALASHLIEA	517.28	2	yes
ALLNIKVKL	602.89	2	yes
ALSDHHIYL	589.30	2	yes
FLLDKKIGV	613.37	2	yes
GRVFIIKSY	**617.36**	2	**yes**
ILMEHIHKL	**642.86**	2	**yes**
LLAAWTARA	**541.31**	2	**yes**
LLFDRPMHV	618.83	2	yes
QLIDKVWQL	**647.37**	2	**yes**
QLVDIIEKV	**604.35**	2	**yes**
RLASYLDRV	601.33	2	yes
RLLDYVVNI	**607.35**	2	**yes**
SLLDIIEKV	**590.85**	2	**yes**
SLLDKIIGA	**540.82**	2	**yes**
SLSEKTVLL	570.83	2	yes
YLLPAIVHI	574.34	2	yes

Table A.2.: Presentation ratios and mRNA expression ratios were calculated between tumour and autologous normal tissue. # indicates that no corresponding MS/MS spectrum was recorded. Gene Symbol and Title can be found at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene. Fragmentation spectra for each peptide can be found at http://www.uni-tuebingen.de/uni/kxi/PaperSupplements/Weinzierl_MCP2006/.

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source	m/z H4-NIC	m/z D4-NIC	charge state
SEPT2	septin 2	RLYPWGVVEV	0.33	0.32	A02	RCC099	661.86	663.93	2
ACVRL1	activin A receptor type II-L 1	SPRKGLLML	-0.65	3.01	B07	RCC100	581.36	583.37	2
ADFP	adip.diff.-related prot.	VRLGSLSTK	4.22	-0.08	B27	RCC099	554.38	556.38	2
ADFP	adip.diff.-related prot.	SLLTSSKGQLQK	2.32	1.71	A03	RCC100	#	741.96	2
ADFP	adip.diff.-related prot.	VRLGSLSTK	-0.81	3.60	B27	RCC110	554.37	556.37	2
ADPRHL2	ADP-ribosylhydrolase L 2	EGLARPLPL	-2.10	0.61	n/a	RCC100	535.83	#	2
AKR1C1	aldo-keto reductase family 1C1	RPELVRPAL	1.97	0.29	B07	RCC100	#	580.37	2
ALDH1L1	aldeh.dehyd. 1L1	RVKTVTFEY	4.04	-0.38	B57	RCC099	#	647.38	2
ALDOA	aldolase A	ALSDHHIYL	3.38	2.72	A02	RCC099	587.31	589.31	2
ALDOA	aldolase A	RTVPPAVTGITF	2.63	2.72	B57	RCC099	#	684.44	2
ALDOA	aldolase A	ALSDHHIYL	2.11	1.01	A02	RCC100	#	589.31	2
ALDOA	aldolase A	ALSDHHIYL	2.23	0.60	A02	RCC110	587.28	589.38	2
ANKRD25	ankyrin repeat domain 25	RLLDYVVNI	-0.06	-0.08	A02	RCC099	#	607.37	2
ANXA4	annexin A4	DEVKFLTV	3.52	1.41	B18	RCC100	549.39	551.38	2
ANXA4	annexin A4	DEVKFLTV	1.95	2.20	B18	RCC110	#	551.37	2
APOL1	apolipoprot. L1	ALADGVQKV	0.78	1.52	A02	RCC099	524.29	526.31	2
APOL1	apolipoprot. L1	ALADGVQKV	0.20	1.90	A02	RCC110	526.38	524.36	2
APP	amyloid beta precursor prot.	LLAAWTARA	0.93	-0.38	A02	RCC099	#	541.35	2
ARF1	ADP-ribosylation factor 1	KLGEIVTTI	-0.62	-0.21	A02	RCC099	560.84	562.88	2
ASCC3L1	activ.signal coint. 1-3-L 1	DEHLITFF	0.88	0.71	B18	RCC100	#	565.87	2
ASCC3L1	activ.signal coint. 1-3-L 1	DEHLITFF	0.23	0.20	B18	RCC110	563.77	#	2
ATOX1	ATX1 antiox.prot. 1 homol.	RVLNKLGGVK	-1.01	n/a	A03	RCC100	636.92	638.92	2
ATP1B1	ATPase, Na+/K+ transp. b1	GRTGGSWFK	-0.07	-0.93	B27	RCC099	#	573.87	2
AUP1	ancient ubiquitous prot. 1	TLAQRVKEV	-1.00	-0.08	A02	RCC099	595.85	#	2
BSG	basigin	RVRSHLAALW	0.58	0.12	B57	RCC099	#	659.41	2
C10orf10	chrom. 10 ORF 10	RPSSVLRTL	2.96	3.01	B07	RCC100	#	569.36	2
C1QG	complement component 1q	RSGVKVVTF	0.96	2.02	B57	RCC099	#	572.37	2
C2orf24	chrom. 2 ORF 24	GRWRGWYTY	0.49	0.72	B27	RCC099	675.37	677.33	2
C2orf24	chrom. 2 ORF 24	GRWRGWYTY	0.42	1.10	B27	RCC110	#	677.38	2
CALD1	caldesmon 1	DEAAFLERL	-0.52	1.20	B18	RCC110	584.80	#	2

continued on next page

Table A.2 – contiuned

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source	m/z H4-NIC	m/z D4-NIC	charge state
CCDC21	coiled-coil dom.cont. 21	RLQMEQMQL	-1.14	n/a	A02	RCC100	641.35	643.38	2
CCNI	cyclin I	LLDRFLATV	-0.93	-0.18	A02	RCC099	576.87	578.83	2
CCNI	cyclin I	SLLDRFLATV	0.01	-0.18	A02	RCC099	620.34	622.34	2
CCT2	chap.cont.TCP1, subun.2	FLLDKKIGV	0.41	0.42	A02	RCC099	611.38	613.37	2
CCT3	chap.cont.TCP1, subun.3	ITTKAISRW	1.22	0.52	B57	RCC099	611.85	613.86	2
CD24	CD24 antigen	RAMVARLGL	5.32	0.71	A02	RCC100	#	548.43	2
CD59	CD59 antigen	SLSEKTVLL	-1.04	-0.78	A02	RCC099	568.85	570.85	2
CD59	CD59 antigen	SLSEKTVLL	-0.32	-0.44	A02	RCC100	#	570.91	2
CD59	CD59 antigen	SLSEKTVLL	-1.17	-0.20	A02	RCC110	#	570.90	2
CIB1	Ca2+integrin bind. 1	FLTKQEILL	2.55	0.42	A02	RCC099	626.37	628.44	2
CIRBP	cold induc.RNA bind.prot.	RSRGFGFVTF	0.13	-0.46	B57	RCC099	#	641.84	2
CLIC1/4	Cl- intracell.chan. 1/4	NLLPKLHIV	-1.34	0.65	A02	RCC099	597.44	599.40	2
CLIC1/4	Cl- intracell.chan. 1/4	NLLPKLHIV	1.18	0.41	A02	RCC100	597.39	599.42	2
CLIC1/4	Cl- intracell.chan. 1/4	NLLPKLHIV	-0.82	0.95	A02	RCC110	597.36	599.48	2
CLIC4	Cl- intracell.chan. 4	FLDGNEMTL	-0.02	0.17	A02	RCC099	572.76	574.83	2
CLIC5	Cl- intracell.chan. 5	NLLPKLHVV	3.00	-2.75	A02	RCC099	590.37	#	2
CLTC	clathrin, heavy pol.pep.	SRVNIPKVLR	-1.47	-0.28	B27	RCC099	664.92	666.91	2
COL18A1	collagen, type XVIII, alpha 1	RLLDVLAPL	-0.17	-1.28	A02	RCC099	557.85	559.89	2
COL6A2	collagen, type VI, alpha 2	KTRVFAVVI	0.60	-0.58	A02	RCC099	#	592.41	2
COPG	coatom.prot.complex, g	AIVDKVPSV	0.77	1.02	A02	RCC099	537.82	539.84	2
COPG	coatom.prot.complex, g	AIVDKVPSV	1.32	0.40	A02	RCC110	537.81	539.89	2
COPG/G2	coatom.prot.complex, g	DEFKVVVV	0.39	0.01	A02	RCC100	541.41	#	2
COPS7A	COP9 const.photom. 7A	ALATLIHQV	-0.21	-0.08	A02	RCC099	535.82	537.85	2
CREB3L1/2	cAMP resp.bind.3-L 1/2	KRTLIAEGY	0.09	0.30	B27	RCC110	599.34	601.41	2
CREB3L1/2	cAMP resp.bind.3-L 1/2	KRTLIAEGY	0.78	0.02	B27	RCC099	599.34	#	2
CSPG4	chondr.sulf.proteoglyc. 4	TMLARLASA	3.38	2.81	A02	RCC100	#	521.81	2
CTNNB1	catenin beta 1	TTSRVLKVL	-0.66	-0.67	A02	RCC099	582.38	584.37	2
CXCL14	chemok.(C-X-C motif)lig. 14	RLLAAALLL	2.09	0.47	A02	RCC099	#	531.90	2
CYHR1	cysteine/histidine-rich 1	HLGPEGRSV	-1.21	n/a	A02	RCC099	528.77	#	2
CYP2J2	cytochrome P450, family 2J2	KLLDEVTYL	4.78	2.32	A02	RCC099	#	622.90	2
DDX48	DEAD box pol.pep. 48	GRKGVAINF	2.48	-0.28	B27	RCC110	554.81	556.81	2
DDX5/17	DEAD box pol.pep. 5/17	YLLPAIVHI	0.55	-0.61	A02	RCC099	572.35	574.38	2
DDX5/17	DEAD box pol.pep. 5/17	YLLPAIVHI	0.18	-0.40	A02	RCC110	572.34	#	2
DHX38	DEAH box pol.pep. 38	VLFGLLREV	0.81	n/a	A02	RCC099	575.82	577.86	2
DNCL2A	dynein, cytoplasmic, light 2A	SLMHSFILK	0.57	-0.09	A03	RCC100	611.85	613.87	2

continued on next page

Table A.2 – contiuned

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source	m/z H4-NIC	m/z D4-NIC	charge state
DNM2	dynamin 2	TLIDLPGITKV	-0.62	-1.68	A02	RCC099	658.90	660.91	2
DST/MACF1	dyst.microt.actin crossl. fact. 1	SRWEKVVQR	-0.22	-0.78	B27	RCC099	667.88	669.93	2
DST/MACF1	dyst.microt.actin crossl. fact. 1	SRWEKVVQR	-0.21	0.13	B27	RCC110	667.87	#	2
EEF2	euk.transl.elongat.fact. 2	ILTDITKGV	-0.91	0.42	A02	RCC099	553.81	#	2
EEF2	euk.transl.elongat.fact. 2	RRWLPAGDAL	0.37	0.42	B27	RCC099	#	632.41	2
EEF2	euk.transl.elongat.fact. 2	ILTDITKGV	-1.27	0.50	A02	RCC110	553.82	#	2
EFHD1	EF-hand domain family D1	KLSEIDVAL	-0.77	-3.18	A02	RCC099	567.81	569.87	2
EFHD1	EF-hand domain family D1	KLSEIDVAL	0.15	-2.99	A02	RCC100	#	569.85	2
EFHD1	EF-hand domain family D1	KLSEIDVAL	0.37	-2.30	A02	RCC110	567.82	#	2
EHD2	EH-domain containing 2	ALASHLIEA	1.07	3.32	A02	RCC099	515.31	517.29	2
EHD2	EH-domain containing 2	ALASHLIEA	1.82	3.41	A02	RCC100	#	517.31	2
EHD2	EH-domain containing 2	ALASHLIEA	-0.21	1.00	A02	RCC110	515.28	#	2
EIF3S6	euk.transl.elongat.fact. 3-6	KRTTVVAQL	-0.10	0.52	B27	RCC099	#	583.88	2
EIF3S8	euk.transl.elongat.fact. 3-8	ISKQFHHQL	0.20	0.59	B57	RCC099	#	644.86	2
EIF4A2	euk.transl.elongat.fact. 4A2	GRKGVAINF	-0.71	0.00	B27	RCC099	554.81	556.85	2
EPAS1	endoth.PAS dom.prot. 1	HSMDMKFTY	-0.94	-0.68	B57	RCC099	#	655.84	2
EPAS1	endoth.PAS dom.prot. 1	HELPLPHSV	-1.18	1.01	B18	RCC100	567.32	#	2
EXOSC6	exosome component 6	ALMPVLNQV	-0.08	-0.08	A02	RCC099	545.32	547.32	2
FBXL5	F-box leu-rich rep.prot. 5	SLFEKGLKNV	-0.65	-0.18	A02	RCC099	662.35	664.38	2
FLJ14803	hypothetical prot. FLJ14803	YLFERIKEL	1.11	0.52	A02	RCC099	#	681.42	2
FLJ21439	hypothetical prot. FLJ21439	KVANIILSY	0.20	-0.26	A03	RCC100	584.36	586.38	2
FLJ32206	hypothetical prot. FLJ32206	GSHFISHLS	-0.89	n/a	n/a	RCC099	545.23	547.32	2
FLNA	filamin A	GTHKVTVLF	0.06	1.02	B57	RCC099	574.92	576.84	2
FLNA	filamin A	GTHTVSVKY	1.01	1.02	B57	RCC099	569.80	#	2
FLNA	filamin A	GVHTVHVTF	-0.24	1.02	B57	RCC099	551.33	553.32	2
FOLH1	folate hydrolase	ALFDIESKV	-1.69	-2.58	A02	RCC099	584.85	586.85	2
G3BP	Ras-GTPase-activating prot.	RRFMQTFVL	0.88	0.47	B27	RCC099	651.92	653.94	2
G3BP	Ras-GTPase-activating prot.	RRFMQTFVL	0.65	0.10	B27	RCC110	651.83	653.93	2
GAPDH	glyc.aldeh.-3-phosph.dehydr.	GRIGRLVTR	1.35	0.32	B27	RCC099	#	568.85	2
GBP4	guanylate binding prot. 4	KRLGTLVVTY	-0.60	n/a	B27	RCC099	648.88	#	2
GNB1	guan.nucleot.bind.prot. B1	HAIPLRSSW	0.15	0.31	B57	RCC099	586.33	588.37	2
GNB2L1	guan.nucleot.bind.prot. b2-L 1	KTIKLWNTL	1.06	1.02	A02	RCC099	#	655.39	2
GNB2L1	guan.nucleot.bind.prot. b2-L 1	YTDNLVRVW	0.24	1.02	B57	RCC099	#	637.85	2
GNB5	guan.nucleot.bind.prot. B5	ILFGHENRV	-3.72	-0.78	A02	RCC099	595.33	#	2

continued on next page

Table A.2 – contiuned

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source	m/z H4-NIC	m/z D4-NIC	charge state
GPX1	glutathione peroxidase 1	LLIENVASL	1.06	1.02	A02	RCC099	538.82	540.83	2
GRHPR	glyoxylate reductase	GRIGQAIAR	0.28	-0.28	B27	RCC099	523.83	525.83	2
GTF2I	gen.transcript.fact. II, i	NRIKFVIKR	-0.52	0.52	B27	RCC099	681.98	456.31	2,3
HCA112	hep.cell.carc.assoc.ag. 112	SSRLLVASW	0.66	1.02	B57	RCC099	541.79	#	2
HDAC1/2	histone deacetylase 1/2	RMLPHAPGV	-0.07	-0.38	A02	RCC099	541.79	#	2
HDAC1/2	histone deacetylase 1/2	RMLPHAPGV	-0.47	0.21	A02	RCC100	541.81	543.82	2
HDAC1/2	histone deacetylase 1/2	RMLPHAPGV	-0.35	-0.05	A02	RCC110	#	543.87	2
HECTD1	HECT domain containing 1	MPRGVVVTL	-0.13	-0.49	B07	RCC100	538.83	540.85	2
HLA-A	m.histocomp.complex I A	HRVDLGTLR	0.93	0.77	B27	RCC099	#	588.40	2
HLA-A	m.histocomp.complex I A	VMAPRTLVL	0.40	1.60	A02	RCC110	#	555.39	2
HLA-A/G	m.histocomp.complex I A/G	VMAPRTLLL	-0.23	1.60	A02	RCC110	#	561.92	2
HLA-B	m.histocomp.complex I B	AAQITQRKW	1.84	1.29	B57	RCC099	#	626.88	2
HLA-B	m.histocomp.complex I B	TAAQITQRKW	1.61	1.29	B57	RCC099	675.37	677.40	2
HLA-B/C	m.histocomp.complex I B/C	DTAAQITQR	-4.41	1.85	A68	RCC110	554.77	#	2
HLA-DPA1	m.histocomp.complex II DPA1	EEFGQAFSF	-0.34	0.66	B18	RCC100	583.77	585.97	2
HLA-DRB1/5	m.histocomp.complex II DPB1/5	LGRPDAEYW	-0.49	1.01	B57	RCC099	606.32	608.38	2
HMGB1	high-mobility group box 1	VAKKLGEMW	-1.05	0.32	B57	RCC099	#	625.84	2
HMOX1	heme oxygenase 1	KIAQKALDL	2.30	3.70	A02	RCC110	#	596.88	2
HNRPC	heterogen.nucl.rib.prot. C	SLLENLEKI	1.14	-0.19	A02	RCC100	603.50	605.47	2
HNRPC	heterogen.nucl.rib.prot. C	SLLENLEKI	0.18	0.00	A02	RCC110	603.34	605.38	2
HNRPM	heterogen.nucl.rib.prot. M	KSRGIGTVTF	-0.74	0.47	B57	RCC099	#	608.88	2
HNRPM	heterogen.nucl.rib.prot. M	LLFDRPMHV	0.43	0.47	A02	RCC099	616.85	618.87	2
HNRPUL1	heterogen.nucl.rib.prot. U-L1	VGLPAAGKTTW	0.68	0.72	B57	RCC099	624.40	626.37	2
HSPA5	heat shock 70kDa prot. 5	TRIPKIQQL	0.24	0.12	B27	RCC099	#	624.43	2
HSPG2	hep.sul.prot.glyc. 2	ALADLDELLIRA	1.56	-0.08	A02	RCC099	#	711.39	2
IFI30	IFN-g-inducib.prot. 30	LLDVPTAAV	-0.34	1.42	A02	RCC099	#	504.34	2
IFI30	IFN-g-inducib.prot. 30	LLDVPTAAV	-0.03	1.80	A02	RCC110	502.29	#	2
IL32	interleukin 32	LVHAVQALW	2.17	1.82	B57	RCC099	#	573.33	2
IMP3	U3 small nucl.ribonuc.prot.	ALLDKLYAL	1.59	0.91	A02	RCC100	#	586.42	2
ITPR1	inosit. 1,4,5-triphos.recpt.1	TLLNVIKSV	-0.53	-0.83	A02	RCC099	567.36	#	2
KIAA0152	KIAA0152 gene product	TAVALLRLL	0.85	0.12	A02	RCC099	#	539.87	2
KIAA0738	KIAA0738 gene product	KLGSVPVTV	-0.20	0.32	A02	RCC099	523.83	525.83	2
KIAA1305	KIAA1305 gene product	TLADIIARL	-0.10	n/a	A02	RCC099	545.85	547.87	2
KLHL24	kelch-L 24	VLLGKVYVV	-0.11	0.57	A02	RCC099	568.88	#	2
KPNA1	karyopherin alpha 1	VMDSKIVQV	-0.51	-0.46	A02	RCC099	583.37	585.37	2

continued on next page

Table A.2 – contiuned

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source	m/z H4-NIC	m/z D4-NIC	charge state
KRT18	keratin 18	ALLNIKVKL	2.06	1.22	A02	RCC099	#	602.93	2
KRT18	keratin 18	RLASYLDRV	-0.14	1.22	A02	RCC099	599.34	601.34	2
KRT18	keratin 18	RLASYLDRV	1.00	0.91	A02	RCC100	#	601.45	2
KRT18	keratin 18	RLASYLDRV	-0.51	-0.20	A02	RCC110	599.34	601.41	2
KRT8	keratin 8	KLSELEAAL	0.56	0.62	A02	RCC099	#	562.82	2
LMNA	lamin A/C	KAGQVVTIW	0.29	0.07	B57	RCC099	574.83	576.90	2
LMNA	lamin A/C	MRARMQQQL	-0.60	0.07	B27	RCC099	633.85	635.85	2
LMO4	LIM domain only 4	KIADRFLLY	-0.86	-0.49	A03	RCC100	643.37	645.37	2
LSP1	lymphocyte-specific prot. 1	KLIDRTESL	0.39	1.90	A02	RCC110	611.35	#	2
LYPLA2	lysophospholipase II	FLEKLLPPV	-0.10	-0.08	A02	RCC099	601.86	603.87	2
MACF1	microt.-actin crossl fact.1	SRWEKVVQR	-0.22	-0.58	B27	RCC099	667.88	669.93	2
MAF	v-maf fibrosarc.oncogene	ALISNSHQL	0.40	2.80	A02	RCC110	#	546.37	2
MAT1A/2A	met.adenosyltransf. Ia/Iia	RRVLVQVSY	0.48	-0.48	B27	RCC099	#	614.92	2
MAT1A/2A	met.adenosyltransf. Ia/Iia	RRVLVQVSY	4.08	0.50	B27	RCC110	612.85	614.96	2
MYH10	myosin, heavy pol.pep. 10	YLLEKSRAV	-0.57	-1.38	A02	RCC099	613.33	#	2
MYL6	myosin, light pol.pep. 6	FVRHILSG	-3.45	0.12	n/a	RCC099	517.29	#	2
MYL6	myosin, light pol.pep. 6	DEMNVKVL	0.28	-0.80	B18	RCC110	547.77	549.88	2
MYO1C	Myosin IC	FLDHVRTSF	-5.69	0.36	A03	RCC100	627.89	#	2
NAT8	N-acetyltransferase 8	ILDTGTIQL	1.65	-1.28	A02	RCC099	539.81	541.82	2
NCSTN	nicastrin	RRSSIQSTF	0.09	0.22	B27	RCC099	593.82	595.88	2
NCSTN	nicastrin	RRSSIQSTF	0.99	0.70	B27	RCC110	595.90	#	2
NDRG1	N-myc downstr.regul.gene 1	KLDPTKTTL	2.75	0.62	A02	RCC099	603.36	605.36	2
NDRG1	N-myc downstr.regul.gene 1	KLDPTKTTL	3.03	1.90	A02	RCC110	603.33	605.41	2
NEDD9	neur.prec.cell devel.downreg. 9	ALQEMVHQV	-0.49	0.42	A02	RCC099	580.30	582.32	2
NPL	N-acet.neuram.pyruv.lyase	KLDQVIIHV	-0.74	-0.48	A02	RCC099	606.38	608.38	2
NYREN18	NEDD8 ultimate buster-1	RSKIAETF	0.29	0.92	B57	RCC099	549.80	#	2
ODC1	ornithine decarboxylase 1	ILDQKINEV	-0.18	0.32	A02	RCC099	609.84	611.85	2
OGG1	8-oxoguan.DNA glycosyl.	VLADQVWTL	-3.07	0.17	A02	RCC099	575.35	#	2
ORMDL3	ORM1-L 3	VLIPKLPQL	-0.07	-0.98	A02	RCC099	584.37	586.40	2
PDCD6IP	progr.cell death 6 interac. prot.	SLFGGSVKL	-0.07	-0.78	A02	RCC099	527.81	#	2
PFN1	profilin 1	RTKSTGGAPTF	-0.38	0.92	B57	RCC099	635.38	637.40	2
PGCP	plasm.glutam.carboxypeptidase	ALASLIRSV	-0.08	-1.08	A02	RCC099	#	519.87	2
PGK1	phosphoglycerate kinase 1	VTRAKQIVW	0.50	1.27	B57	RCC099	624.36	626.40	2
PHGDHL1	phosph.glyc.dehydrogen. L1	RLFPPLRQR	0.19	0.41	A03	RCC100	644.41	646.42	2
PIGR	polymeric Ig receptor	FSVVINQLR	-4.73	1.72	A03	RCC099	590.43	#	2

continued on next page

Table A.2 – contiuned

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source	m/z H4-NIC	m/z D4-NIC	charge state
PLS3	plastin 3	KSSDIAKTF	-0.90	-0.68	B57	RCC099	#	595.40	2
PLVAP	plasmalem.vesic.assocc.prot.	KVKTLEVEI	0.50	0.52	A02	RCC099	624.36	626.40	2
PLXNB2	plexin B2	TYTDRVFFL	9.58	-0.50	C	RCC110	#	635.48	2
POLR2C	polymerase II pol.pep.C,33kDa	KLSDLQTQL	3.52	0.10	A02	RCC110	#	598.93	2
PPHLN1	periphilin 1	KTKEIEQVY	-0.37	-0.28	B57	RCC099	663.88	665.90	2
PPP1CA	prot. phosphatase 1 alpha	SIIGRLLEV	-3.83	0.02	A02	RCC099	552.85	#	2
PPP1CA	prot. phosphatase 1 alpha	SIIGRLLEV	0.47	0.35	A02	RCC110	552.83	#	2
PPP1CA/B/C	prot. phosphatase 1 alpha/b/c	KYPENFFLL	6.20	0.35	C	RCC110	#	661.40	2
PPP1R3C	prot. phosphatase 13C	GRMENLASYR	-0.78	0.92	B27	RCC099	#	653.40	2
PRKDC	prot. kinase, DNA-activated	DEFKIGELF	0.50	0.31	B18	RCC100	#	624.92	2
PSMB1	proteasome subun. b type, 1	RRFFPYYVY	0.66	0.37	B27	RCC099	708.43	710.44	2
PSMB1	proteasome subun. b type, 1	RRFFPYYVY	0.23	0.55	B27	RCC110	708.36	#	2
PSME3	proteasome acti. subun.3	QLVDIIEKV	1.43	-0.75	A02	RCC110	602.36	#	2
PSME3	proteasome acti. subun.3	QLVDIIEKV	-0.12	-0.05	A02	RCC099	#	604.38	2
PTGFRN	prostgland.F2 recept.neg.regul.	RLASRPLLL	-0.17	-0.29	A02	RCC100	572.46	574.46	2
PTRF	polym. I transcript.rel.fact.	SLLDKIIGA	-0.07	0.82	A02	RCC099	538.86	540.86	2
PTRF	polym. I transcript.rel.fact.	DEVKLPAKL	-0.04	0.81	B18	RCC100	601.35	603.39	2
PTRF	polym. I transcript.rel.fact.	SLLDKIIGA	0.99	0.81	A02	RCC100	538.84	540.85	2
PTRF	polym. I transcript.rel.fact.	DEVKLPAKL	-0.18	0.90	B18	RCC110	601.44	603.47	2
PTRF	polym. I transcript.rel.fact.	SLLDKIIGA	-0.54	0.90	A02	RCC110	#	540.90	2
RAMP2	recept.activ.modif.prot. 2	RPAALRLLL	-1.45	0.56	B07	RCC100	564.37	566.40	2
RASL11A	RAS-L, family 11, member A	YLLPKDIKL	-0.37	n/a	A02	RCC099	646.45	648.46	2
RBBP4/7	retinoblastoma bind.prot. 4/7	HTAKISDFSW	-0.13	0.14	B57	RCC099	#	671.85	2
RBBP4/7	retinoblastoma bind.prot. 4/7	TAKISDFSW	0.43	0.14	B57	RCC099	601.34	603.39	2
RBM4/30	RNA binding motif prot. 4/30	KLHGVNINV	-0.07	-0.73	A02	RCC099	570.89	572.88	2
RBM4/30	RNA binding motif prot. 4/30	KLHGVNINV	-0.10	-0.03	A02	RCC110	570.82	572.87	2
RBPMS	RNA bind.prot.mult.splicing	TSKQPVGF	0.63	0.42	B57	RCC099	#	507.81	2
RCN1	reticulocalbin 1	RRLGLALGL	1.71	0.42	B27	RCC099	#	539.38	2
RCN1	reticulocalbin 1	RRLGLALGL	0.73	0.70	B27	RCC110	537.33	539.39	2
RPA2	replication prot. A2, 32kDa	GRAPISNPGM	0.66	0.42	B27	RCC099	#	554.82	2
RPL15	ribosomal prot. L15	EVILIDPFHK	-4.51	-0.05	A68	RCC110	679.37	#	2
RPL19	ribosomal prot. L19	ILMEHIHKL	0.45	0.92	A02	RCC099	640.89	642.89	2
RPL19	ribosomal prot. L19	ILMEHIHKL	1.07	0.41	A02	RCC100	#	642.89	2
RPL19	ribosomal prot. L19	ILMEHIHKL	-0.34	0.70	A02	RCC110	640.86	642.97	2
RPL8	ribosomal prot. L8	GRIDKPILK	0.39	1.02	B27	RCC099	614.86	616.91	2

continued on next page

Table A.2 – contiuned

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source	m/z H4-NIC	m/z D4-NIC	charge state
RPL8	ribosomal prot. L8	GRIDKPILK	0.85	0.70	B27	RCC110	614.87	616.98	2
RPS11	Ribosomal prot. S11	GRILSGVVTK	0.31	0.12	B27	RCC099	588.88	590.89	2
RPS16	ribosomal prot. S16	ISKALVAYY	1.12	0.82	B57	RCC099	587.82	589.83	2
RPS16	ribosomal prot. S16	KLLEPVLLL	1.09	0.82	A02	RCC099	#	594.94	2
RPS25	ribosomal prot. S25	HRAQVIYTR	0.11	0.52	B27	RCC099	#	626.91	2
RPS4X	ribosomal prot. S4	GRIGVITNR	0.05	0.02	B27	RCC099	545.86	547.87	2
RPS4X	ribosomal prot. S4	GRIGVITNR	-0.90	0.30	B27	RCC110	545.81	#	2
RPS8	ribosomal prot. S8	KTRIIDVVY	0.81	1.02	B57	RCC099	627.35	629.39	2
RSL1D1	ribosomal L1 dom.cont. 1	HIIENIVAV	0.47	1.30	A02	RCC110	#	558.90	2
SBLF	stoned B-L factor	SVAGKIHTV	-2.19	-0.28	A02	RCC099	529.80	531.87	2
SCAP2	src assoc.phosphoprot. 2	KRGDVIYIL	-0.05	0.16	n/a	RCC100	612.37	614.39	2
SCAP2	src assoc.phosphoprot. 2	KRGDVIYIL	-0.79	-0.90	B27	RCC110	612.36	614.47	2
SCD	stearoyl-CoA desaturase	ARLPLRLFL	3.96	1.52	B27	RCC099	#	604.46	2
SCD	stearoyl-CoA desaturase	ITAGAHRLW	1.50	1.52	B57	RCC099	#	567.31	2
SCOC	short coiled-coil prot.	KTRLINQVL	-0.51	-0.58	B27	RCC099	#	618.43	2
SDCBP	syndecan binding prot.	LMDHTIPEV	0.00	0.92	A02	RCC099	580.36	582.32	2
SEC14L1	SEC14-L 1	QLIDKVWQL	-0.24	-0.53	A02	RCC099	645.40	647.41	2
SEC14L1	SEC14-L 1	QLIDKVWQL	-1.25	-0.20	A02	RCC110	645.37	647.43	2
SF3B3	splic.fact.3b, subun.3	GRVLIGVGK	-0.29	0.62	B27	RCC099	523.32	#	2
SKP1A	S-phase kin.assoc.prot.1A	AILKKVIQW	0.68	0.02	A02	RCC099	#	646.45	2
SLC17A3	solute carrier family 17-3	ARYGIALVL	4.86	0.62	B27	RCC099	#	542.89	2
SNRPG	small nucl.ribon.prot. G	IMLEALERV	0.32	0.80	A02	RCC110	589.83	#	2
SPTBN1	spectrin b, non-erythroc. 1	DEMKVLVL	-0.87	0.86	B18	RCC100	547.30	549.33	2
SPTBN1	spectrin b, non-erythroc. 1	DEMKVLVL	-0.69	0.30	B18	RCC110	547.30	549.38	2
SRXN1	sulfiredoxin 1 homolog	TLSDLRVYL	0.43	0.32	A02	RCC099	#	594.85	2
SSR1	signal sequence receptor	VLFRGGPRGLLAVA	0.89	0.37	A02	RCC099	765.99	767.98	2
SSR1	signal sequence receptor	VLFRGGPRGSLAVA	-1.35	0.37	A02	RCC099	752.93	754.97	2
SSR1	signal sequence receptor	VLFRGGPRGLLAVA	-2.00	0.70	A02	RCC110	765.95	#	2
SSR1	signal sequence receptor	VLFRGGPRGSLAVA	-2.42	0.70	A02	RCC110	752.94	#	2
SSR2	signal sequence receptor b	DVSHTVVLR	0.50	0.60	A68	RCC110	565.80	#	2
STAT3	s.transd.activ.transcript.3	EERIVELF	-0.25	-0.09	B18	RCC100	570.32	572.34	2
STAT3	s.transd.activ.transcript.3	EERIVELF	0.10	-0.80	B18	RCC110	570.30	#	2
STAT3	s.transd.activ.transcript.3	EELQQKVSY	0.38	-0.80	B18	RCC110	635.81	#	2
SYNCRIP	syn.tagm.bind.RNA interact.prot.	RLFVGSIPK	-0.34	0.31	A03	RCC100	582.35	584.39	2
TAGLN	transgelin	TGYGRPRQI	-0.44	0.02	n/a	RCC099	576.79	#	2

continued on next page

Table A.2 – contiuned

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source	m/z H4-NIC	m/z D4-NIC	charge state
TCF3	transcription factor 3	RVRDINEAF	-0.42	-0.25	B57	RCC099	#	614.83	2
TCN2	transcobalamin II	MRHLGAFLF	0.89	0.12	B27	RCC099	#	600.85	2
TEGT	testis enhanced gene transcript	KRLGLLAGF	0.49	-0.10	B27	RCC110	561.34	563.43	2
TLN1	talin 1	LLDHVLLTL	0.82	0.02	A02	RCC099	571.33	573.36	2
TMED10	transmemb.traff.prot. 10	FLLGPRLVLA	-0.91	-0.08	A02	RCC099	602.41	604.38	2
TMED10	transmemb.traff.prot. 10	LLGPRLVLA	1.49	-0.08	A02	RCC099	528.86	530.87	2
TMED10	transmemb.traff.prot. 10	LLGPRLVLA	-1.21	-0.40	A02	RCC110	528.84	#	2
TMEM16F	transmembrane prot. 16F	VLDDKLVFV	-0.05	0.02	A02	RCC099	#	599.86	2
TMEM16F	transmembrane prot. 16F	VLDDKLVFV	0.28	0.61	A02	RCC100	#	599.91	2
TMEM16F	transmembrane prot. 16F	VLDDKLVFV	0.38	0.40	A02	RCC110	597.84	#	2
TMEM41B	transmembrane prot. 41B	HLINYIIFL	-0.10	-1.23	A02	RCC099	#	628.35	2
TMEM66	transmembrane prot. 66	KGWDGYDVQW	0.39	0.72	B57	RCC099	700.83	702.86	2
TMEM66	transmembrane prot. 66	RRLDPIPQL	-3.55	0.72	B27	RCC099	606.87	608.94	2
TMEM66	transmembrane prot. 66	RRLDPIPQL	1.35	0.40	B27	RCC110	606.86	608.95	2
TNIK	TRAF2-NCK interact.kinase	ITKDVVLQW	0.16	0.10	B57	RCC099	#	626.93	2
TNS1	tensin 1	HAKVLEFGW	0.81	0.32	B57	RCC099	617.40	619.34	2
TNS1	tensin 1	FLIETGPRGV	0.45	0.32	A02	RCC099	#	599.40	2
TRIM22	tripartite motif-containing 22	HLANIVERV	0.55	1.60	A02	RCC110	#	580.34	2
TRRAP	transcript.dom.assoc.prot.	TLADLVHHV	0.08	0.47	A02	RCC099	555.32	557.30	2
TRRAP	transcript.dom.assoc.prot.	TLADLVHHV	0.48	-0.04	A02	RCC100	555.35	557.33	2
TRRAP	transcript.dom.assoc.prot.	TLADLVHHV	-1.05	-0.05	A02	RCC110	555.30	#	2
TSC2	tuberous sclerosis 2	SLLDIIEKV	0.13	-1.08	A02	RCC099	#	591.40	2
TXNIP	thioredoxin interacting prot.	KIKSFEVVF	-0.89	0.05	A03	RCC099	643.37	645.40	2
UBB	ubiquitin B	QIFVKTLTGK	1.04	0.16	A03	RCC100	662.42	664.46	2
UGT1A6	UDP glucuron.transfer.1-A6	ALGKIPQTV	-4.55	1.12	A02	RCC099	537.32	#	2
UMOD	uromodulin	RAFSSLGLLK	-5.49	-7.19	A03	RCC100	619.85	#	2
UNC84B	unc-84 homolog B	RIRPTAVTL	-0.58	0.11	A02	RCC100	566.36	568.37	2
USMG5	upregul.skelet.musc.grow.5	YQFTGIKKY	-0.69	-0.28	B27	RCC099	668.89	670.94	2
USP9X	ubiquit.specif.peptidase 9	RLWGEPVNL	-0.23	-0.10	A02	RCC110	#	596.88	2
UTRN	utrophin	KLLDPEDVAVQL	-0.66	-0.48	A02	RCC099	743.93	746.00	2
VIM	vimentin	ARLDLERKV	1.13	1.82	B27	RCC099	623.93	625.93	2
VIM	vimentin	NLAEDIMRL	0.27	1.82	A02	RCC099	#	592.29	2
VIM	vimentin	NYIDKVRFL	1.02	1.82	C	RCC099	657.88	659.86	2
VIM	vimentin	LERKVESL	1.61	1.90	B18	RCC110	560.91	562.89	2
WDR78	WD repeat domain 78	TSVVYDVAW	0.32	0.52	B57	RCC099	572.83	574.83	2

continued on next page

Table A.2 – contiuned

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source	m/z H4-NIC	m/z D4-NIC	charge state
XPO1	exportin 1	VLIDYQRNV	0.09	0.32	A02	RCC099	#	614.86	2
YTHDF1/2/3	YTH dom.fam. 1/2/3	GRVFIIKSY	0.91	-0.01	B27	RCC099	615.38	617.38	2
YTHDF1/2/3	YTH dom.fam. 1/2/3	GRVFIIKSY	-0.45	0.20	B27	RCC110	#	617.45	2
ZA20D2	zinc finger A20 dom.cont.2	IRKENPVVV	0.62	-0.08	B27	RCC099	600.86	#	2
ZFR	zinc finger RNA binding prot.	DEGKVIRF	0.23	-0.07	B18	RCC110	#	557.87	2

Table A.3.: MASCOT search results for 50 randomly chosen peptides identified in RCC099. With 50 randomly chosen sequence tags derived from HLA peptides found in RCC099 MASCOT searches were performed against a database containing the EBI IPI database in a normal and a reversed format. Peptides with only one MASCOT hit are marked in bold, peptides resulting in MASCOT hits with a significant score (>63) are highlighted in bold and italic.

Sequence	$[M-H]^+$ calc.	MASCOT searched sequence tag	No. hits IPI	No. hits rIPI	max. MASCOT score IPI	max. MASCOT score rIPI
RLLAAALLL	1062.78	1062.90 seq(n-R[IL][IL]AAA[IL][IL][IL])	**1**	0	***90***	n.a.
LVHAVQALW	1145.72	1145.76 seq(c-AVQALW)	**1**	0	***78***	n.a.
KLHGVNINV	1144.73	1144.86 seq(n-K) tag(530.41,GVN[IL],913.64)	1	1	***77***	***67***
LLDRFLATV	1156.75	1156.76 seq(n-[IL][IL]DRF[IL]A)	**1**	0	***72***	n.a.
KLDQVIIHV	1215.80	1215.86 seq(n-K) tag(508.4,QV[IL][IL],961.78)	**1**	0	***72***	n.a.
HAKVLEFGW	1237.72	1237.78 seq(n-H) seq(c-EFGW)	**1**	0	***70***	n.a.
SLLDKIIGA	1080.72	1080.82 seq(n-S[IL][IL]D) tag(708.5,[IL][IL],934.66)	**1**	0	***69***	n.a.
LGRPDAEYW	1215.65	1215.76 seq(n-[IL]) seq(c-AEYW)	**1**	0	***68***	n.a.
RLLDYVVNI	1213.77	1213.84 seq(n-R) seq(c-YVVN[IL])	**1**	0	***67***	n.a.
KLLDEVTYL	1244.76	1244.90 seq(n-K) tag(621.48,EVT,950.7) comp(*[Y])	**1**	0	***67***	n.a.
FLEKLLPPV	1206.80	1206.92 seq(n-F[IL]EK[IL][IL])	**1**	0	***65***	n.a.
ARYGIALVL	1084.73	1084.88 seq(n-A) seq(c-[IL]A[IL]V[IL])	**1**	0	63	n.a.
KLSELEAAL	1124.71	1124.74 seq(n-K) seq(c-[IL]EAA[IL])	**1**	0	63	n.a.
GRVFIIKSY	1233.79	1233.86 seq(n-GRVF[IL][IL])	**1**	0	63	n.a.
KTRLINQVL	1235.83	1235.86 seq(n-K) seq(c-NQV[IL])	**1**	0	63	n.a.
TAVALLRLL	1078.77	1078.84 seq(n-TAVA[IL][IL])	2	0	62	n.a.
RSGVKVVTF	1143.74	1143.84 seq(n-R) seq(c-VVTF)	**1**	0	62	n.a.
TLADIIARL	1094.73	1094.84 seq(n-T[IL]AD[IL][IL])	**1**	0	61	n.a.
ILMEHIHKL	1284.80	1284.78 seq(n-[IL][IL]MEH)	**1**	0	61	n.a.
RLLDVLAPL	1118.77	1118.88 seq(n-R[IL][IL]DV[IL])	2	0	60	n.a.
KAGQVVTIW	1152.73	1152.90 seq(c-VVT[IL]W)/1152.9 seq(c-VVT[IL][ADGEVS][ADGEVS])	**1**	0	60	n.a.
FLLGPRLVLA	1207.83	1207.86 seq(n-F[IL][IL]) seq(c-[IL]V[ILAPS][ILAPS])	**1**	0	60	n.a.
HLINYIIFL	1254.80	1254.80 seq(n-H[IL][IL]NY)	**1**	0	59	n.a.
KTKEIEQVY	1330.79	1330.90 seq(c-[IL]EQVY)	1	1	59	60
RRLGLALGL	1077.76	1077.86 seq(n-R) seq(c-[IL]A[IL]G[IL])	**1**	0	58	n.a.
RRSSIQSTF	1190.70	1190.86 seq(c-[IL]QSTF)	**1**	0	58	n.a.
SLLDIIEKV	1180.77	1180.90 seq(n-S[IL][IL]D[IL][IL])	2	0	57	n.a.
ARLPLRLFL	1207.84	1208.02 seq(n-AR) seq(c-LFL)	**1**	0	57	n.a.
LLGPRLVLA	1060.76	1060.84 seq(n-[IL][IL]) seq(c-[IL]V[IL]A)	**1**	0	56	n.a.
LLAAWTARA	1081.69	1081.80 seq(n-[IL][IL]AAW)	2	0	55	n.a.

continued on next page

Table A.3 – contiuned

Sequence	$[M-H]^+$ calc.	MASCOT searched sequence tag	No. hits IPI	No. hits rIPI	max. MASCOT score IPI	max. MASCOT score rIPI
ALFDIESKV	1172.71	1172.80 seq(n-A[IL]FD) seq(c-V)	**1**	0	55	n.a.
TLSDLRVYL	1188.74	1188.80 seq(n-T[IL]SD[IL])	3	0	55	n.a.
ILTDITKGV	1106.70	1106.72 seq(n-[IL][IL]TD) seq(c-V)	2	2	55	50
VLFGLLREV	1154.77	1154.82 seq(n-V[IL]FG[IL][IL])	**1**	0	54	n.a.
QLVDIIEKV	1207.78	1207.86 seq(n-Q[IL]VD[IL])	2	0	54	n.a.
NLAEDIMRL	1183.69	1183.68 seq(n-N[IL]AED)	1	1	54	51
GRIGQAIAR	1050.69	1050.76 seq(n-G) seq(c-A[IL]AR)	1	1	53	46
ILFGHENRV	1189.69	1189.66 seq(n-[IL][IL]FG) seq(c-V)	**1**	0	52	n.a.
ITKDVVLQW	1252.78	1252.96 seq(c-V[IL]QW)	2	0	52	n.a.
ALADLDELLIRA	1421.88	1421.88 seq(n-A[IL]AD[IL]D)	**1**	0	52	n.a.
ALADGVQKV	1051.67	1051.72 seq(n-A[IL]ADG)	**1**	0	51	n.a.
VLDDKLVFV	1198.76	1198.82 seq(n-V[IL]) tag(722.5,[IL]V,934.66) comp(*[F])	1	2	51	54
VGLPAAGKTTW	1251.76	1251.84 seq(n-VG[IL]PA)	**1**	0	50	n.a.
LLDHVLLTL	1145.77	1145.82 seq(c-V[IL][IL]T[IL])	3	0	47	n.a.
KIKSFEVVF	1289.81	1289.90 seq(c-EVVF)	**1**	0	47	n.a.
QLIDKVWQL	1293.81	1293.92 seq(n-Q[IL][IL]D) comp(*[W])	**1**	0	47	n.a.
FLIETGPRGV	1197.74	1197.90 seq(n-F[IL][IL]ET)	**1**	0	46	n.a.
ISKALVAYY	1178.73	1178.84 seq(n-[IL]SKA)	2	1	46	45
ITAGAHRLW	1133.70	1133.72 seq(n-[IL]TAG)	7	2	46	44
KRTTVVAQL	1166.78	1166.86 seq(c-VAQ[IL])	4	2	44	40

Table A.4.: HLA ligands identified in more than one RCC patient. 31 HLA ligands were identified in two tumour and normal tissue pairs, 11 ligands in all analysed samples. Thus in total 64 HLA ligand comparisons were drawn. Gene Symbol and Title can be found at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene.

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source
ADFP	adip.diff.-related prot.	VRLGSLSTK	−0.81	3.60	B27	RCC110
ADFP	adip.diff.-related prot.	VRLGSLSTK	4.22	−0.08	B27	RCC099
ALDOA	aldolase A	ALSDHHIYL	2.11	1.01	A02	RCC100
ALDOA	aldolase A	ALSDHHIYL	2.23	0.60	A02	RCC110
ALDOA	aldolase A	ALSDHHIYL	3.38	2.72	A02	RCC099
ANXA4	annexin A4	DEVKFLTV	3.52	1.41	B18	RCC100
ANXA4	annexin A4	DEVKFLTV	1.95	2.20	B18	RCC110
APOL1	apolipoprot. L1	ALADGVQKV	0.20	1.90	A02	RCC110
APOL1	apolipoprot. L1	ALADGVQKV	0.78	1.52	A02	RCC099
ASCC3L1	activ.signal coint. 1-3-L 1	DEHLITFF	0.88	0.71	B18	RCC100
ASCC3L1	activ.signal coint. 1-3-L 1	DEHLITFF	0.23	0.20	B18	RCC110
C2orf24	chrom. 2 ORF 24	GRWRGWYTY	0.42	1.10	B27	RCC110
C2orf24	chrom. 2 ORF 24	GRWRGWYTY	0.49	0.72	B27	RCC099
CD59	CD59 antigen	SLSEKTVLL	−0.32	−0.44	A02	RCC100
CD59	CD59 antigen	SLSEKTVLL	−1.17	−0.20	A02	RCC110
CD59	CD59 antigen	SLSEKTVLL	−1.04	−0.78	A02	RCC099
CLIC1/4	Cl- intracell.chan. 1/4	NLLPKLHIV	1.18	0.41	A02	RCC100
CLIC1/4	Cl- intracell.chan. 1/4	NLLPKLHIV	−0.82	0.95	A02	RCC110
CLIC1/4	Cl- intracell.chan. 1/4	NLLPKLHIV	−1.34	0.65	A02	RCC099
COPG	coatom.prot.complex, g	AIVDKVPSV	1.32	0.40	A02	RCC110
COPG	coatom.prot.complex, g	AIVDKVPSV	0.77	1.02	A02	RCC099
CREB3L1	cAMP respons.element binding 3-L 1	KRTLIAEGY	0.09	0.30	B27	RCC110
CREB3L1	cAMP respons.element binding 3-L 1	KRTLIAEGY	0.78	0.02	B27	RCC099
DDX5/17	DEAD box pol.pep. 5/17	YLLPAIVHI	0.55	−0.61	A02	RCC099
DDX5/17	DEAD box pol.pep. 5/17	YLLPAIVHI	0.18	−0.40	A02	RCC110
DST/MACF1	dyst.microt.actin crossl. fact. 1	SRWEKVVQR	−0.21	0.13	B27	RCC110
DST/MACF1	dyst.microt.actin crossl. fact. 1	SRWEKVVQR	−0.22	−0.78	B27	RCC099
EEF2	eukaryotic translation elongat. factor 2	ILTDITKGV	−1.27	0.50	A02	RCC110
EEF2	eukaryotic translation elongat. factor 2	ILTDITKGV	−0.91	0.42	A02	RCC099
EFHD1	EF-hand domain family D1	KLSEIDVAL	0.15	−2.99	A02	RCC100
EFHD1	EF-hand domain family D1	KLSEIDVAL	0.37	−2.30	A02	RCC110
EFHD1	EF-hand domain family D1	KLSEIDVAL	−0.77	−3.18	A02	RCC099

continued on next page

Table A.4 – contiuned

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source
EHD2	EH-domain containing 2	ALASHLIEA	1.82	3.41	A02	RCC100
EHD2	EH-domain containing 2	ALASHLIEA	−0.21	1.00	A02	RCC110
EHD2	EH-domain containing 2	ALASHLIEA	1.07	3.32	A02	RCC099
G3BP	Ras-GTPase-activating prot.	RRFMQTFVL	0.65	0.10	B27	RCC110
G3BP	Ras-GTPase-activating prot.	RRFMQTFVL	0.88	0.47	B27	RCC099
HDAC1/2	histone deacetylase 1/2	RMLPHAPGV	−0.47	0.21	A02	RCC100
HDAC1/2	histone deacetylase 1/2	RMLPHAPGV	−0.35	−0.05	A02	RCC110
HDAC1/2	histone deacetylase 1/2	RMLPHAPGV	−0.07	−0.38	A02	RCC099
HNRPC	heterogen. nuclear ribonucleoprot. C	SLLENLEKI	1.14	−0.19	A02	RCC100
HNRPC	heterogen. nuclear ribonucleoprot. C	SLLENLEKI	0.18	0.00	A02	RCC110
IFI30	IFN-g-inducib.prot. 30	LLDVPTAAV	−0.03	1.80	A02	RCC110
IFI30	IFN-g-inducib.prot. 30	LLDVPTAAV	−0.34	1.42	A02	RCC099
KRT18	keratin 18	RLASYLDRV	1.00	0.91	A02	RCC100
KRT18	keratin 18	RLASYLDRV	−0.51	−0.20	A02	RCC110
KRT18	keratin 18	RLASYLDRV	−0.14	1.22	A02	RCC099
MAT1A/2A	met.adenosyltransf. Ia/Iia	RRVLVQVSY	4.08	0.50	B27	RCC110
MAT1A/2A	met.adenosyltransf. Ia/Iia	RRVLVQVSY	0.48	−0.48	B27	RCC099
NCSTN	nicastrin	RRSSIQSTF	0.99	0.70	B27	RCC110
NCSTN	nicastrin	RRSSIQSTF	0.09	0.22	B27	RCC099
NDRG1	N-myc downstr.regul.gene 1	KLDPTKTTL	3.03	1.90	A02	RCC110
NDRG1	N-myc downstr.regul.gene 1	KLDPTKTTL	2.75	0.62	A02	RCC099
PPP1CA	prot. phosphatase 1 alpha	SIIGRLLEV	0.47	0.35	A02	RCC110
PPP1CA	prot. phosphatase 1 alpha	SIIGRLLEV	−3.83	0.02	A02	RCC099
PSMB1	proteasome subun. b type, 1	RRFFPYYVY	0.23	0.55	B27	RCC110
PSMB1	proteasome subun. b type, 1	RRFFPYYVY	0.66	0.37	B27	RCC099
PSME3	proteasome acti. subun.3	QLVDIIEKV	−0.12	−0.05	A02	RCC099
PSME3	proteasome acti. subun.3	QLVDIIEKV	1.43	−0.75	A02	RCC110
PTRF	polym. I transcript.rel.fact.	DEVKLPAKL	−0.04	0.81	B18	RCC100
PTRF	polym. I transcript.rel.fact.	DEVKLPAKL	−0.18	0.90	B18	RCC110
PTRF	polym. I transcript.rel.fact.	SLLDKIIGA	0.99	0.81	A02	RCC100
PTRF	polym. I transcript.rel.fact.	SLLDKIIGA	−0.54	0.90	A02	RCC110
PTRF	polym. I transcript.rel.fact.	SLLDKIIGA	−0.07	0.82	A02	RCC099
RBM4	RNA binding motif prot. 4	KLHGVNINV	−0.10	−0.03	A02	RCC110
RBM4	RNA binding motif prot. 4	KLHGVNINV	−0.07	−0.73	A02	RCC099
RCN1	reticulocalbin 1	RRLGLALGL	0.73	0.70	B27	RCC110
RCN1	reticulocalbin 1	RRLGLALGL	1.71	0.42	B27	RCC099

continued on next page

Table A.4 – contiuned

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	HLA restrict.	Source
RPL19	ribosomal prot. L19	ILMEHIHKL	1.07	0.41	A02	RCC100
RPL19	ribosomal prot. L19	ILMEHIHKL	−0.34	0.70	A02	RCC110
RPL19	ribosomal prot. L19	ILMEHIHKL	0.45	0.92	A02	RCC099
RPL8	ribosomal prot. L8	GRIDKPILK	0.85	0.70	B27	RCC110
RPL8	ribosomal prot. L8	GRIDKPILK	0.39	1.02	B27	RCC099
RPS4X	ribosomal prot. S4	GRIGVITNR	−0.90	0.30	B27	RCC110
RPS4X	ribosomal prot. S4	GRIGVITNR	0.05	0.02	B27	RCC099
SCAP2	src assoc.phosphoprot. 2	KRGDVIYIL	−0.05	0.16	n/a	RCC100
SCAP2	src assoc.phosphoprot. 2	KRGDVIYIL	−0.79	−0.90	B27	RCC110
SEC14L1	SEC14-L 1	QLIDKVWQL	−1.25	−0.20	A02	RCC110
SEC14L1	SEC14-L 1	QLIDKVWQL	−0.24	−0.53	A02	RCC099
SPTBN1	spectrin b, non-erythroc. 1	DEMKVLVL	−0.87	0.86	B18	RCC100
SPTBN1	spectrin b, non-erythroc. 1	DEMKVLVL	−0.69	0.30	B18	RCC110
SSR1	signal sequence receptor	VLFRGGPRGLLAVA	−2.00	0.70	A02	RCC110
SSR1	signal sequence receptor	VLFRGGPRGLLAVA	0.89	0.37	A02	RCC099
SSR1	signal sequence receptor	VLFRGGPRGSLAVA	−2.42	0.70	A02	RCC110
SSR1	signal sequence receptor	VLFRGGPRGSLAVA	−1.35	0.37	A02	RCC099
STAT3	s.transd.activ.transcript.3	EERIVELF	−0.25	−0.09	B18	RCC100
STAT3	s.transd.activ.transcript.3	EERIVELF	0.10	−0.80	B18	RCC110
TMED10	transmemb.traff.prot. 10	LLGPRLVLA	−1.21	−0.40	A02	RCC110
TMED10	transmemb.traff.prot. 10	LLGPRLVLA	1.49	−0.08	A02	RCC099
TMEM16F	transmembrane prot. 16F	VLDDKLVFV	0.28	0.61	A02	RCC100
TMEM16F	transmembrane prot. 16F	VLDDKLVFV	0.38	0.40	A02	RCC110
TMEM16F	transmembrane prot. 16F	VLDDKLVFV	−0.05	0.02	A02	RCC099
TMEM66	transmembrane prot. 66	RRLDPIPQL	1.35	0.40	B27	RCC110
TMEM66	transmembrane prot. 66	RRLDPIPQL	−3.55	0.72	B27	RCC099
TRRAP	transcript.dom.assoc.prot.	TLADLVHHV	0.48	−0.04	A02	RCC100
TRRAP	transcript.dom.assoc.prot.	TLADLVHHV	−1.05	−0.05	A02	RCC110
TRRAP	transcript.dom.assoc.prot.	TLADLVHHV	0.08	0.47	A02	RCC099
YTHDF1/2/3	YTH dom.fam. 1/2/3	GRVFIIKSY	−0.45	0.20	B27	RCC110
YTHDF1/2/3	YTH dom.fam. 1/2/3	GRVFIIKSY	0.91	−0.01	B27	RCC099

Table A.5.: The majority of HLA ligands significantly over-presented on tumour tissue derived from tumour associations antigens. Out of the 30 redundant source proteins (Figure 4, area I-III) 19 were described to be tumour associated. 10 HLA ligands were found uniquely presented on tumour tissue (marked in bold). 8 of their source proteins were reported to be tumour associated.

Gene Symbol	Gene Title	Sequence	$_2log$ HLA ligand ratio	$_2log$ mRNA ratio	Source	Literature
ADFP	adip.diff.-related prot.	SLLTSSKGQLQK	2.32	1.71	RCC100	[1]
ADFP	adip.diff.-related prot.	**VRLGSLSTK**	4.22	−0.08	RCC099	[1]
AKR1C1	aldo-keto reductase family 1C1	RPELVRPAL	1.97	0.29	RCC100	[2]
ALDH1L1	aldeh.dehyd. 1L1	RVKTVTFEY	4.04	−0.38	RCC099	[3]
ALDOA	aldolase A	ALSDHHIYL	2.11	1.01	RCC100	-
ALDOA	aldolase A	ALSDHHIYL	2.23	0.60	RCC110	-
ALDOA	aldolase A	RTVPPAVTGITF	2.63	2.72	RCC099	-
ALDOA	aldolase A	ALSDHHIYL	3.38	2.72	RCC099	-
ANXA4	annexin A4	DEVKFLTV	1.95	2.20	RCC110	[4]
ANXA4	annexin A4	**DEVKFLTV**	3.52	1.41	RCC100	[4]
C10orf10	chrom. 10 ORF 10	RPSSVLRTL	2.96	3.01	RCC100	-
CD24	CD24 antigen	RAMVARLGL	5.32	0.71	RCC100	[5]
CIB1	Ca2+integrin bind. 1	FLTKQEILL	2.55	0.42	RCC099	[6]
CLIC5	Cl- intracell.chan. 5	NLLPKLHVV	3.00	−2.75	RCC099	-
CSPG4	chondr.sulf.proteoglyc. 4	TMLARLASA	3.38	2.81	RCC100	[7]
CXCL14	chemok.(C-X-C motif)lig. 14	**RLLAAALLL**	2.09	0.47	RCC099	[8]
CYP2J2	cytochrome P450, family 2J2	KLLDEVTYL	4.78	2.32	RCC099	-
DDX48	DEAD box pol.pep. 48	GRKGVAINF	2.48	−0.28	RCC110	[9]
EHD2	EH-domain containing 2	ALASHLIEA	1.82	3.41	RCC100	-
HLA-B	m.histocomp.complex I B	TAAQITQRKW	1.61	1.29	RCC099	-
HLA-B	m.histocomp.complex I B	AAQITQRKW	1.84	1.29	RCC099	-
HMOX1	heme oxygenase 1	**KIAQKALDL**	2.30	3.70	RCC110	[10]
HSPG2	hep.sul.prot.glyc. 2	ALADLDELLIRA	1.56	−0.08	RCC099	[11]
IL32	interleukin 32	LVHAVQALW	2.17	1.82	RCC099	-
IMP3	U3 small nucl.ribonuc.prot.	ALLDKLYAL	1.59	0.91	RCC100	-
KRT18	keratin 18	ALLNIKVKL	2.06	1.22	RCC099	[12]
MAT1A/2A	met.adenosyltransf. Ia/Iia	**RRVLVQVSY**	4.08	0.50	RCC110	-
NAT8	N-acetyltransferase 8	ILDTGTIQL	1.65	−1.28	RCC099	[13]
NDRG1	N-myc downstr.regul.gene 1	KLDPTKTTL	2.75	0.62	RCC099	[14]
NDRG1	N-myc downstr.regul.gene 1	KLDPTKTTL	3.03	1.90	RCC110	[14]
PLXNB2	plexin B2	**TYTDRVFFL**	9.58	−0.50	RCC110	[15]
POLR2C	polymerase II pol.pep.C,33kDa	**KLSDLQTQL**	3.52	0.10	RCC110	[16]
PPP1CA/B/C	prot. phosphatase 1 alpha/b/c	**KYPENFFLL**	6.20	0.35	RCC110	[17]
RCN1	reticulocalbin 1	RRLGLALGL	1.71	0.42	RCC099	-
SCD	stearoyl-CoA desaturase	ITAGAHRLW	1.50	1.52	RCC099	[18]
SCD	stearoyl-CoA desaturase	**ARLPLRLFL**	3.96	1.52	RCC099	[18]
SLC17A3	solute carrier family 17-3	**ARYGIALVL**	4.86	0.62	RCC099	-
VIM	vimentin	LERKVESL	1.61	1.90	RCC110	[19]

[1] S. M. Schmidt, K. Schag, M. R. Muller, T. Weinschenk, S. Appel, O. Schoor, M. M. Weck, F. Grunebach, L. Kanz, S. Stevanovic, H. G. Rammensee, and P. Brossart. Induction of adipophilin-specific cytotoxic T lymphocytes using a novel HLA-A2-binding peptide that mediates tumor cell lysis. *Cancer Res.*, 64(3):1164–1170, 2004.

[2] C. Y. Chen, C. P. Hsu, N. Y. Hsu, C. S. Shih, T. Y. Lin, and K. C. Chow.

Expression of dihydrodiol dehydrogenase in the resected stage I non-small cell lung cancer. *Oncol. Rep.*, 9(3):515–519, 2002.

[3] S. A. Krupenko and N. V. Oleinik. 10-formyltetrahydrofolate dehydrogenase, one of the major folate enzymes, is down-regulated in tumor tissues and possesses suppressor effects on cancer cells. *Cell Growth Differ.*, 13(5):227–236, 2002.

[4] U. Zimmermann, S. Balabanov, J. Giebel, S. Teller, H. Junker, D. Schmoll, C. Protzel, C. Scharf, B. Kleist, and R. Walther. Increased expression and altered location of annexin IV in renal clear cell carcinoma: a possible role in tumour dissemination. *Cancer Lett.*, 209(1):111–118, 2004.

[5] S. Schindelmann, J. Windisch, R. Grundmann, R. Kreienberg, R. Zeillinger, and H. Deissler. Expression profiling of mammary carcinoma cell lines: correlation of in vitro invasiveness with expression of CD24. *Tumour. Biol.*, 23 (3):139–145, 2002.

[6] H. Lee, E. C. Yi, B. Wen, T. P. Reily, L. Pohl, S. Nelson, R. Aebersold, and D. R. Goodlett. Optimization of reversed-phase microcapillary liquid chromatography for quantitative proteomics. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, 803(1):101–110, 2004.

[7] J. Yang, M. A. Price, C. L. Neudauer, C. Wilson, S. Ferrone, H. Xia, J. Iida, M. A. Simpson, and J. B. McCarthy. Melanoma chondroitin sulfate proteoglycan enhances FAK and ERK activation by distinct mechanisms. *J. Cell Biol.*, 165(6):881–891, 2004.

[8] S. R. Schwarze, J. Luo, W. B. Isaacs, and D. F. Jarrard. Modulation of CXCL14 (BRAK) expression in prostate cancer. *Prostate*, 64(1):67–74, 2005.

[9] Q. Xia, X. T. Kong, G. A. Zhang, X. J. Hou, H. Qiang, and R. Q. Zhong. Proteomics-based identification of DEAD-box protein 48 as a novel autoantigen, a prospective serum marker for pancreatic cancer. *Biochem. Biophys. Res. Commun.*, 330(2):526–532, 2005.

[10] P. O. Berberat, Z. Dambrauskas, A. Gulbinas, T. Giese, N. Giese, B. Kunzli, F. Autschbach, S. Meuer, M. W. Buchler, and H. Friess. Inhibition of heme oxygenase-1 increases responsiveness of pancreatic cancer cells to anticancer treatment. *Clin. Cancer Res.*, 11(10):3790–3798, 2005.

[11] E. M. Gonzalez, M. Mongiat, S. J. Slater, R. Baffa, and R. V. Iozzo. A novel interaction between perlecan protein core and progranulin: potential effects on tumor growth. *J. Biol. Chem.*, 278(40):38113–38116, 2003.

[12] U. Woelfle, G. Sauter, S. Santjer, R. Brakenhoff, and K. Pantel. Down-regulated expression of cytokeratin 18 promotes progression of human breast cancer. *Clin. Cancer Res.*, 10(8):2670–2674, 2004.

[13] C. Lilla, E. Verla-Tebit, A. Risch, B. Jager, M. Hoffmeister, H. Brenner, and J. Chang-Claude. Effect of NAT1 and NAT2 genetic polymorphisms on colorectal cancer risk associated with exposure to tobacco smoke and meat consumption. *Cancer Epidemiol. Biomarkers Prev.*, 15(1):99–107, 2006.

[14] H. Cangul, K. Salnikow, H. Yee, D. Zagzag, T. Commes, and M. Costa. Enhanced expression of a novel protein in human cancer cells: a potential aid to cancer diagnosis. *Cell Biol. Toxicol.*, 18(2):87–96, 2002.

[15] P. Conrotto, S. Corso, S. Gamberini, P. M. Comoglio, and S. Giordano. Interplay between scatter factor receptors and B plexins controls invasive growth. *Oncogene*, 23(30):5131–5137, 2004.

[16] R. J. White. RNA polymerases I and III, growth control and cancer. *Nat. Rev. Mol. Cell Biol.*, 6(1):69–78, 2005.

[17] M. Fujita, C. Seta, J. Fukuda, S. Kobayashi, and T. Haneji. Induction of apoptosis in human oral squamous carcinoma cell lines by protein phosphatase inhibitors. *Oral Oncol.*, 35(4):401–408, 1999.

[18] N. Scaglia, J. M. Caviglia, and R. A. Igal. High stearoyl-CoA desaturase protein and activity levels in simian virus 40 transformed-human lung fibroblasts. *Biochim. Biophys. Acta*, 1687(1-3):141–151, 2005.

[19] S. Singh, S. Sadacharan, S. Su, A. Belldegrun, S. Persad, and G. Singh. Overexpression of vimentin: role in the invasive phenotype in an androgen-independent model of prostate cancer. *Cancer Res.*, 63(9):2306–2311, 2003.

Table A.6.: HLA ligands identified from LCL721.174 and LCL721.45. HLA presentation ratios, protein and mRNA expression ratios were calculated between LCL721.174 and LCL721.45. Signal sequence prediction was performed with SignalP (http://www.cbs.dtu.dk/services/SignalP/). SYFPEITHI and BIMAS scores were calculated using web based software (http://www.syfpeithi.de, http://www-bimas.cit.nih.gov/molbio/hla_bind). Gene Symbol can be found at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene.

Sequence	Gene Symbol	position	signal seq.	HLA	SYFP-EITHI score	Bimas score	$_2log$ (721.174 vs. 721.45) HLA	protein	mRNA
LLSAEPVPA	CD79B	20-28	28 | 29	A*0201	18	8	9.12	n.s.d.	0.30
LLGPRLVLA	TMP21	22-31	31 | 32	A*0201	24	19	7.98	n.d.	-0.80
SLWGQPAEA	COL4A5	18-26	26 | 27	A*0201	23	41	6.96	n.a.	-0.10
VLAPRVLRA	RCN1	21-29	29 | 30	A*0201	24	19	6.91	-0.04	-0.40
ALVVQVAEA	HEXB	34-42	42 | 43	A*0201	24	11	6.86	n.d.	-0.70
HGVFLPLV	KIAA0247	21-28	39 | 40	B*5101	26	n.a.	6.53	n.d.	-0.10
LLAAWTARA	APP	9-17	17 | 18	A*0201	22	8	6.52	n.d.	0.00
VLLKARLVPA	KIAA1946	19-28	28 | 29 or 33 | 34	A*0201	23	72	5.58	n.d.	-1.10
KMDASLGNLFA	FAM3C	30-40	24 | 25	A*0201	23	6	5.19	n.d.	0.40
LLFSHVDHVIA	SLC8A1	25-35	35 | 36	A*0201	22	27	4.81	n.d.	721.174 only
FLGPWPAAS	LRPAP1	22-30	28 | 29 or 32 | 33	A*0201	16	0	4.57	n.d.	-0.40
SLYALHVKA	VKORC1	23-31	31 | 32	A*0201	16	16	4.53	n.d.	n.d.
LLLSAEPVPA	CD79B	19-28	28 | 29	A*0201	20	31	4.48	n.d.	0.30
AMAPPSHLLL	PELP1	473-482	21 | 22	A*0201	25	15	4.19	n.d.	721.174 only
MAPLALHLL	IL4I1	1-9	21 | 22	B*5101	21	n.a.	4.17	n.s.d.	n.d.
FLLGPRLVLA	TMP21	22-31	31 | 32	A*0201	27	149	4.16	n.s.d.	-0.80
LLLDVPTAAV	IFI30	15-24	26 | 27	A*0201	30	128	4.15	n.s.d.	-0.20
LLLDVPTAAVQA	IFI30	15-26	26 | 27	A*0201	20	128	3.89	n.s.d.	-0.20
LLLDVPTAA	IFI30	15-23	26 | 27	A*0201	20	128	3.88	n.s.d.	-0.20
LLDVPTAAV	IFI30	16-24	26 | 27	A*0201	28	47	3.85	n.s.d.	-0.20
VLFRGGPRGLLAVA	SSR1	19-32	20 | 21	A*0201	23	0	3.68	n.s.d.	-1.57
ALLSSLNDF	NIF3L1	5-13	no	A*0201	19	1	3.68	n.d.	0.10
IITKEVLAP	EST sequence	frame 3	n.a.	A*0201	15	0	3.68	n.d.	n.d.
QLQEGKNVIGL	TAGLN2	166-176	no	A*0201	23	18	2.98	0.25	-0.10
ILAPAGSLPKI	TERE1	328-338	no	A*0201	n.a.	n.a.	2.63	n.d.	0.60
MASRWGPLIG	Cab45	8-17	36 | 37	B*5101	21	260	2.19	n.d.	-0.10
AVLALVLAPAGA	NRP1	10-21	21 | 22	A*0201	23	19	2.11	n.d.	n.d.
LAPRVLRA	RCN1	22-29	29 | 30	A*0201	n.a.	n.a.	2.08	-0.04	-0.40

continued on next page

Table A.6 – contiuned

Sequence	Gene Symbol	position	signal seq.	HLA	SYFP-EITHI score	Bimas score	$_2log$ (721.174 vs. 721.45) HLA	protein	mRNA
AALLDVRSVP	GDF5	275-284	27 | 28	A*0201	24	2	1.94	n.d.	n.d.
SLPKKLALL	HSPC023	72-80	no	A*0201	29	49	1.70	n.d.	0.40
KAPVTKVAA	PDLIM1	241-249	no	n.a.	n.a.	n.a.	0.81	1.32	1.60
NPLPSKETI	TMSB4X	27-35	no	B*5101	26	586	0.29	-0.32	n.d.
MFPLVKSAL	COX7B	1-9	no	B*5101	18	n.a.	-0.15	n.d.	0.00
MAPRTLVL	HLA-A	4-11	24 | 25	B*5101	24	n.a.	-0.93	n.d.	-0.80
TLLGHEFVL	CDC27	605-613	no	A*0201	24	188	-1.29	n.d.	0.70
LPHVPLGVI	AUP1	374-382	yes	B*5101	26	629	-1.74	n.d.	0.50
YLTAEILEL	HIST1H2AJ	58-66	61 | 62	A*0201	28	226	-1.86	0.49	n.d.
LPREILNLI	LOC116064	259-267	no	B*5101	26	692	-1.87	n.d.	0.65
LLDRFLATV	CCNI	72-80	no	A*0201	31	413	-2.25	n.d.	0.20
GSHSMRYF	HLA-A, -B, -C, -G	25-32	24 | 25	n.a.	n.a.	n.a.	-2.34	n.d.	n.d.
HLINYIIFL	TMEM41B	240-248	no	A*0201	27	9	-2.54	n.d.	n.d.
YVPRAILV	TUBB3	59-66	no	B*5101	14	n.a.	-2.81	n.a.	n.d.
TLAEIAKVEL	NONO	120-129	no	A*0201	28	88	-2.86	0.41	0.70
RIIEETLAL	ARPC2	9-17	no	A*0201	26	12	-3.04	-0.04	-0.15
DGLVVLKI	EIF3S3	42-49	no	B*5101	27	n.a.	-3.07	n.s.d.	0.20
MAPRTLLL	HLA-C	4-11	24 | 25	B*5101	22	n.a.	-3.12	n.d.	-0.98
VMAPRTLVL	HLA-A	3-11	24 | 25	A*0201	24	n.a.	-3.23	n.d.	-0.80
IPPIQVTKV	NCOR2	973-981	no	B*5101	26	286	-3.42	n.d.	-0.20
KLWEMDNMLI	PP784	56-65	no	A*0201	25	2544	-3.49	n.d.	n.d.
YAYDGKDYLAL	HLA-A, -B, -C	140-150	24 | 25	B*5101	23	173	-3.51	n.d.	n.d.
IPPLIKSV	DSCAM	1573-1580	17 | 18	B*5101	25	n.a.	-3.52	n.d.	n.d.
QLDDLKVEL	RPL35	20-28	no	A*0201	26	35	-3.53	n.s.d.	0.90
YAYDGKDYL	HLA-G	140-148	24 | 25	B*5101	23	173	-3.79	n.d.	-1.15
IARVLTVI	RPL35	54-61	no	B*5101	25	n.a.	-3.91	n.s.d.	0.90
DALDVANKIGII	RPL23A	145-156	no	B*5101	33	484	-3.97	0.25	0.00
LPFRVLLV	WDR54	80-87	no	B*5101	30	n.a.	-3.98	n.d.	n.d.
IPYQDLPHL	MGLL	14-22	no	B*5101	24	286	-4.01	n.d.	-0.70
TIIDTKGVTAL	PAF53	174-184	no	A*0201	25	4	-4.01	n.d.	1.45
FLTKQEILL	CIB1	21-29	no	A*0201	20	98	-4.10	n.d.	0.20
VPYPLPKI	NPEPPS	254-261	no	B*5101	28	n.a.	-4.13	-0.3	0.07
SLLDRFLATV	CCNI	71-80	no	A*0201	33	9221	-4.21	n.d.	0.20
LLIDDKGTIKL	CDC2	134-144	no	A*0201	25	8	-4.30	0.34	0.55
ILHEIAVLEL	STK17B	77-86	no	A*0201	31	342	-4.35	n.d.	-0.10

continued on next page

Table A.6 – contiuned

Sequence	Gene Symbol	position	signal seq.	HLA	SYFP-EITHI score	Bimas score	$_2log$ (721.174 vs. 721.45) HLA	protein	mRNA
DALDKIRLI	TRA1	110-118	yes	B*5101	30	532	-4.35	-0.38	-0.70
HLANIVERV	TRIM22	73-81	no	A*0201	26	3	-4.39	n.d.	-0.30
ASPSSIRSL	MCM5	132-140	no	n.a.	n.a.	n.a.	-4.45	0.03	1.00
LLATLAAAML	C16orf30	177-186	25 | 26	A*0201	26	36	-4.47	n.d.	n.d.
HIIENIVAV	RSL1D1	225-233	no	A*0201	29	3	-4.48	0.36	n.d.
FLIRESETL	LYN	153-161	no	A*0201	25	48	-4.48	n.s.d.	0.00
QSPVALRPL	FLJ22875	n.a.	n.a.	B*5101	19	n.a.	-4.51	n.d.	0.20
VLPKLYV	RPS26	63-69	no	B*5101	n.a.	n.a.	-4.64	-0.05	0.90
DAKQLTTTI	NEDD9	541-549	no	B*5101	27	440	-4.66	n.d.	-2.60
YPFKPPKI	UBE2L3	61-68	no	B*5101	35	n.a.	-4.76	-0.18	0.68
MAPRSLLL	HLA-F	1-8	21 | 22	B*5101	22	n.a.	-4.78	-2.13	-1.30
LPRSTVINI	IFITM1	19-27	no	B*5101	26	572	-4.79	721.45 only	-0.80
DAYLPLRL	POLD1	515-522	no	B*5101	26	n.a.	-4.84	n.d.	1.10
DAYLVHLI	DDX49	232-239	no	B*5101	29	n.a.	-5.03	n.d.	0.80
AIVDKVPSV	COPG	147-155	no	A*0201	29	90	-5.04	n.s.d.	-0.70
YPPPEVRNI	SF3A1	44-52	no	B*5101	26	572	-5.06	-0.46	0.70
IMLEALERV	SNRPG	68-76	no	A*0201	28	1460	-5.10	n.s.d.	0.40
YPFKPPKV	UBE2E2	114-121	no	B*5101	33	n.a.	-5.11	n.d.	n.d.
KMDPIISRV	BTG2	73-81	28 | 29	A*0201	25	162	-5.14	n.d.	-0.40
DPYKVYRIV	IRF4	120-128	no	B*5101	30	1065	-5.38	0.12	0.05
DALGKLISI	GCN1L1	692-700	no	B*5101	30	484	-5.42	n.s.d.	0.30
DPYEVSYRI	BTG2	105-113	28 | 29	B*5101	33	1936	-5.65	n.d.	-0.40
ALSDHHIYL	ALDOA	216-224	no	A*0201	23	482	-5.70	0.09	-0.23
DAWKLPKI	CYBB	378-385	25 | 26 or 34 | 35	B*5101	33	n.a.	-5.72	n.d.	0.00
IANHQVLII	DHX16	412-420	no	B*5101	25	260	-5.78	n.d.	-0.80
DAFTIKTV	IFIH1	514-521	no	B*5101	29	n.a.	-5.86	n.d.	n.d.
TLFDYEVRL	UHRF1	57-65	no	A*0201	25	324	-6.07	n.d.	1.40
VLAVLGAVVAV	HLA-C	ca. >280	24 | 25	A*0201	27	52	-6.21	n.s.d.	-0.98
LAPLIQVI	HIP1	222-229	no	B*5101	29	n.a.	-6.36	n.d.	n.d.
DAFRQPSLFY	MAP3K5	158-167	no	B*5101	29	n.a.	-6.41	n.d.	0.40
RMLPHAPGV	HDAC1	371-379	no	A*0201	24	186	-6.44	-0.41	0.50
DAPHPPLKI	OAZ1	69-77	no	B*5101	29	400	-6.44	n.d.	n.d.
DALLKFSHI	TEGT	11-19	45 | 46	B*5101	29	400	-6.44	n.d.	0.10
KLLDPEDISV	SPTBN1	249-258	no	A*0201	26	3693	-6.46	n.a.	0.45
DAHIYLNHI	TYMS	254-262	no	B*5101	28	484	-6.50	721.45 only	0.20

continued on next page

Table A.6 – contiuned

Sequence	Gene Symbol	position	signal seq.	HLA	SYFP-EITHI score	Bimas score	$_2log$ (721.174 vs. 721.45) HLA	protein	mRNA
RLLEVPVML	ISOC2	49-57	no	A*0201	26	324	-6.51	n.d.	n.d.
DAAAKALRI	STAU	70-78	no	B*5101	29	400	-6.51	n.s.d.	-0.55
DALRSILTI	MARS	703-711	24 | 25	B*5101	31	532	-6.52	721.45 only	-0.80
SAPYGRITL	CYFIP2	832-840	no	B*5101	18	n.a.	-6.57	n.s.d.	-0.20
IPPHVVKV	G3BP	270-277	no	B*5101	27	n.a.	-6.60	n.a.	0.70
DGLRDLPSI	LPIN1	456-464	no	B*5101	23	176	-6.86	n.d.	-0.93
VPYLFKKV	L3MBTL2	411-418	no	B*5101	23	n.a.	-6.92	n.d.	1.20
KLIDRTESL	LSP1	135-143	no	A*0201	28	150	-6.94	n.a.	-0.60
NLLPKLHIV	CLIC1	179-187	no	A*0201	30	243	-7.02	-1.76	-1.30
IAPTGHSL	SEPT6	155-162	no	B*5101	20	n.a.	-7.02	-0.03	0.10
DGVAVLKV	HSPD1	399-406	no	B*5101	25	n.a.	-7.06	0.59	0.10
ILMEHIHKL	RPL19	137-145	no	A*0201	32	1267	-7.07	0.71	0.20
YLLPAIVHI	DDX5	148-156	no	A*0201	30	408	-7.08	-0.05	-0.20
KIADFGWSV	AURKC	181-189	no	A*0201	25	3911	-7.09	n.d.	n.d.
NPYDSVKKI	UBD	25-33	no	B*5101	32	1065	-7.11	n.d.	-3.10
KILEDVVGV	TPX2	465-473	no	A*0201	30	1168	-7.28	n.d.	1.10
LPHHRVIEV	NOMO3	1030-1038	no	B*5101	23	260	-7.29	n.d.	n.d.
FPLDPQLAKMVI	DHX15	565-576	no	A*0201	21	0	-7.46	n.s.d.	0.30
DAYVLPKLY	RPS26	60-68	no	B*5101	23	7	-7.55	-0.05	0.90
QLIDKVWQL	SEC14L1	593-601	no	A*0201	28	1511	-7.59	n.s.d.	-0.47
DGPRVFRV	FARSLA	121-128	no	B*5101	25	n.a.	-7.60	721.45 only	n.d.
LPEDKPRLI	QTRTD1	243-251	no	B*5101	26	629	-7.61	n.d.	0.50
DAMKYTIVV	ATP5A1	217-225	no	B*5101	28	220	-7.62	-1.85	0.30
KLLEPVLLL	RPS16	50-58	no	A*0201	31	2609	-7.67	n.a.	0.40
ALIEKLVEL	POLA2	22-30	no	A*0201	35	201	-7.74	0.47	0.90
ALAEIAKAEL	SFPQ	343-352	no	A*0201	30	88	-7.76	n.d.	-0.20
ALADGVQKV	APOL1	160-168	27 | 28	A*0201	33	656	-7.80	n.d.	721.45 only
LPHLLPLL	TNPO2	392-399	no	B*5101	23	n.a.	-7.84	n.d.	-0.20
LLIENVASL	GPX1	40-48	18 | 19	A*0201	34	342	-7.87	n.d.	0.30
ALSNLEVKL	URP2	326-334	no	A*0201	27	21	-7.92	-0.05	0.00
TSPKALVI	CDYL	148-155	no	B*5101	17	n.a.	-7.95	n.d.	-1.10
DAARFPII	HM13	74-81	no	B*5101	30	n.a.	-8.21	721.45 only	0.00
IPNRIPKI	KIAA1430	111-118	no	B*5101	33	n.a.	-8.28	n.d.	0.00
VPLDKQITI	GNL3	295-303	no	B*5101	28	692	-8.31	n.s.d.	n.d.
ALLDKLYAL	C15orf12	78-86	no	A*0201	32	745	-8.41	n.d.	0.50

continued on next page

Table A.6 – contiuned

Sequence	Gene Symbol	position	signal seq.	HLA	SYFP-EITHI score	Bimas score	$_2log$ (721.174 vs. 721.45) HLA	protein	mRNA
LPPEVNRI	SF3B14	13-20	no	B*5101	26	n.a.	-8.50	n.s.d.	n.d.
ALTELLAKI	CCND2	227-235	no	A*0201	30	98	-8.51	n.d.	-1.03
TGYLNTVTV	GNB2L1	192-200	no	B*5101	22	88	-8.62	n.a.	0.40
TLIDLPGITRV	MX1 or MX2	175-185/222-232	no	A*0201	24	15	-8.64	-1.13	n.d.
ALATLIHQV	COPS7A	26-34	no	A*0201	30	160	-8.66	n.d.	0.30
VLFHNLPSL	CD53	42-50	24 | 25 or 29 | 30	A*0201	29	309	-8.73	n.d.	0.10
SIIGRLLEV	PPP1CA	11-19	no	A*0201	31	22	-8.81	721.45 only	0.00
HPDSHQLFI	G3BP	336-344	no	B*5101	21	440	-8.85	n.a.	0.70
HLPETKFSEL	AIM1	942-951	no	A*0201	24	6	-8.92	n.d.	-0.60
KLHGVNINV	RBM4	59-67	no	A*0201	23	243	-8.93	-0.02	0.25
HPYRLILTV	TYK2	479-487	40 | 41	B*5101	26	484	-9.01	n.d.	-0.20
YPFHVPLL	LTA	142-149	34 | 35	B*5101	27	n.a.	-9.02	n.d.	-1.70
YAFNMKATV	HSPA8	545-553	no	B*5101	24	315	-9.09	0.2	0.30
IPYHIVNIV	SARS	357-365	no	B*5101	27	572	-9.10	-0.19	-0.70
DAYSFSRKI	CNAP1	626-634	no	B*5101	31	880	-9.12	n.d.	0.80
DANPYDSVKKI	UBD	23-33	no	B*5101	32	1065	-9.26	n.d.	-3.10
VLIPKLPQL	ORMDL3	134-142	38 | 39	A*0201	31	84	-9.44	n.d.	n.d.
QLVDIIEKV	PSME3	114-122	no	A*0201	27	321	-9.97	0.17	0.85
IPYPRPIHL	LOC57149	79-87	no	B*5101	25	260	-10.11	n.d.	0.00
LPAPIEKTI	IGHG3	402-410	19 | 20	B*5101	25	692	-10.49	0.81	0.40
DVANKIGII	RPL23A	148-156	no	B*5101	26	n.a.	n.a.	0.25	0.00
LDVPTAAVQA	IFI30	17-26	26 | 27	A*0201	10	0	n.a.	n.s.d.	-0.20
VLIDYQRNV	XPO1	784-792	no	A*0201	25	97	n.a.	n.d.	0.25
RVFENIVAV	TRAF1	195-203	no	A*0201	24	212	n.a.	721.45 only	-0.20
DAFKIWVI	SWAP70	114-121	no	B*5101	34	n.a.	n.a.	n.d.	-0.80
MAPARLFAL	SDC4	1-9	18 | 19	B*5101	20	72	n.a.	n.d.	721.45 only
LPLRLFLII	SCD	132-140	no	B*5101	26	520	n.a.	n.d.	0.23
VLIGEFLEKV	RNMT	182-191	no	A*0201	31	506	n.a.	n.d.	-0.20
SLAQYLINV	PCBP2	346-354	no	A*0201	29	160	n.a.	-0.02	n.d.
RLIPIIVLL	MYO1G	708-716	no	A*0201	30	49	n.a.	n.d.	-0.65
SLYREILFL	MYCPBP	1800-1808	no	A*0201	28	865	n.a.	n.d.	n.d.
DGLTVHLVI	UBQLN1	98-106	no	B*5101	29	176	n.a.	n.a.	0.00
QLLEKVIEL	ECHDC1	79-87	no	A*0201	31	745	n.a.	n.d.	0.05

n.d. not detected, n.s.d. no significant detection (n $<$ 3) , n.a. no definite ratio available

Table A.7.: 21-25 aa precursors of HLA ligands presented TAP independently but not located in the signal sequence were digested in vitro with constitutive proteasomes (c20s) and immunoproteasomes (i20s). The HLA ligand sequence is highlighted in bold. Proteasomal cleavage efficiency was scored from 0 to 3. Score 0: no correct C-terminus was generated; score 1: <5% peptides with the correct C-terminus; score 2: >=5% and <10% peptides with correct C-terminus; score3: >=10% peptides with correct C-terminus.

Sequence	Gene Symbol	Gene Title	HLA	HLA $_2log$ (721.174 / 721.45)	c20S score	i20s score
QRVKEV**LPHVPLGVI**QRDLAKTG	AUP1	ancient ubiquitous protein 1	B*5101	0.30	3	3
HNNLLTY**LPREILNLI**HLEELSLR	LOC116064	hypothetical protein LOC116064	B*5101	0.27	3	3
ETFALASS**LLDRFLATV**KAHPKYLS	CCNI	cyclin I	A*0201	0.21	3	3
KASS**SLPKKLALL**KAPAKKKG	HSPC023	HSPC023 protein	A*0201	1.70	2	2
NYAYAY**TLLGHEFVL**TEELDKA	CDC27	cell division cycle 27	A*0201	-1.29	1	1
NEASSHK**YVPRAILV**DLEPGTMDSV	TUBB3	tubulin, beta 3	B*5101	0.14	0	2
KLKLDVGE**AMAPPSHLLL**PVPCK	PELP1	Pro-Glu-Leu-rich protein 1	A*0201	4.19	0	1
SGFRSV**KAPVTKVAA**SIGNAQK	PDLIM1	PDZ and LIM domain 1	n.a.	0.81	0	1
QQVERHRE**HLINYIIFL**RITPFL	TMEM41B	transmembrane protein 41B	A*0201	0.17	0	0
RLETR**TLAEIAKVEL**DNMPLRGK	NONO	non-POU domain containing, octamer-binding	A*0201	0.14	0	0
DSAVKQVQI**DGLVVLKI**IKHYQEE	EIF3S3	eukar. transl. initiation factor 3, subunit 3g	B*5101	0.12	0	0
LAAVLE**YLTAEILEL**AGNAAR	HIST1H2AJ	histone 1, H2aj	A*0201	0.27	0	0
SGRQP**AALLDVRSVP**GLDGSG	GDF5	growth differentiation factor 5	A*0201	1.94	0	0